Power System Resilience under Natural Disasters

Yushi Tan

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Daniel S. Kirschen, Chair

Payman Arabshahi, Chair

Arindam K. Das

Michael R. Wagner

Program Authorized to Offer Degree:
Department of Electrical Engineering

University of Washington

Abstract

Power System Resilience under Natural Disasters

Yushi Tan

Co-Chairs of the Supervisory Committee:
Close Professor Daniel S. Kirschen
Department of Electrical Engineering

Professor Payman Arabshahi
Department of Electrical Engineering

Power systems are not likely to remain unscathed by natural disasters such as earthquake, hurricanes, ice storms, as evident from the recent Hurricane Harvey and Hurricane Irma. The outages will last days or even weeks because of the amount of damaged components. And the impacts are affecting the economies, public health and communities especially those that are already facing challenges. This motivates us to study methods of improving resilience in both operational stage and planning stage. We believe this is an interdisciplinary research from several aspects,

1. There has been no consensus on the definition of power system resilience under natural disasters. And in fact, this research direction only becomes hot in recent 4 or 5 years. However, the concept of infrastructure resilience has been prevailing and well-studied in civil engineering. After summarizing previous efforts on defining and quantifying of resilience including those adapted to power systems, we base our work on the resilient measure derived from operability trajectory and develop an equivalent measure of harm that has clearer power system meanings.

2. The knowledge of power systems guides us to focus on electricity distribution systems,

where we believe the resilience has more potential for improvement. We start with the case of fully automated radial distribution network, and then move on to partially automated radial distribution network and finally find a way to handle the uncertainties in repair time. After consulting with industry experts, we relax certain operational constraints to make the problems (slightly but enough) easier to solve without compromising their practicality in field. Built upon the operation problems, we formulate the quantification and assessment of resilience in the planning stage, which will help electric utilities decide how best to spread the budget to improve the resilience.

3. Unfortunately, none of the problems described above are easy to solve in terms of the computational complexity. In particular, the operational problems might need to be solved in real time repeatedly and MILP formulations, though straightforward, are too slow in practice. We adopt the settings of scheduling theory and propose the first of its kind, soft precedence constraints, to model the relaxed load flow equations in radial distribution networks. And for the assessment of resilience in the planning stage, we simplify the operational problem by using a single crew approximation with only a constant away from optimal. This allows us to reformulate the distribution systems hardening problem into a combinatorial optimization with the flavor of the multiple knapsack problem.

To summarize, this research aims to develop good algorithms and heuristics for problems under the framework of power system resilience adapted from the concept of infrastructure resilience.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Table Number Page

GLOSSARY

ARRA: the American Recovery and Reinvestment Act

DA: Distribution Automation

DER: Distributed Energy Resources

DG: Distributed Generation (Generator)

DOE: Department of Energy

EMS: Energy Management System

EPRI: Electric Power Research Institute

ERCOT: Electricity Reliability Council of Texas

GUROBI: a commercial optimization solver

HVAC: Heating, ventilation, and air conditioning

IEEE: Institute of Electrical and Electronics Engineers

JULIA: a high-level dynamic programming language for numerical analysis

JUMP: Julia for Mathematical Optimization

LP: Linear Programming

MILP: Mixed Integer Linear Programming

NERC: North America Electric Reliability Corporation

NIAC: National Infrastructure Advisory Council

PMU: Phase Measurement Units

RCS: Remote Controlled Switch

SAIDI: System Average Interruption Duration Index

SAIFI: System Average Interruption Frequency Index

SCADA: Supervisory control and data acquisition

SGIG: the Smart Grid Investment Grant Program

STAIDI: the Storm Average Interruption Duration Index

STAIFI: the Storm Average Interruption Frequency Index

ACKNOWLEDGMENTS

It has been a long 5 years.

It has been a short 5 years.

I still remember my attitude towards heuristics 5 years ago when I believed the MILP is the solution to everything. There is no chance that I could imagine I would present quite a few heuristic algorithms during my final exam. I still recall that how reluctant I was when my mom was trying to teach me to cook before I came to UW. Well, right now, I am capable of cooking some Chinese dishes and some "western" style dishes as well. Friends seem to enjoy most of them whenever Ruiyi and I hosted a party. These are merely 2 examples of how the Ph.D. career turned me into whom I never thought I could be. And the current me could not travel this long without the help and accompany of many.

I would like to thank my friends, Jianghui Yu, Haowen Cao, Haixiao Yu and Ying Huang, Tianxing Cheng, Jingcheng Che, Jiaying Shi, Yi Luan, Yuzong Liu and Lida Lin, Xingwei Wu, Xiaojing Zhu, Mingwei Tang, Junjie Yan, Shi Chen, Huizhong Guo, Yuankun Li, Xi Zhu, Wenruo Bai, Hongyang Sun and Xiaoyun Yang, Jiarui Li, Danqing Zhang, Haoran Cai, Di Fan, Shanchao Li, Simin Peng, Vivian Zhu, Tong Zhang, Zheng Li, Xinbo Geng, Wenting Wang, Zhongwei Wang, Jingcheng Xu, Zhe Yu, Xiaofei Zhao, Xin Hu. For those who live outside of Seattle, I really appreciate your accompany over Wechat all these years. For those who we have spent more than a few happy moments in Seattle, I really cherish the moments (espcially those in wechat) and hope we could re-unite here in the future.

I want to give special thanks to Yishen Wang, Zeyu Wang and Zhongming Ye for offering so much help when I just arrived and throughout the first four years, and for still being friends with me after living together in Argonne for 3 months. And I also must acknowledge all the members of REAL lab including but not limited to: Dr. Ting Qiu, Dr. Iker Diaz de Zerio Mendaza, Dr. Yangwu Shen, Dr. Jesus Elmer Contreras, Agustina Gonzalez, Rebecca Breiding, Dr. Yury Dvorkin, Dr. Mushfiqur Sarker, Dr. Ahlmahz Negash, Ahmad Milyani, Abeer Almaimouni, Dr. Pan Li, Dr. Bolun Xu and Minwen Weng, Dr. Yang Wang, Dr. Ricardo Fernandez-Blanco, Kelly Kozdras, Daniel Olsen, Dr. Hao Wang, Dr. Chenghui Tang, Yuanyuan Shi, Ryan Elliott, Dr. Jingkun Liu, Dr. Yao Chang, Yi Wang, Leonardo Macedo, Mareldi Ahumada.

I have spent 2 rewarding intership. It has been a great pleasure working from Aidan Touhy, Eamonn Lannoye and Robert Entriken with EPRI and Feng Qiu, Chen Chen and Jianhui Wang from Argonne National Lab. I am glad to meet Yang Wang, Zhan Qin, Yuna Zhang, Gang Huang, Yanlong Sun and re-unite wtih Shujun Xin, Runze Chen and Cheng Wang. I also want to express my condolences to Yanlong's family. Finally, the color printer and the paper with high qualities allow me to print a lot of papers and book chapters, leading to the idea in Chapter 2.6 and eventually this whole dissertation.

I am always grateful to Prof. Chongqing Kang, Prof. Mark O'Malley, Prof. Qixin Chen and Prof. Chen Shen for providing references when I applied for the graduate school and helping me come to this amazing University of Washington. I also appreciate the help from Prof. Ning Zhang for helping me communicate with Prof. Mark O'Mally and for answering my questions when I started my Ph.D. Finally, I want to thank Prof. Xuemin Zhang for providing references when I applied for the concurrent Master degree in Statistics.

I want to thank Professor Wagner for serving as my reading committee and Drton for agreeing to be the GSR and for being such a great instructor when I took STAT 559, 538 and 502. My gratitude also goes to Prof. Arabshahi and Prof. Das for the fruitful be-weekly discussion over the last 3 years and for offering your insights towards every detail within this dissertation. I also want to thank Dr. Ortega-Vazquez for giving me so much advice, for sharing your exercise route within the CSE/EE building and for being a terrific professor.

I want to express my appreciation towards my supervisor Professor Daniel S. Kirschen. He has been a great supervisor throughout the years, which is a consensus among all the lab members. He supported me applying for a concurrent MS degree in Statistics, connected with various professionals and researchers including engineers in SNOPUD and project manager in LANL and recommended me two internship opportunities with Electric Power Research Institute and Argonne National Lab.

I want to thank my parents Zhenyu Tan and Baili Shi, my parents-in-law Yanbin Li and Xiuling Li for being supportive throughout these 5 years, for not rushing me to graduate and to find a job and for not having rushed me too hard to publish papers.

Finally, there is this one person that I shared almost every single detail of these 5 years with, my wife, Ruiyi Li. I could not be more grateful and lucky to have her accompany ever since high school. I am looking forward to the brand-new post-graduation life ahead of us.

DEDICATION

to Ruiyi Li

Chapter 1

INTRODUCTION

1.1 Impact of Natural Disasters on Power Systems

Natural disasters, such as Hurricane Sandy in November 2012, the Christchurch Earthquake of February 2011 or the June 2012 Mid-Atlantic and Midwest Derecho, caused major damage to the electricity distribution networks and deprived homes and businesses of electricity for prolonged periods. Such power outages carry heavy social and economic costs. Estimates of the annual cost of power outages caused by severe weather between 2003 and 2012 range from $18 billion to $33 billion on average (Executive Office of the President 2013).

Hurricanes often cause storm surges that flood substations and corrode metal, electrical components and wiring (The City of New York 2013). Earthquake can trigger ground liquefaction that damage buried cables and dislodge transformers (Kwasinski et al. 2014). Wind and ice storms bring down trees, breaking overhead cables and utility poles (Infrastructure Security and Energy Restoration, Office of Electricity Delivery and Energy Reliability, U.S. Department of Energy 2012). As the duration of an outage increases, its economic and social costs rise exponentially. We summarize the impacts of different natural disasters on power systems in Table 1.1, modified upon the work by Breiding (2015). See also (Wang et al. 2015, Reed et al. 2009) for discussions of the impacts of natural disasters on power grids and (Vugrin et al. 2011, Reinhorn et al. 2010) for its impact on other infrastructures.

Finally, we want to take the example of Hurricane Harvey to illustrate an almost worst cases scenario for impacts on power systems. Hurricane Harvey made landfall at peak intensity on San Jose Island, just east of Rockport, with winds of 130 mph (215 km/h) at approximately 10 p.m. on August 25, 2017. The storm gradually weakened to a tropical storm by the evening of August 26, 2017. Based on the report by Electric Reliability Council

Table 1.1: Natural Disaster Impacts on Power Systems

Natural Disaster	Effect
Wind Storms	Downed transmission lines due to strong winds and or trees tripping Damaged distribution poles due to strong winds and or trees tripping
Ice Storms	Downed transmission lines due to ice loading and/or strong winds Damaged distribution poles due to ice loading and/or strong winds
Hurricanes	Downed transmission lines due to airborne debris and/or strong winds Damaged distribution poles due to airborne debris and/or strong winds
Floods	Damaged substations, transformers, and/or underground distribution lines due to water seepage
Tornados	Downed transmission lines due to ice loading and/or strong winds Damaged distribution poles due to ice loading and/or strong winds
Earthquakes	Damaged transformers due to inadequate anchorage during shaking Damaged underground lines due to ground liquefaction Damaged overhead lines due to poles shaking in opposite directions Damaged overhead lines due high weight loading of distribution poles

of Texas (2017), since the hurricane first made landfall, six 345kV transmission lines in the ERCOT system have experienced storm-related Forced Outages. Approximately 52% of the 138kV facilities and 34% of the 69kV facilities remain outaged as of the morning of August 30th. In addition to transmission outages, approximately 8,000 MW of generation is outaged and approximately 3,000 MW is derated due to storm-related causes as of the morning of August 30th. On the distribution side, we take the local utility company CenterPoint Energy 2.4 million metered customers across 5,000 square miles in and around Houston, Texas. Based on the presentation by Kenny Mercado, Senior Vice President, Electric Operations (2017), Hurricane Harvey leads to a 755 million total minutes outage over 10 days and a 308 SAIDI minutes. 8 substations were out of service and 9 substations were inaccessible due to high water. Thoughout the restoration process, 293 total electric circuits were locked out and 4,494 total electric fuses were out.

When the natural phenomenon is in the upper range or beyond what is expected, as in the case of Hurricane Harvey, power systems may not be able to survive such events relatively unscathed. Physical damage to grid components must be repaired before power can be restored (The GridWise Alliance 2013, NERC 2014). Therefore in these cases, the ability to repair the damage quickly to restore at least a basic service that helps communities return to a more normal life becomes the crucial aspect.

1.2 Defining Power System Resilience

The concept of power system resilience stems from the context of civil and industrial engineering, which has been well studied since the publication of (on Critical Infrastructure Protection 1997). As pointed out by Reed et al. (2009), our approach only focuses on engineering resilience, although more comprehensive definitions of resilience considers social, economic and environmental factors. After about 20 years, there are still various definitions of resilience. Consider, among others (Mili & Center 2011, Bruneau et al. 2003, O'Rourke 2007), the three definitions from the electrical engineering literature, government advisory report and policy directives:

- The resilience of the distribution system is based on three elements: prevention, recovery, and survivability. System recovery refers to the use of tools and techniques to quickly restore service to as many affected customers as practical. Survivability refers to the use of innovative technologies to aid consumers, communities, and institutions in continuing some level of normal function without complete access to the grid. (Electric Power Research Institute 2013)

- Infrastructure resilience is the ability to reduce the magnitude and/or duration of disruptive events. The effectiveness of a resilient infrastructure or enterprise depends upon its ability to anticipate, absorb, adapt to, and/or rapidly recover from a potentially disruptive event. And they also organized the main features into a sequence of events named by the NIAC resilience construct, which we reproduce in Figure 1.1. (National Infrastructure Advisory Council 2010)

- Resilience is the ability to anticipate, prepare for, and adapt to changing climate conditions and withstand, respond to, and recover rapidly from disruptions. (Obama 2013)

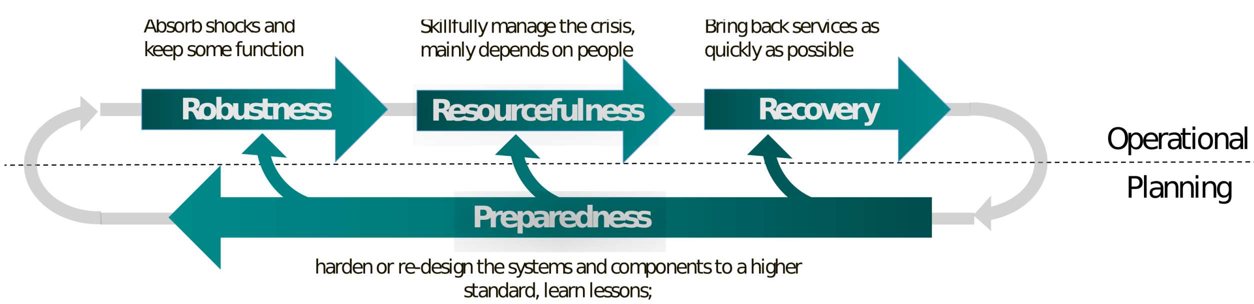


Figure 1.1: Interactions between the four aspects of resilience

Specific definitions of resilience are less important than the fundamental concepts of resilience (National Infrastructure Advisory Council 2010). Although the definitions are not exactly the same, all such definitions contain more or less the following four aspects or features of resilience in our points of view:

Preparedness refers to the application of engineering designs and advanced technologies that harden the distribution system to limit damage (Electric Power Research Institute 2013) or re-design the systems and components to a higher standard (Ton & Wang 2015);

Robustness is the ability to absorb shocks (National Infrastructure Advisory Council 2010) and aid customers in continuing some level of normal function (Electric Power Research Institute 2013);

Resourcefulness the ability to skillfully manage the crisis, including but not limited to identifying problems, establishing restoration plans, mobilizing resources and communicating decisions to the people who will implement them after the event takes place (National Infrastructure Advisory Council 2010, Bruneau et al. 2003). Resourcefulness mainly depends on people, instead of technology.

Recovery is the capacity to bring services back as quickly as possible (National Infrastructure Advisory Council 2010).

These aspects might not cover all facets of concept of resilience, but they turns out to be instructive for utilities to manage and practice and also contains most of the recent researches regarding resilience. Although the definition and the four aspects is applicable to all critical infrastructures, we feel no need to restate them in the power system context but instead we will list some corresponding utility practices and power system researches in the following sections.

Here we want to distinguish the concept of resilience with other similar concepts in power systems. Electric utilities normally declare that ensuring the reliability of their systems, i.e., satisfying the customer load requirements is the primary mission. Theoretically, reliability is defined as the probability of the system performing its function adequately, for the period of time intended, under the operation conditions intended (Prada 1999). To accommodate the characteristics and requirements of power systems, power system reliability normally refers to the related concepts, indices and evaluation techniques. The North America Electric Reliability Corporation (NERC) enforces reliability standards on utility companies. An detailed review on deterministic and probabilistic approaches for operations and planning considering reliability standards could be seen by Strbac et al. (2016). These techniques assume that only a small number of components will fail at the same time and that most part of the system should operate undisturbed. However, both aspects do not apply to the case with natural disasters. In essence, the fact that tens of hundreds of components could be destroyed and that customers will be left without power for days is the very reason that power system resilience has drawn much attention recently.

Another concept to compare is self-healing, for which there is no formal or unanimous definition. However, since its introduction into power systems (Amin 2001), self-healing capability involves two parts, monitoring and controlling unforeseen events, with an emphasis on utilizing the advanced information, sensing, control and communication technologies and without human intervention. Self-healing is related to the robustness aspect of resilience in terms of the objective of minimizing the adverse impact. But self-healing is applicable not only to natural disasters but also minor disturbances. In most current framework of self-healing grids, the adverse impact could be reduced, but not literally 'healed', since repairs and replacement of damages would require involvement of crew.

An electric power system broadly consists of two parts, transmission system and distribution system. Although the topic of this proposal is about power system resilience, we will only try to model distribution system for several reasons. Transmission systems usually span a wide area whereas natural disasters mostly happen locally, so distribution systems are more likely to be severely damaged. Transmission systems are meshed networks, while distribution systems are mostly radial to reach as many customers as possible. Since the power flows through transmission systems towards distribution systems, repairs and replacements in transmission system are prioritized and those in distribution systems are the bottleneck of restoring power to all consumers.

As we mentioned above, resilience is a general concept for any infrastructure and literatures on resilience Miles (2011), Berkeley III et al. (2010) to natural disasters emphasizes the concept of 'community resilience' to highlight the fact that the different types of infrastructures are interdependent. Post-disaster recovery plans must take all these interdependencies into account. Reed et al. (2009) list the 11 interdependent infrastructures including:

(1) Electric power delivery, with subsystems distribution, transmission, and generation;
(2) Telecommunications, with subsystems of cable, cellular, Internet, landlines, and media;
(3) Transportation, with subsystems air travel, roadways, fueling: gas stations, mass transit, rail, and water and port facilities;
(4) Utilities, with subsystems water supply, sewage treatment, sanitation, oil delivery and natural gas delivery;
(5) Building support, with subsystems HVAC, elevators, security and plumbing;
(6) Business, with subsystems computer systems, hotels, insurance, gaming, manufacturing, marine-maritime, mines, restaurants and retail;
(7) Emergency Services, with subsystems 911, ambulance, fire, police and shelters;
(8) Financial systems, with subsystems ATM, banks, credit cards and stock exchange;
(9) Food supply, with subsystems distribution, storage, preparation, and production;
(10) Government, with subsystems of offices and services;
(11) Health care, with subsystems of hospitals and public health.

And an example of electric power infrastructure dependencies (Rinaldi et al. 2001) is provided in Figure 1.2.

1.3 Quantifying Resilience

It has been argued by Ton & Wang (2015) that developing resilience metrics would help the public utility commissions and other regulators to guide decisions for policy, planning, investments and operations and to manage trade-offs. Many researches have been focusing on quantifying resilience from various aspects, including graph theory, complex network, civil

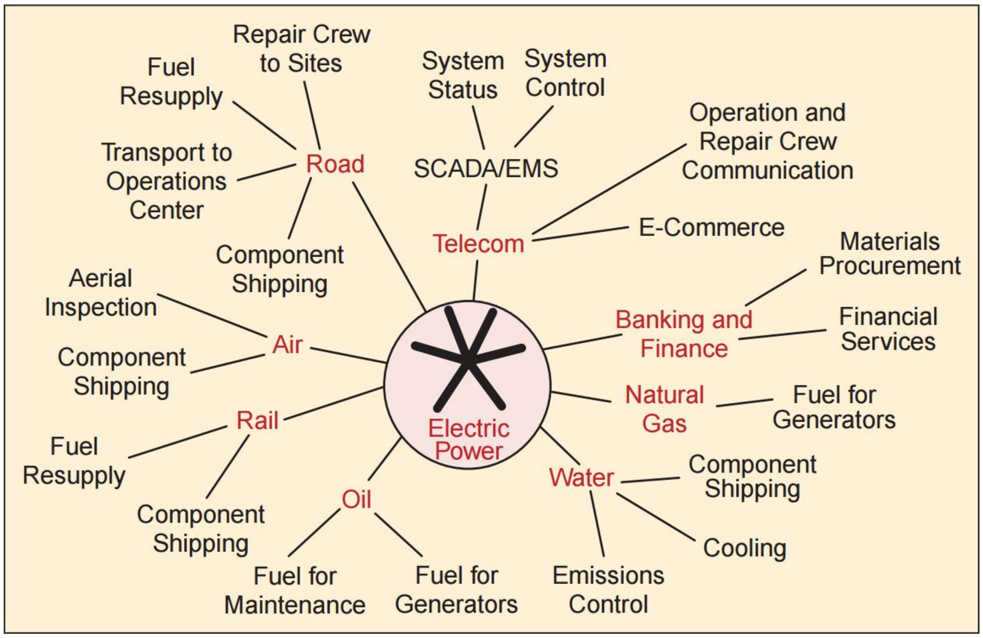


Figure 1.2: An example of electric power infrastructure dependencies

engineering and power systems. We will list some of them including the one we adapted from the civil engineering context.

Network science based metrics Some network science based resilience metrics in current literatures are related to centrality. Centrality actually measures the importance of a vertex or an edge. There are various ways to measure centrality, including but not limited to degree centrality, closeness centrality, betweenness centrality (Barthelemy 2004) and eigenvector centrality. An electrical centrality measure is proposed by Hines & Blumsack (2008) and the authors also examine the power system is a scale-free network. As indicated by Chanda & Srivastava (2016), the smaller the value of maximum centrality measure among all nodes and/or edges, the network is more resilient. In this sense, centrality is more of an indicator of robustness. Malfunction of any node or edge will not significantly lower the system performance.

Percolation based metrics Percolation theory stems from physics. The process of a section being dysfunctional is called bond percolation. It describes the behavior of connected clusters when there is a random breakdown in the network. A key metric is percolation threshold p_c, which is defined as the maximum fraction of edges removed while the network still have a giant component (Cohen & Havlin 2010). In the case of natural disasters, the average size of the small clusters is also a good indication of how many generation resources are necessary to maintain the system functionality.

Reliability Indices STAIFI, as defined by where the SAIFI value is evaluated for the duration of the extreme event (Brown et al. 1997), is proposed by Reed et al. (2010) to analyze system performance for Hurricane Katrina. SAIDI and SAIFI are defined as (*IEEE Guide for Electric Power Distribution Reliability Indices - Redline* 2012)

$$\text{SAIDI} = \frac{\sum \text{Customers minutes of interruption}}{\text{Total number of customers served}} \tag{1.1}$$

$$\text{SAIFI} = \frac{\sum \text{Total number of customer interruptions}}{\text{Total number of customers served}} \tag{1.2}$$

Trajectory-based metrics In the civil engineering context, resilience can be illustrated using the "operability trajectory", $Q(t)$, as shown in Figure 1.3, adopted from (Reed et al. 2009). This is also the so-called "resilience triangle" (Bruneau et al. 2003). The trajectory shows the increase in infrastructure functionality over time and is an effective visual indicator of the 'goodness of the restoration process'. Robustness is quantified by the depth of functionality drop at time zero (without any loss of generality, we assume that the the disaster occurs at time $t = 0$ and the restoration process commences immediately afterward), while the quality of the recovery process is quantified by the ramp up time of the operability trajectory to full/satisfactory functionality, post time zero. Obviously, we desire that an infrastructure system exhibit a relatively small drop in functionality at time zero and a quick ramp up time to full/satisfactory functionality, post time zero. Consequently, the ideal operability trajectory is defined by $Q_{ideal}(t) = 1, \forall t \geq 0$, assuming that operability is measured in fractional units instead of percentages. These two metrics can naturally be combined into an unifying measure of resilience (O'Rourke 2007). Letting T be the restoration time horizon, a resilience measure, R, can be defined as follows (Reed et al. 2009):

$$R = \int_0^T Q(t)dt, \tag{1.3}$$

The closer $Q(t)$ is to $Q_{ideal}(t)$, the greater is the area under $Q(t)$, and therefore the greater is the resilience measure. It is interesting to note that this definition of resilience is similar to the notion of 'area under (RoC) curve' (AUC), a criterion which is widely used in signal processing, communications, and machine learning.

Instead of maximizing the resilience measure defined in eqn. 5.1, we could choose to minimize the quantity $\int_0^T Q_{ideal}(t)dt - \int_0^T Q(t)dt$, which is the area over the $Q(t)$ curve, bounded from above by $Q_{ideal}(t)$. This area, informally, the 'other side of resilience', can be interpreted as a measure of 'aggregate harm' H. In a power system, it can be shown using the Lebesgue integral that minimizing this area is equivalent to minimizing the quantity $\sum_n w_n T_n$, where w_n can be interpreted as the contribution of node n to the overall loss in functionality of the system or the importance of node n and T_n is the time to restore node n. *Therefore, the objective for operational problems is to minimize the measure $\sum_n w_n T_n$, given a specific disaster scenario, while the objective for planning problems is to minimize $\sum_n w_n T_n$ in an expected sense, where the expectation is over all possible disaster scenarios.* We will elaborate the quantification with further details in Chapter 5. Note that the aggregate harm H can be seen as a weighted version of SAIDI.

Another approach is to fit the trajectory $Q(t)$ by a function $1 - e^{-bt}$ and the parameter b governs how rapidly the restoration process occurs (Reed et al. 2010). However, in some cases, the trajectory is not close to this function (see Fig. 2.4 for example in Chapter 2) and the error from the fitting itself might affect the accurateness of this measure.

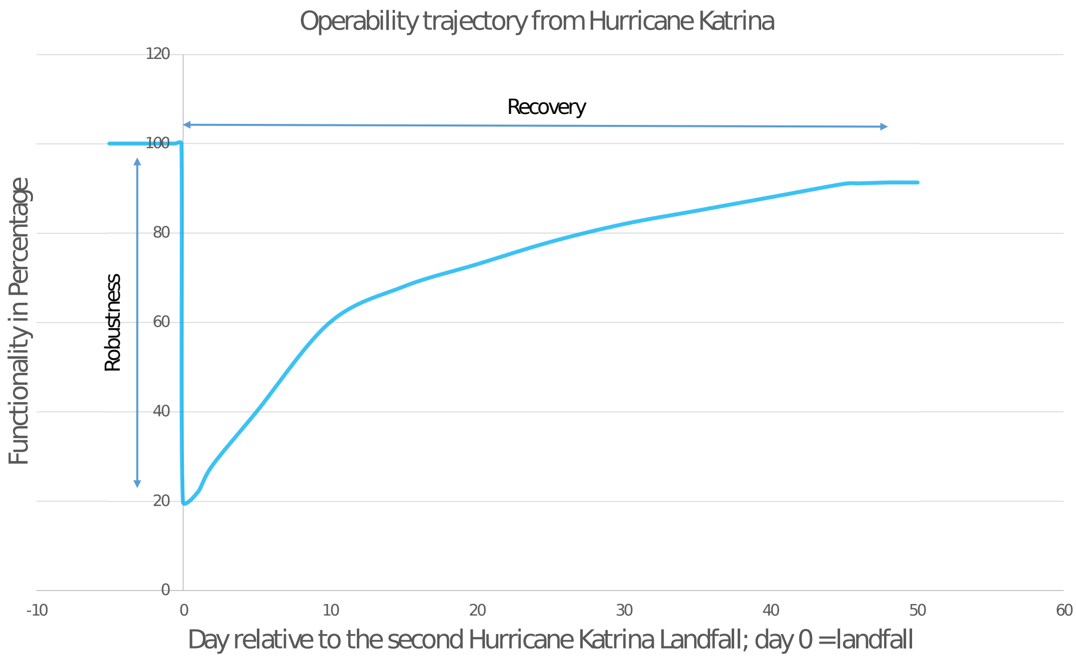


Figure 1.3: Operability trajectory after Hurricane Katrina (Reed et al. 2009).

Panteli et al. (2017) argue that this approach cannot capture other highly critical resilience dimensions of typical power systems, for example, how fast the system functionality degrades once the event hits a critical infrastructure or how long the infrastructure remains in one or more post-event degraded states before restoration is initiated and while it is fully accomplished. Therefore a resilience quantification framework is proposed, building upon the concept of a resilience trapezoid, as shown in Fig. 1.4, that depicts all the phases that a critical infrastructure, including power systems, might reside in during an event, as well as the transition between these states.

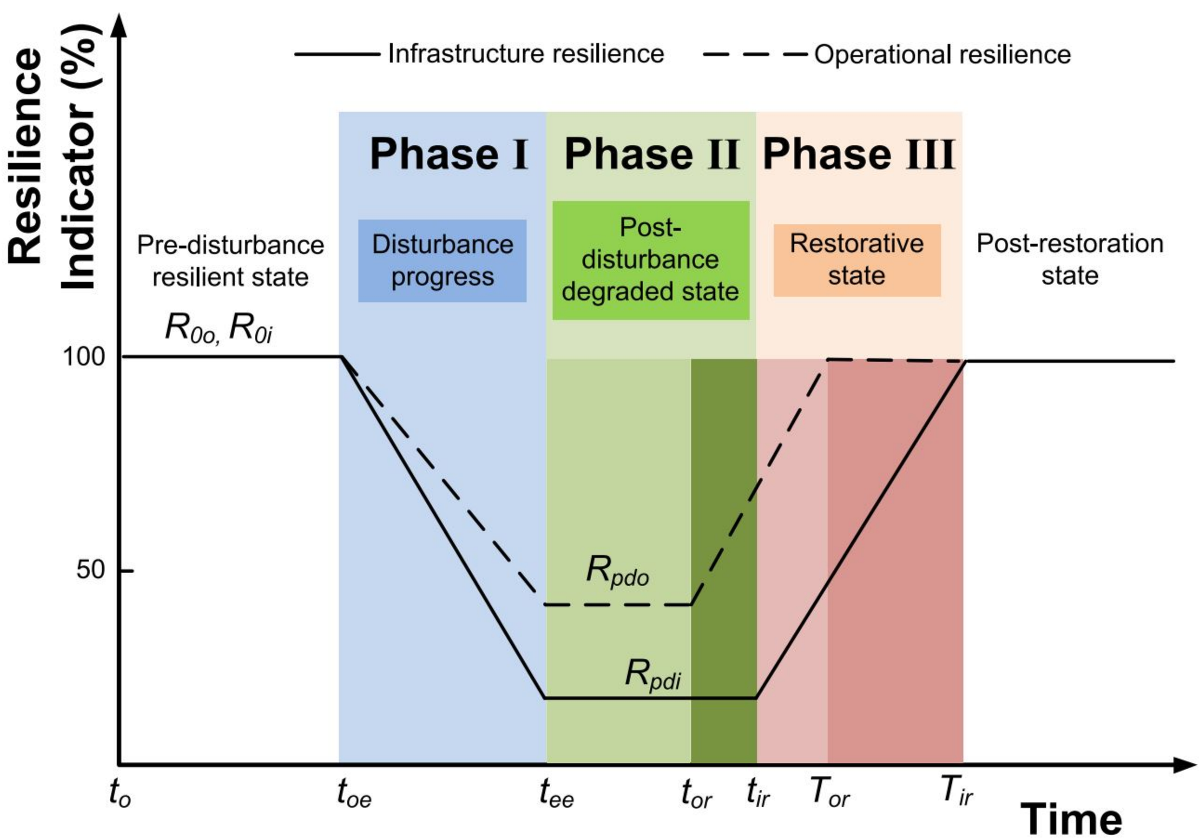


Figure 1.4: An illustration of Resilience Trapezoid by Panteli et al. (2017)

Notice that they also propose an operational resilience and infrastructure resilience.

> We will not elaborate the difference here but refer interesting readers to Section II.A of (Panteli et al. 2017). Instead, as we argue later, we will focus on electricity distribution systems. Therefore the emphasis of this research is about the recovery process, which is depicted by the resilience construct. And the tree-like topology structure unifies the two resilience metrics. But we think that the trapezoid approach might contribute to the resilience assessment of transmission network.

Finally we want to mention with passing that there are several efforts of combining multiple metrics to quantify resilience from all aspects (Ouyang & Dueñas-Osorio 2014, Willis & Loa 2015, Chanda & Srivastava 2016).

1.4 Literature review on previous research

Many resilience-related publications are conceptual. Only a few seek to resolve the problem in a technical and reasonable manner. In this section, we try to summarize and classify previous researches that feature power system resilience in each one of the three aspects, excluding resourcefulness. Again, this does not undermine the importance of resourcefulness in any means.

1.4.1 Robustness

Most of the current efforts on resilient (self-healing) distribution systems focus on robustness and in essence investigate the service restoration problem. Service restoration tackles the problem of re-energizing a part of the local, low voltage distribution network that has been disconnected following a fault. This can usually be done by optimal switching actions to isolate the faulted and healthy parts. With Distributed Generator (DG) and Distributed Energy Resources (DERs), this idea has become even more attractive. It has been proposed to incorporate DG at the end of Distribution Networks so that the system can still supply a majority of customers even with multiple faults (Chen et al. 2015) and the model is expanded by Wang et al. (2016) by considering line configurations and the fact that DGs cannot supply highly three-phase unbalanced power. Moreover, with the accurate and fast measurements from Phase Measurement Units (PMUs), the state of Distribution Networks can be monitored and the service restoration can be done autonomously without human involvement. If the process of fault detection, fault location, service restoration can be automatic, then this is the so-called Distribution Automation. The level of distribution automation may

Many other researcher may not touch the service restoration problem directly, but the loss of load at the time of extreme event is treated as the objective of some resilient operation. All such research may be classified into this aspect.

1.4.2 Rapid Recovery

The earliest effort on this kind of problem, though not necessarily done by the community of power systems, dates back to 1990s by Nojima & Kameda (1992) and there are a few publications from earthquake engineering on post-earthquake restorations (Xu et al. 2007, Çağnan et al. 2006). Several groups of researchers have been working on recovery from disasters in power systems recently. Coffrin & Van Hentenryck (2014) propose a technique to co-optimize the sequence of repairs, the load pick-ups and the generation dispatch in transmission system as well as their subsequent publication (Van Hentenryck & Coffrin 2015). They analyzed how different power flow approximation techniques could lead to the practicality of restoration plans. Nurre et al. (2012) and (2014) formulate an integrated network design and scheduling (INDS) problem, analyze the complexity of problems with different settings and propose a heuristic dispatch rule. Sharkley et al. Sharkey et al. (2015) consider the problem of restoration interdependent infrastructures after disasters and evaluate the benefit of information-sharing among different infrastructure operators. A time-stage MILP formulation for multiple team repair scheduling is proposed by Ouyang & Fang (2017). However, as we imply in the abstract and will show in Chapter 2, MILP formulation may not be efficient for real time disaster relief. Arif et al. (2017) propose a pre-processing strategy of clustering repair tasks of damaged components based on their distances from the depots and the availability of resources to reduce the size of MILP.

The current industry routines for restoration after extreme events, for the example of FirstEnergy Group, Edison Electric Institute, use dispatch rules after finishing those that threaten life safety and emergency services. We will compare these rules with our developed heuristics in Chapter 2.

1.4.3 Preparedness

Defending critical infrastructures at the transmission level has been a major research focus over the past decade (Brown et al. 2006, Bier et al. 2007, Yuan et al. 2014). In general, this research adopted the setting of Stackelberg game and formulated the problem with a tri-level defender-attacker-defender model. Such a model, in the case of natural disasters, assumes that the nature has perfect anticipation of how the defender will optimally operate the system after the attack. Arguably, this model may be more suitable for malicious attack, where attackers have a budget to spread the targets. In recent years, several researchers have investigated different methods for distribution systems hardening, but most focus solely on the robustness, i.e., worst-case load shedding at the onset of disaster. However, the objective of all these upgrade strategies are related to robustness. Detailed reviews can be seen in Chapter 5.

In practice, the utility companies are implementing hardening activities based on observations of which components failed during past disasters. While such measures will undoubtedly be useful, they do not necessarily represent the most effective way to enhance the overall resilience of the system. Large infrastructure investments may therefore not be targeted at the most effective solutions. To overcome this problem, electric utilities and government agencies in all areas that could be affected by a natural disaster need a rigorous method for assessing the relative value of various investments.

1.5 Scope and outline of this thesis

In this dissertation, we limit ourselves only to electricity distribution networks. While transmission networks can be affected by natural disasters, they are usually built to a higher standard than distribution network and thus less likely to be severely damaged. Furthermore, distribution network tend to be more concentrated than transmission networks and are thus more likely to be seriously degraded by severe weather events. While the utility companies would obviously repair the transmission system first because that's where the power comes from, repairing the distribution network usually takes the largest amount of time simply due to the fact that there are many more components that are likely to be damaged. It is thus a bottleneck and really determines how long the customers would lose power.

By the recent example of Hurricane Harvey, ERCOT was able to meet total electricity demand in part because of the lower levels of demand despite the significant amount of generation and transmission outages. This is partially because electricity demand in ERCOT was significantly lower than usual for the time of year mainly (See Fig. 1.5 by U.S. Energy Information Administation (2017) for a hourly plot of demands during the restoration process) because of the customer outages in storm-affected areas in south Texas and along the Gulf Coast and cooler temperatures across much of the state. From the perspective of transmission side, the demand was served and the market remain stable and therefore the power systems worked just fine. This shows that quantifying the resilience on the transmission side might underestimate the harm caused by the extreme events.

To quantify the aggregate harm defined by the total weighted restoration time adapted from the trajectory-based resilience metric, we need to develop an optimization problem that considers the process of repair physical damages with limited resources. And the normally pro-longed process after extreme events supports this argument.

Finally, we want to provide useful tools for industry by fitting their current practice and guidelines with minimal changes. This belongs to the future work of this dissertation.

Based on our understanding of power system resilience, a framework of key directions of researches are shown in Figure 1.6. The basis of all other problems is the post-disaster restoration covered in Chapter 2. The algorithms are extended to the case with limited capability of distribution automation (Chapter 3) and the case when the estimates of repair

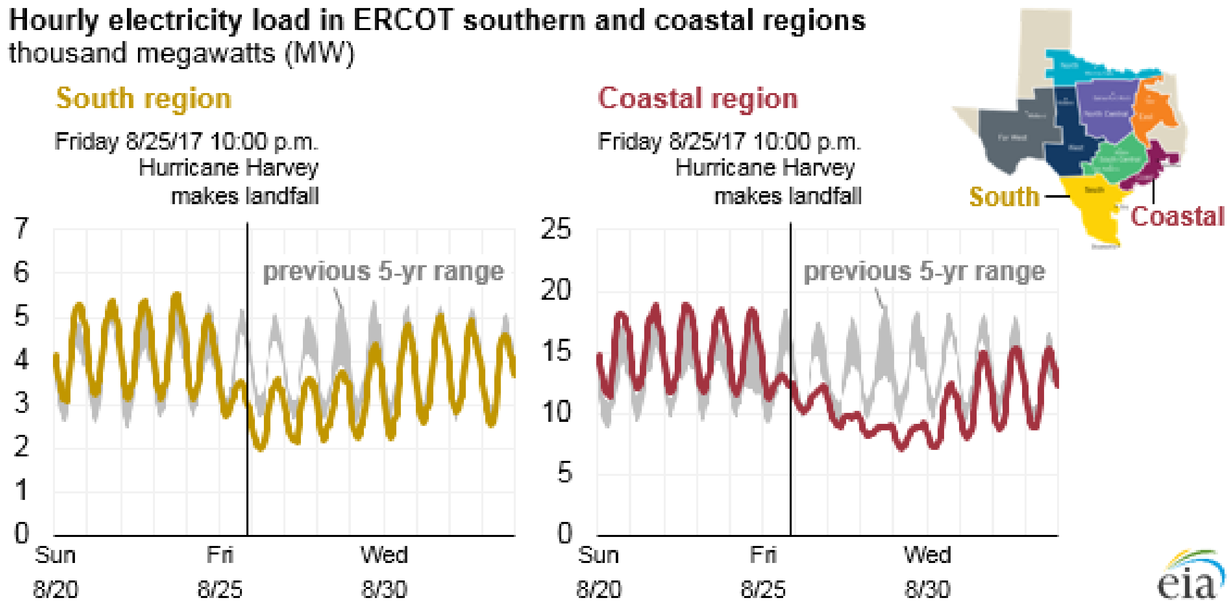


Figure 1.5: Hourly demand in ERCOT southern and costal regions byU.S. Energy Information Administation (2017)

time are not accurate (Chapter 4). We proposed a hardening strategy for distribution systems in Chapter 5. The rest of this framework constitutes the ongoing work and potential research direction, which will be briefly explained in Chapter 6.

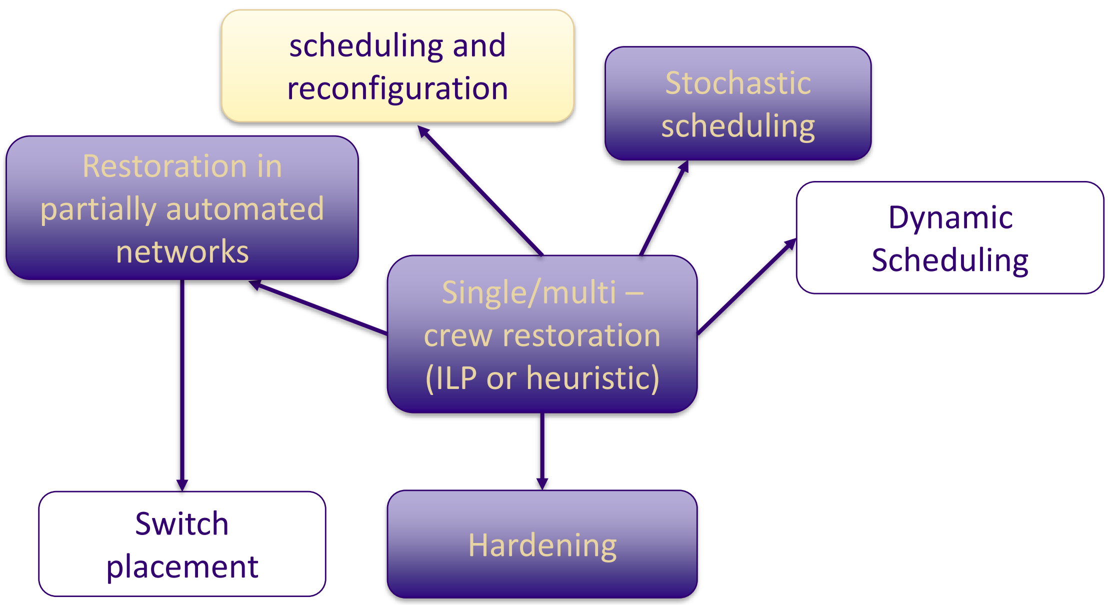


Figure 1.6: Research Path Regarding Power System Resilience

Chapter 2

SCHEDULING POST-DISASTER REPAIRS IN RADIAL DISTRIBUTION NETWORKS

2.1 Introduction

It is important to distinguish the repair scheduling problem in distribution networks discussed in this paper from the blackout restoration problem and the service restoration problem. Blackouts are large scale power outages (such as the 2003 Northeast US and Canada blackout) caused by an instability in the power generation and the high voltage transmission systems. This instability is triggered by an electrical fault or failure and is amplified by a cascade of component disconnections. Restoring power in the aftermath of a blackout is a different scheduling problem because most system components are not damaged and only need to be re-energized. See (Adibi & Fink 1994, 2006) for a discussion of the blackout restoration problem and (Sun et al. 2011) for a mixed-integer programming approach for the optimal generator start-up strategy. On the other hand, service restoration focuses on re-energizing a part of the local, low voltage distribution grid that has been automatically disconnected following a fault on a single component or a very small number of components. This can usually be done by isolating the faulted components and re-energizing the healthy parts of the network using switching actions. The service restoration problem thus involves finding the optimal set of switching actions. The repair of the faulted component is usually assumed to be taking place at a later time and is not considered in the optimization model. Several approaches have been proposed for the optimization of service restoration such as heuristics (Toune et al. 2002, Hou et al. 2011), knowledge based systems (Ma et al. 1992), and dynamic programming (Pérez-Guerrero et al. 2008).

Unlike the outages caused by system instabilities or localized faults, outages caused by natural disasters require the repair of numerous components in the distribution grid before consumers can be reconnected. The research described in this paper therefore aims to schedule the repair of a significant number of damaged components, so that the distribution network can be progressively re-energized in a way that minimizes the cumulative harm over the total restoration horizon. Fast algorithms are needed to solve this problem because it must be solved immediately after the disaster and may need to be re-solved multiple times as more detailed information about the damage becomes available. Relatively few papers address this problem. Coffrin & Van Hentenryck (2014) propose a technique to co-optimize the sequence of repairs, the load pick-ups and the generation dispatch. However, the sequencing of repair does not consider the fact that more than one repair crew could work at the same time. Nurre et al. (2012) formulate an integrated network design and scheduling (INDS) problem with multiple crews, which focuses on selecting a set of nodes and edges for installation in general infrastructure systems and scheduling them on work groups. They also propose a heuristic dispatch rule based on network flows and scheduling.

The rest of the chapter is organized as follows. We first review the basics of scheduing theory in Section 2.2. In Section 2.3, we define the problem of optimally scheduling multiple repair crews in a radial electricity distribution network after a natural disaster and model the problem by machine scheduling with soft precedence constraints. We also show that this problem is at least strongly $\mathcal{NP}$-hard. We present 4 solution techniques both exact and approximate, including integer linear programming (2.4), LP-based list scheduling algorithm (2.5) and a conversion algorithm (2.6) with a heuristic dispatch rule. In Section 2.7, we apply these methods to several standard test models of distribution networks.

2.2 Preliminaries in Scheduling Theory

Scheduling is a decision-making process that deals with the allocation of resources to tasks over time to optimize one or more objectives. It is used on a regular basis in many manufacturing and services industries (Pinedo 2012). For example, in a multi-task computer operating system need to schedule the time that the CPU or CPUs devotes to different programs. The actual processing time of each program might not be known exactly in advance and each task has a priority factor. Then the objective is to minimize the expected sum of the weighted completion time of all tasks.

There are many models for even deterministic scheduling. A framework and notation originally proposed by Graham et al. (1979) is adapted to capture the structure of many models widely considered. We will briefly review the notation along with some basic definitions related to this problem. A detailed introduction can be seen in the classic book by Pinedo (2012). Each job has a processing time p_j and weight w_j. A scheduling problem is described by a triplet $\alpha \mid \beta \mid \gamma$. The α field describes the machine environment. The β field provides the information of processing constraints and characteristics and may contain zero, one or more entries. The γ field is the objective.

The machine environments in α fields can be 1, which represents the single machine case and P_m, which represents the case of m identical parallel machines.

Possible entries in the β fields are:

Preemption(*prmp*) implies that it is not necessary to keep a job on one machine, once started, until it completes. So a job can be interrupted when it is still processing. Only when *prmp* is included in the β field, preemption is allowed.

Precedence constraints(*prec*) requires that certain jobs need to completed before another job is allowed to start. A precedence graph is a directed, acyclic graph where nodes represent tasks and where arcs, say from node i to node j, imply a precedence relation between task i and task j, i.e., task i must be completed before task j can start. If each job has at most one immediate predecessor, then the precedence graph must be a directed tree and the constraints are named *outtree*.

Define C_j as the time job j finishes on the last machine on which it requires processing. Then we consider the following three objectives:

Makespan(C_{max}) defined as $\max(C_1, \cdots, C_2)$, is the completion time of the last job.

Total weighted completion time($\sum w_j C_j$) considers the weights and the totaling holding of the schedule. It is sometimes referred to the weighted flow time. One special case is the sum of completion time($\sum C_j$).

Discounted total weighted completion time($\sum w_j(1 - e^{-rC_j})$) further takes into account the costs are discounted at a rate of r, $0 < r < 1$, per unit time.

Finally, we want to point out the difference between a sequence, a schedule. A sequence normally denotes the order of jobs to be processed on a given machine. A schedule is an allocation of jobs under a more complicated setting, most likely on multiple machines. The concept of a scheduling policy will also be introduced in Chapter 4.

2.3 Problem Formulation

A distribution network can be represented by a graph G with a set of nodes N and a set of edges (a.k.a, lines) L. We assume that the network topology G is radial, which is a valid assumption for most electricity distribution networks. Let $S \subset N$ represent the set of source nodes which are initially energized and $D = N \setminus S$ represent the set of sink nodes where consumers are located. An edge in G represents a distribution feeder or some other connecting component. In this chapter, we assume there is a switch on every edge so that the switch could be closed to energize the load as soon as the edge becomes intact. Results on the more general case that considers the partially automated radial distribution networks can be found in Chapter 2.

Severe weather can damage these components, resulting in a widespread disruption of power supply to the consumers. Let L^D and $L^I = L \setminus L^D$ denote the sets of damaged and intact edges, respectively. Each damaged edge $l \in L^D$ requires a repair time p_l which depends on the extent of the damage and the location of l. We assume that it would take every crew the same amount of time to repair the same damaged line. Without any loss of generality, we assume that there is only one source node in G. If an edge is damaged, all downstream nodes lose power due to lack of electrical connectivity. In this paper, we consider the case where multiple crews work simultaneously and independently on the repair of separate lines, along with the special case where a single crew must carry all the repairs. Finally, based on conversations with an industry expert, we make the assumption that crew travel times in a typical distribution network are small compared to actual repair times and can be ignored as a first order approximation. Therefore, our goal is to find a schedule by which the damaged lines should be repaired such that the aggregate harm due to loss of electric energy is minimized. We define this harm as follows:

$$\sum_{n \in N} w_n T_n, \tag{2.1}$$

where w_n is a positive quantity that captures the importance of node n and T_n is the time required to restore power at node n. The importance of a node depends on multiple factors, including but not limited to, the amount of load connected to it, the type of load served, and interdependency with other critical infrastructures. For example, re-energizing a node supplying a major hospital should receive a higher priority than a node supplying a similar amount of residential load. Similarly, it is conceivable that a node that provides electricity to a water sanitation plant would be assigned a higher priority. These priorities need to be assigned by the utility and their determination is outside the scope of this paper. We simply assume knowledge of the w_n's in the context of this paper.

The time to restore node n, T_n, is approximated by the *energization time* E_n, which is defined as the time node n first connects to the source node. Voltage and stability issues are not a major concern in distribution networks because they are progressively restored from a source with enough generation capacity. Even if a rigorous power flow model were to be used, the actual demands after re-energization are not known and would be hard to forecast. As a result, we model network connectivity using a simple network flow model, i.e., as long as a sink node is connected to the source, we assume that all the load on this node can be supplied without violating operating constraints. For simplicity, we treat the three-phase distribution network as if it were a single phase system. Our analysis could be extended to a three-phase system using a multi-commodity flow model, as similar to the work by Yamangil et al. (2015*a*).

2.3.1 Soft Precedence Constraints

We construct two simplified directed radial graphs to model the effect that the topology of the distribution network has on scheduling. The first graph, G', is called the 'damaged component graph'. All nodes in G that are connected by intact edges are contracted into a supernode in G'. The set of edges in G' is the set of damaged lines in G, L^D. From a computational standpoint, the nodes of G' can be obtained by treating the edges in G as undirected, deleting the damaged edges/lines, and finding all the connected components of the resulting graph. The set of nodes in each such connected component represents a (super)node in G'. The edges in G' can then be placed straightforwardly by keeping track of which nodes in G are mapped to a particular node in G'. The directions to these edges follow trivially from the network topology. G' is useful in the ILP formulation introduced in Section 2.4.

The second graph, P, is called a 'soft precedence constraint graph', which is constructed as follows. The nodes in this graph are the damaged lines in G and an edge exists between two nodes in this graph if they share the same node in G'. Computationally, the precedence constraints embodied in P can be obtained by replacing lines in G' with nodes and the nodes in G' with lines. Such a graph enables us to consider the hierarchal relationship between damaged lines, which we define as *soft precedence constraints*.

Consider, for example, the IEEE 13-node test feeder (Kersting 2001) shown in Fig. 2.1. Assume that there are four damaged lines, $650-632$, $632-645$, $684-611$ and $671-692$. The corresponding G' and P are shown in Fig. 2.2. Following the procedure discussed above and assuming that node 650 is the source, it can be verified that the precedence constraints are: (i) $(650-632) \rightarrow (632-645)$, using the path from 650 to 646, (ii) $(650-632) \rightarrow (684-611)$, using the path from 650 to 611, and (iii) $(650-632) \rightarrow (671-692)$, using the path from 650 to 675. While these constraints can be concatenated into one tree, as shown in Fig. 2.2 (b), it is quite possible to end up with multiple disjoint trees (forest). For example, if the damaged lines were

$632-645$, $645-646$, $671-684$, $684-611$ and $684-652$ instead, the precedence graph would constitute of two disjoint trees, i.e., $P = [\mathcal{T}_1, \mathcal{T}_2]$, where $\mathcal{T}_1 = [(632-645) \rightarrow (645-646)]$ and $\mathcal{T}_2 = [(671-684) \rightarrow (684-652);\ (671-684) \rightarrow (684-611)]$. If $645-646$ is deleted from the set of damaged lines, $P = \mathcal{T}_2$ effectively.

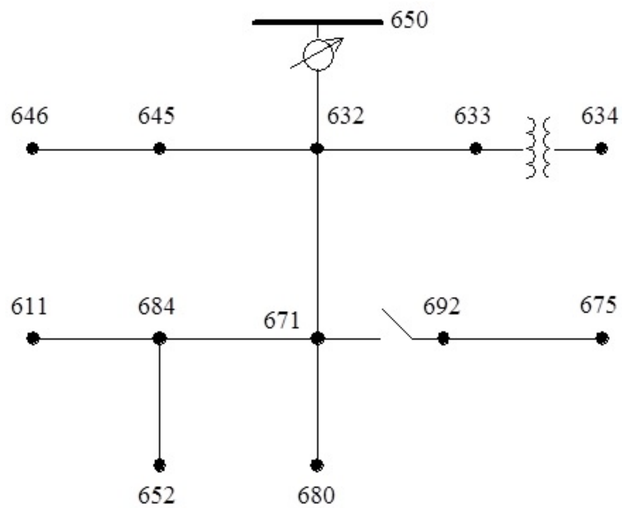
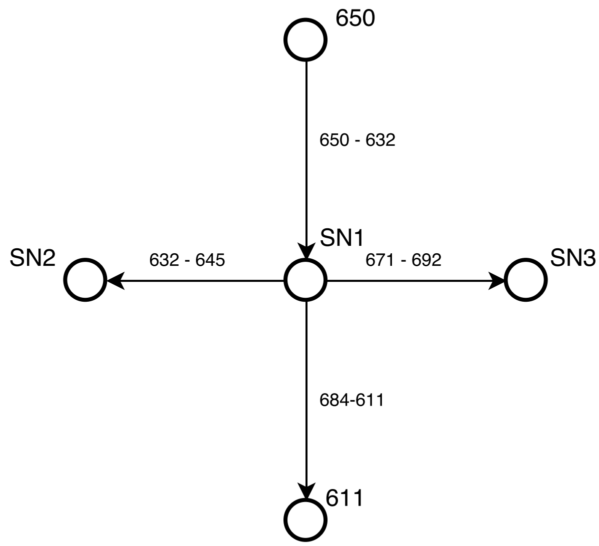


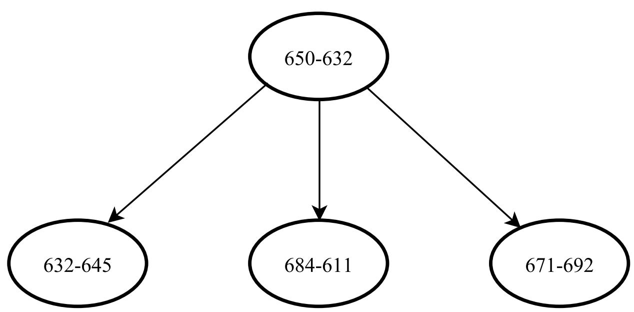
Figure 2.1: IEEE 13 Node Test Feeder

In this graph, the supernode SN_1 comprises of the nodes $\{632, 633, 634, 671, 680, 684, 652\}$, $\mathrm{SN}_2 = \{645, 646\}$, and $\mathrm{SN}_3 = \{692, 675\}$. The set of edges in this graph is the set of damaged lines. (b) The corresponding soft precedence graph, P. An edge exists between the nodes $(650-632)$ and $(684-611)$ because they share the same node, SN_1, in G'. As this graph shows, line $(650, 632)$ must be repaired first, allowing for node 632 to be energized. The next three lines that need to be repaired (in any order, since there aren't any precedence constraints among them) are the leaf nodes in the graph, before power can be restored to nodes 645/646, 611 and 692/675.

A substantial body of research exists on scheduling with precedence constraints. In general, the precedence constraint $i \prec j$ requires that job i be completed before job j is started, or equivalently, $C_j \geq C_i$, where C_j is the completion time of job j. Such precedence



(a) G' graph



(b) P graph

Figure 2.2: (a) The damaged component graph, G', obtained from Fig. 2.1, assuming that the damaged lines are $650-632$, $632-645$, $684-611$ and $671-692$. (b) The corresponding soft precedence graph, P.

constraints, however, are not applicable in post-disaster restoration. While it is true that a sink node in an electrical network cannot be energized unless there is an intact path (i.e., all damaged lines along that path have already been repaired) from the source (feeder) to this sink node, this does not mean that multiple lines on some path from the source to the sink cannot be repaired concurrently.

We keep track of two separate time vectors: the completion times of line repairs, denoted by C_l's, and the energization times of nodes, denoted by E_n's. While we have so far associated the term 'energization time' with nodes in the given network topology, G, it is also possible to define energization times on the lines. Consider the example in Fig. 2.2. The precedence graph, P, requires that the line $650-632$ be repaired prior to the line $671-692$. If this (soft) precedence constraint is met, as soon as the line $671-692$ is repaired, it can be energized, or equivalently, all nodes in SN_3 (nodes 692 and 675) in the damaged component graph, G', can be deemed to be energized. The energization time of the line $671-692$ is therefore identical to the energization times of nodes 692 and 675. Before generalizing the above example, we need to define some notations. Given a directed edge l, let $h(l)$ and $t(l)$ denote the head and tail node of l. Let $l = h(l) \rightarrow t(l)$ be any edge in the damaged component graph G'. Provided the soft precedence constraints are met, it is easy to see that $E_l = E_{t(l)}$, where E_l is the energization time of line l and $E_{t(l)}$ is the energization time of the node $t(l)$ in G. Analogously, the weight of node $t(l)$, $w_{t(l)}$, can be interpreted as a weight on the line l, w_l. The soft precedence constraint, $i \prec_S j$, therefore implies that line j cannot be energized unless line i is energized, or equivalently, $E_j \geq E_i$, where E_j is the energization time of line j.

Proposition 2.1. *Given any feasible schedule of post-disaster repairs, the energization time E_j always satisfies,*

$$E_j = \max_{i \preceq_S j} C_i \tag{2.2}$$

Proof. Denote pred(·) as the immediate predecessor of · . When j has no predecessor, $E_j = C_j = \max_{i \preceq_S j} C_i$. Now assume the proposition holds for the jobs with the depth of n in the precedence graph P. Then for any job j with a depth of $n+1$, it always satisfies that,

$$E_j = \max\{C_j, E_{\text{pred}(j)}\} = \max\{C_j, \max_{i \prec_S j} C_i\} = \max_{i \preceq_S j} C_i \tag{2.3}$$

□

So far, we have modeled the problem of scheduling post-disaster repairs in radial distribution networks as a parallel machine scheduling with *outtree soft precedence constraints* to minimize the total weighted energization time, or equivalently, $P|outtree\ soft\ prec|\sum w_j E_j$, following Graham's notation.

2.3.2 Complexity Analysis

In this section, we study the complexity of the scheduling problem $P|outtree\ soft\ prec|\sum w_j E_j$ and show that it is at least strongly $\mathcal{NP}$-hard.

Theorem 2.1. *The problem of scheduling post-disaster repairs in electricity distribution networks is at least strongly $\mathcal{NP}$-hard.*

Proof. We show this problem is at least strongly $\mathcal{NP}$-hard using a reduction from the well-known identical parallel machine scheduling problem $P \mid\mid \sum w_j C_j$ defined as follows,

$P \mid\mid \sum w_j C_j$: Given a set of jobs J in which j has processing time p_j and weight w_j, find a parallel machine schedule that minimizes the total weighted completion time $\sum w_j C_j$, where C_j is the time when job j finishes. $P \mid\mid \sum w_j C_j$ is strongly $\mathcal{NP}$-hard (Pinedo 2012, Brucker 2007).

Given an instance of $P \mid\mid \sum w_j C_j$ defined as above, construct a star network G_S with a source and $|J|$ sinks. Each sink j has a weight w_j and the line between the source and sink j has a repair time of p_j. Whenever a line is repaired, the corresponding sink can be energized. Therefore the energization time of sink j is equal to the completion time of line j. If one could solve the problem of scheduling post-disaster repairs in electricity distribution networks to optimality, then one can solve the problem in G_S optimally and equivalently solve $P \mid\mid \sum w_j C_j$. □

2.4 Integer Linear Programming (ILP) formulation

With an additional assumption in this section that all repair times are integers, we model the post-disaster repair scheduling problem using time-indexed decision variables, x_l^t, where $x_l^t = 1$ if line l is being repaired by a crew at time period t. Variable y_l^t denotes the repair status of line l where $y_l^t = 1$ if the repair is done by the end of time period $t-1$ and ready to energize at time period t. Finally, $u_i^t = 1$ if node i is energized at time period t. Let T denote the time horizon for the restoration efforts. Although we cannot know T exactly until the problem is solved, a conservative estimate should work. Since $T_i = \sum_{t=1}^{T}(1-u_i^t)$ by discretization, the objective function of eqn. 2.1 can be rewritten as:

$$\text{minimize} \quad \sum_{t=1}^{T} \sum_{i \in N} w_i (1 - u_i^t) \tag{2.4}$$

This problem is to be solved subject to two sets of constraints: (i) repair constraints and (ii) network flow constraints, which are discussed next. We mention in passing that the above time-indexed ILP formulation provides a strong relaxation of the original problem (Nurre et al. 2012) and allows for modeling of different scheduling objectives without changing the structure of the model and the underlying algorithm.

2.4.1 Repair Constraints

Repair constraints model the behavior of repair crews and how they affect the status of the damaged lines and the sink nodes that must be re-energized. The three constraints below are used to initialize the binary status variables y_l^t and u_i^t. Eqn. 2.5 forces $y_l^t = 0$ for all lines which are damaged initially (i.e., at time $t = 0$) while eqn. 2.6 sets $y_l^t = 1$ for all lines which are intact. Eqn. 2.7 forces the status of all source nodes, which are initially energized, to be equal to 1 for all time periods.

$$y_l^1 = 0, \quad \forall l \in L^D \tag{2.5}$$

$$y_l^t = 1, \quad \forall l \in L^I, \ \forall t \in [1, T] \tag{2.6}$$

$$u_i^t = 1, \quad \forall i \in S, \quad \forall t \in [1, T] \tag{2.7}$$

where T is the restoration time horizon. The next set of constraints is associated with the binary variables x_l^t. Eqn. 2.8 constrains the maximum number of crews working on damaged lines at any time period t to be equal to m, where m is the number of crews available.

$$\sum_{l \in L^D} x_l^t \leq m, \ \forall t \in [1, T] \tag{2.8}$$

Observe that, compared to the formulation by Nurre et al. (2012), there are no crew indices in our model. Since these indices are completely arbitrary, the number of feasible solutions can increase in crew indexed formulations, leading to enhanced computation time. For example, consider the simple network $i \to j \to k \to l$, where node i is the source and all edges require a repair time of 5 time units. If 2 crews are available, suppose the optimal repair schedule is: 'assign team 1 to $i \to j$ at time $t = 0$, team 2 to $j \to k$ at $t = 0$, and team 1 to $k \to l$' at $t = 5$. Clearly, one possible equivalent solution conveying the same repair schedule and yielding the same cost, is: 'assign team 2 to $i \to j$ at $t = 0$, team 1 to $j \to k$ at $t = 0$, and team 1 to $k \to l$ at $t = 5$'. In general, formulations without explicit crew indices may lead to a reduction in the size of the feasible solution set. Although the optimal repair sequences obtained from such formulations do not natively produce the work assignments to the different crews, this is not an issue in practice because operators can choose to let a crew work on a line until the job is complete and assign the next repair job in the sequence to the next available crew (the first m jobs in the optimal repair schedule can be assigned arbitrarily to the m crews).

Finally, constraint eqn. 2.9 formalizes the relationship between variables x_l^t and y_l^t. It mandates that y_l^t cannot be set to 1 unless at least p_l number of x_l^τ's, $\tau \in [1, t-1]$, are equal to 1, where p_l is the repair time of line l.

$$y_l^t \leq \frac{1}{p_l} \sum_{\tau=1}^{t-1} x_l^\tau, \ \forall l \in L^D, \ \forall t \in [1, T] \tag{2.9}$$

While we do not explicitly require that a crew may not leave its current job unfinished and take up a different job, it is obvious that such a scenario cannot be part of an optimal repair schedule.

2.4.2 Network flow constraints

We use a modified form of standard flow equations to simplify power flow constraints. Specifically, we require that the flows, originating from the source nodes (eqn. 2.10), travel through lines which have already been repaired (eqn. 2.11). Once a sink node receives a flow, it can be energized (eqn. 2.12).

$$\sum_{l \in \delta_G^-(i)} f_l^t \geq 0, \qquad \forall t \in [1, T], \forall i \in S \tag{2.10}$$

$$-M \times y_l^t \leq f_l^t \leq M \times y_l^t, \quad \forall t \in [1, T], \forall l \in L \tag{2.11}$$

$$u_i^t \leq \sum_{l \in \delta_G^+(i)} f_l^t - \sum_{l \in \delta_G^-(i)} f_l^t, \qquad \forall t \in [1, T], \forall i \in D \tag{2.12}$$

In eqn. 2.11, M is a suitably large constant, which, in practice, can be set equal to the number of sink nodes, $M = |D|$. In eqn. 2.12, $\delta_G^+(i)$ and $\delta_G^-(i)$ denote the sets of lines on which power flows into and out of node i in G respectively.

2.4.3 Valid inequalities

Valid inequalities typically reduce the computing time and strengthen the bounds provided by the LP relaxation of an ILP formulation. We present the following shortest repair time path inequalities, which resemble the ones by Nurre et al. (2012). A node i cannot be energized until all the lines between the source s and node i are repaired. Since the lower bound to finish all the associated repairs is $\lfloor SRTP_i/m \rfloor$, where m denotes the number of crews available and $SRTP_i$ denotes the shortest repair time path between s and i, the following inequality is valid:

$$\sum_{t=1}^{\lfloor SRTP_i/m \rfloor - 1} u_i^t = 0, \quad \forall i \in N \tag{2.13}$$

To summarize, the multi-crew distribution system post-disaster repair problem can be formulated as:

$$\begin{aligned} \text{minimize} \quad & \text{eqn. } 2.4 \\ \text{subject to} \quad & \text{eqns. } 2.5 \sim 2.13 \end{aligned} \tag{2.14}$$

2.5 List scheduling algorithms based on linear relaxation

A majority of the approximation algorithms used for scheduling is derived from linear relaxations of ILP models, based on the scheduling polyhedra of completion vectors developed by Queyranne (1993), Schulz et al. (1996). We briefly restate the definition of scheduling polyhedra and then introduce a linear relaxation based list scheduling algorithm followed by a worst case analysis of the algorithm.

2.5.1 Linear relaxation of scheduling with soft precedence constraints

A set of valid inequalities for m identical parallel machine scheduling was presented by Schulz et al. (1996):

$$\sum_{j \in A} p_j C_j \geq f(A) := \frac{1}{2m} \left(\sum_{j \in A} p_j \right)^2 + \frac{1}{2} \sum_{j \in A} p_j^2 \quad \forall A \subset N \tag{2.15}$$

Theorem 2.2 (Schulz et al. (1996))**.** *The completion time vector C of every feasible schedule on m identical parallel machines satisfies inequalities (2.15).*

The objective of the post-disaster repair and restoration is to minimize the harm, quantified as the total weighted energization time. With the previously defined soft precedence constraints and the valid inequalities for parallel machine scheduling, we propose the following LP relaxation:

$$\underset{C,E}{\text{minimize}} \quad \sum_{j \in L^D} w_j E_j \tag{2.16}$$

$$\text{subject to} \quad C_j \geq p_j, \ \forall j \in L^D \tag{2.17}$$

$$E_j \geq C_j, \ \forall j \in L^D \tag{2.18}$$

$$E_j \geq E_i, \ \forall (i \to j) \in P \tag{2.19}$$

$$\sum_{j \in A} p_j C_j \geq \frac{1}{2m} \left(\sum_{j \in A} p_j \right)^2 + \frac{1}{2} \sum_{j \in A} p_j^2, \ \ \forall A \subset L^D \tag{2.20}$$

where P is the soft precedence graph discussed in Section 2.3 (see also Fig. 2.2). Eqn. 2.17 constrains the completion time of any damaged line to be lower bounded by its repair time; eqn. 2.18 ensures that any line cannot be energized until it has been repaired; eqn. 2.19 models the soft precedence constraints and eqn. 2.20 characterizes the scheduling polyhedron.

The above formulation can be simplified by recognizing that the C_j's are redundant intermediate variables. Combining eqns. 2.18 and 2.20, we have:

$$\sum_{j \in A} p_j E_j \ \geq \ \sum_{j \in A} p_j C_j \ \geq \ \frac{1}{2m} \left(\sum_{j \in A} p_j \right)^2 + \frac{1}{2} \sum_{j \in A} p_j^2, \quad \forall A \subset L^D \tag{2.21}$$

which indicates that the vector of E_j's satisfies the same valid inequalities as the vector of C_j's. After some simple algebra, the LP-relaxation can be reduced to:

$$\underset{E}{\text{minimize}} \quad \sum_{j \in L^D} w_j E_j \tag{2.22}$$

$$\text{subject to} \quad E_j \geq p_j, \ \forall j \in L^D \tag{2.23}$$

$$E_j \geq E_i, \ \forall (i \to j) \in P \tag{2.24}$$

$$\sum_{j \in A} p_j E_j \geq \frac{1}{2m} \left(\sum_{j \in A} p_j \right)^2 + \frac{1}{2} \sum_{j \in A} p_j^2, \ \ \forall A \subset L^D \tag{2.25}$$

We note that although there are exponentially many constraints in the above model, the separation problem for these inequalities can be solved in polynomial time using the ellipsoid method as shown by Schulz et al. (1996).

2.5.2 LP-based approximation algorithm

List scheduling algorithms, which are among the simplest and most commonly used approximate solution methods for parallel machine scheduling problems (Queyranne & Schulz 2006), assign the job at the top of a priority list to whichever machine is idle first. An LP relaxation provides a good insight into the priorities of jobs and has been widely applied to scheduling with hard precedence constraints. We adopt a similar approach in this paper. Algorithm 1, based on a sorted list of the LP midpoints, summarizes our proposed approach. We now develop an approximation bound for Algorithm 1.

Algorithm 1 Algorithm for single/multiple crew repair scheduling in distribution networks, based on LP midpoints

---

Let E^{LP} denote any feasible solution to the constraint eqns. 4.4b - 4.4d. Define the LP mid points to be $M_j^{LP} := E_j^{LP} - p_j/2, \ \forall j \in L^D$. Create a job priority list by sorting the M_j^{LP}'s in an ascending order. Whenever a crew is free, assign to it the next job from the priority list. The first m jobs in the list are assigned arbitrarily to the m crews.

Proposition 2.2. *Let E_j^H denote the energization time respectively of line j in the schedule constructed by Algorithm 1. Then the following must hold,*

$$E_j^H \leq 2E_j^{LP}, \quad \forall j \in L^D \tag{2.26}$$

Proof. Let S_j^H, C_j^H and E_j^H denote the start time, completion time respectively of some line j in the schedule constructed by Algorithm 1. Define $M := \left[M_j^{LP} : j = 1, 2, \ldots, \left|L^D\right|\right]$. Let $\tilde{M}$ denote M sorted in ascending order, $\tilde{\mathcal{I}}_j$ denote the position of some line $j \in L^D$ in $\tilde{M}$, and $\left\{k : \tilde{\mathcal{I}}_k \leq \tilde{\mathcal{I}}_j, k \neq j\right\} := R$ denote the set of jobs whose LP midpoints are upper bounded by M_j^{LP}. First, we claim that $S_j^H \leq \frac{1}{m}\sum_{i \in R} p_i$. To see why, split the set R into m subsets, corresponding to the schedules of the m crews, i.e., $R = \bigcup_{k=1}^{m} R^k$. Since job j is assigned to the first idle crew and repairs commence immediately, we have:

$$S_j^H = \min\left\{\sum_{i\in R^k} p_i : k = 1,2,\ldots m\right\} \le \frac{1}{m}\sum_{k=1}^{m}\sum_{i\in R^k} p_i = \frac{1}{m}\sum_{i\in R} p_i\,, \tag{2.27}$$

where the inequality follows from the fact that the minimum of a set of positive numbers is upper bounded by the mean. Next, noting that $M_j^{LP} = E_j^{LP} - p_j/2$, we rewrite eqn. 4.4d as follows:

$$\sum_{j\in A} p_j M_j^{LP} \ge \frac{1}{2m}\left(\sum_{j\in A} p_j\right)^2, \ \forall A \subset L^D \tag{2.28}$$

Now, letting $A = R$, we have:

$$\left(\sum_{i\in R} p_i\right) M_j^{LP} \ \ge\ \sum_{i\in R} p_i M_i^{LP} \ \ge\ \frac{1}{2m}\left(\sum_{i\in R} p_i\right)^2, \tag{2.29}$$

where the first inequality follows from the fact that $M_j^{LP} \ge M_i^{LP}$ for any $i \in R$. Combining eqns. 3.6 and 3.9, it follows that $S_j^H \le 2M_j^{LP}$. Consequently, $C_j^H = S_j^H + p_j \le 2M_j^{LP} + p_j = 2E_j^{LP}$. Then,

$$E_j^H = \max_{i \preceq_S j} C_i^H \le \max_{i \preceq_S j} 2E_i^{LP} = 2E_j^{LP}, \tag{2.30}$$

where the last equality follows trivially from the definition of a soft precedence constraint. □

Theorem 2.3. *Algorithm 1 is a 2-approximation.*

Proof. Let E_j^* denote the energization time of line j in the optimal schedule. Then, with E_j^{LP} being the solution of the linear relaxation,

$$\sum_{j\in L^D} w_j E_j^{LP} \le \sum_{j\in L^D} w_j E_j^* \tag{2.31}$$

Finally, from eqns. 2.26 and 2.31, we have:

$$\sum_{j\in L^D} w_j E_j^H \le 2\sum_{j\in L^D} w_j E_j^{LP} \le 2\sum_{j\in L^D} w_j E_j^* \tag{2.32}$$

□

2.6 An algorithm for converting the optimal single crew repair sequence to a multi-crew schedule

In practice, many utilities schedule repairs using a priority list (Xu et al. 2007), which leaves much space for improvement. We analyze the repair and restoration process as it would be done with a single crew because this provides important insights into the general structure of the multi-crew scheduling problem. Subsequently, we provide an algorithm for converting the single crew repair sequence to a multi-crew schedule, which is inspired by similar previous work by Chekuri & Khanna (2004), and analyze its worst case performance. Finally, we develop a multi-crew dispatch rule and compare it with the current practices of FirstEnergy Group and Edison Electric Institute.

2.6.1 Single crew restoration in distribution networks

We show that this problem is equivalent to $1 \mid outtree \mid \sum w_j C_j$, which stands for scheduling to minimize the total weighted completion time of N jobs with a single machine under 'outtree' precedence constraints. Outtree precedence constraints require that each job may have at most one predecessor. Given the manner in which we derive the soft precedence, it is easy to see that P will indeed follow outtree precedence requirements, i.e. each node in P will have at most one predecessor, as long as the network topology G does not have any cycles. We will show by the following lemma that the soft precedence constraints degenerate to the precedence constraints with one repair team.

Proposition 2.3. *Given one repair crew, the optimal schedule in a radial distribution system must follow outtree precedence constraints, the topology of which follows the soft precedence graph P.*

Proof. Given one repair crew, each schedule can be represented by a sequence of damaged lines. Let $i - j$ and $j - k$ be two damaged lines such that the node (j, k) is the immediate successor of node (i, j) in the soft precedence graph P. Let π be the optimal sequence and π' another sequence derived from π by swapping $i - j$ and $j - k$. Denote the energization times of nodes j and k in π by E_j and E_k respectively. Similarly, let E'_j and E'_k denote the energization times of nodes j and k in π'. Define $f := \sum_{n \in N} w_n E_n$.

Since node k cannot be energized unless node j is energized and until the line between it and its immediate predecessor is repaired, we have $E'_k = E'_j$ in π' and $E_k > E_j$ in π. Comparing π and π', we see that node k is energized at the same time, i.e., $E'_k = E_k$, and therefore, $E'_j > E_j$. Thus:

$$\begin{aligned} f(\pi') - f(\pi) &= (w_j E'_j + w_k E'_k) - (w_j E_j + w_k E_k) \\ &= w_j(E'_j - E_j) + w_k(E'_k - E_k) > 0 \end{aligned} \tag{2.33}$$

Therefore, any job swap that violates the outtree precedence constraints will strictly increase the objective function. Consequently, the optimal sequence must follow these constraints. □

It follows immediately from Proposition 2.3 that:

Lemma 2.1. *Single crew repair scheduling in distribution networks is equivalent to* $1 \mid$ *outtree* $\mid \sum_j w_j C_j$*, where the outtree precedences are given in the soft precedence constraint graph* P*.*

2.6.2 Recursive scheduling algorithm for single crew restoration scheduling

As shown above, the single crew repair scheduling problem in distribution networks is equivalent to $1 \mid$ *outtree* $\mid \sum w_j C_j$, for which an optimal algorithm exists (Adolphson & Hu 1973). We will briefly discuss this algorithm and the reasoning behind it. Details and proofs can be found by (Brucker 2007). Let $J^D \subseteq L^D$ denote any subset of damaged lines. Define:

$$w\left(J^D\right) := \sum_{j \in J^D} w_j, \quad p\left(J^D\right) := \sum_{j \in J^D} p_j, \quad q\left(J^D\right) := \frac{w\left(J^D\right)}{p\left(J^D\right)}$$

Algorithm 2, adapted from (Brucker 2007) with a change of notation, finds the optimal repair sequence by recursively merging the nodes in the soft precedence graph P. The input to this algorithm is the precedence graph P. Let $N(P) = \{1, 2, \ldots |N(P)|\}$ denote the set of nodes in P (representing the set of damaged lines, L^D), with node 1 being the designated root.

The predecessor of any node $n \in P$ is denoted by pred(n). Lines $1 - 7$ initialize different variables. In particular, we note that the predecessor of the root is arbitrarily initialized to be 0 and its weight is initialized to $-\infty$ to ensure that the root node is the first job in the optimal repair sequence. Broadly speaking, at each iteration, a node $j \in N(P)$ (j could also be a group of nodes) is chosen to be merged into its immediate predecessor $i \in N(P)$ if $q(j)$ is the largest. The algorithm terminates when all nodes have been merged into the root. Upon termination, the optimal single crew repair sequence can be recovered from the predecessor vector and the element $A(1)$, which indicates the last job finished.

We conclude this section by noting that Algorithm 2 requires the precedence graph P to have a defined root. However, as illustrated in Section 2.3, it is quite possible for P to be a forest, i.e., a set of disjoint trees. In such a situation, P can be modified by introducing a dummy root node with a repair time of 0 and inserting directed edges from this dummy root to the roots of each individual tree in the forest. This fictitious root will be the first job in the repair sequence returned by the algorithm, which can then be stripped off.

2.6.3 Conversion algorithm and an approximation bound

A greedy procedure for converting the optimal single crew sequence to a multiple crew schedule is given in Algorithm 3. We now prove that it is a $\left(2 - \frac{1}{m}\right)$ approximation algorithm. We start with two lemmas that provide lower bounds on the minimal harm for an m-crew schedule, in terms of the minimal harms for single crew and ∞-crew schedules. Let $H^{1,*}$, $H^{m,*}$ and $H^{\infty,*}$ denote the minimal harms when the number of repair crews is 1, some arbitrary m ($2 \leq m < \infty$), and ∞ respectively.

Proposition 2.4. $H^{m,*} \geq \frac{1}{m} H^{1,*}$

Proof. Given an arbitrary m-crew schedule S^m with harm H^m, we first construct a 1-crew repair sequence, S^1. We do so by sorting the energization times of the damaged lines in S^m in ascending order and assigning the corresponding sorted sequence of lines to S^1. Ties, if any, are broken according to precedence constraints or arbitrarily if there is none. By

Algorithm 2 Optimal algorithm for single crew repair and restoration in distribution networks.

$w(1) \leftarrow -\infty$; pred$(1) \leftarrow 0$;
for $n = 1$ to $|N(P)|$ **do**
  $A(n) \leftarrow n$; $B_n \leftarrow \{n\}$; $q(n) \leftarrow w(n)/p(n)$;
end for
for $n = 2$ to $|N(P)|$ **do**
  pred$(n) \leftarrow$ parent of n in P;
end for
nodeSet $\leftarrow \{1, 2, \cdots, |N(P)|\}$;
while nodeSet $\neq \{1\}$ **do**
  Find $j \in$ nodeSet such that $q(j)$ is largest; % `ties can be broken arbitrarily`
  Find i such that pred$(j) \in B_i$, $i = 1, 2, \ldots |N(P)|$;
  $w(i) \leftarrow w(i) + w(j)$;
  $p(i) \leftarrow p(i) + p(j)$;
  $q(i) \leftarrow w(i)/p(i)$;
  pred$(j) \leftarrow A(i)$;
  $A(i) \leftarrow A(j)$;
  $B_i \leftarrow \{B_i, B_j\}$; % `',' denotes concatenation`
  nodeSet $\leftarrow$ nodeSet $\setminus \{j\}$;
end while

Algorithm 3 Algorithm for converting the optimal single crew schedule to an m-crew schedule

Treat the optimal single crew repair sequence as a priority list, and, whenever a crew is free, assign to it the next job from the list. The first m jobs in the single crew repair sequence are assigned arbitrarily to the m crews.

construction, for any two damaged lines i and j with precedence constraint $i \prec j$, the completion time of line i must be strictly smaller than the completion time of line j in S^1, i.e., $C_i^1 < C_j^1$. Additionally, $C_i^1 = E_i^1$ because the completion and energization times of lines are identical for a 1-crew repair sequence which also meets the precedence constraints of P.

Next, we claim that $E_i^1 \leq mE_i^m$, where E_i^1 and E_i^m are the energization times of line i in S^1 and S^m respectively. In order to prove it, we first observe that:

$$E_i^1 = C_i^1 = \sum_{\{j:\ E_j^m \leq E_i^m\}} p_j \quad \leq \sum_{\{j:\ C_j^m \leq E_i^m\}} p_j \,, \tag{2.34}$$

where the second equality follows from the manner we constructed S^1 from S^m and the inequality follows from the fact that $C_j^m \leq E_j^m \Rightarrow \{j :\ E_j^m \leq E_i^m\} \subseteq \{j :\ C_j^m \leq E_i^m\}$ for any m-crew schedule. In other words, the number of lines that have been energized before line i is energized is a subset of the number of lines on which repairs have been completed before line i is energized. Next, we split the set $\{j :\ C_j^m \leq E_i^m\} := R$ into m subsets, corresponding to the schedules of the m crews in S^m, i.e., $R = \bigcup_{k=1}^m R^k$, where R^k is a subset of the jobs in R that appear in the k^{th} crew's schedule. It is obvious that the sum of the repair times of the lines in each R^k can be no greater than E_i^m. Therefore,

$$E_i^1 \leq \sum_{\{j:\ C_j^m \leq E_i^m\}} p_j := \sum_{j \in R} p_j = \sum_{k=1}^{m} \left(\sum_{j \in R^k} p_j \right) \leq mE_i^m \tag{2.35}$$

Proceeding with the optimal m-crew schedule $S^{m,*}$ instead of an arbitrary one, it is easy to see that $E_i^1 \leq mE_i^{m,*}$, where $E_i^{m,*}$ is the energization time of line i in $S^{m,*}$. The lemma then follows straightforwardly.

$$H^{m,*} = \sum_{i \in L^D} w_i E_i^{m,*} \geq \sum_{i \in L^D} w_i \, \frac{1}{m} \, E_i^1 = \frac{1}{m} \sum_{i \in L^D} w_i \, E_i^1 = \frac{1}{m} H^1 \geq \frac{1}{m} H^{1,*} \tag{2.36}$$

□

Before stating the next lemma, we provide an example which illustrates some of the ideas in the proof of the previous lemma. Consider the graph $G = (a - b - c - d - e)$, where node a is the source. From left to right, the lines are numbered 1, 2, 3, and 4, with repair times 10, 40, 20 and 30 respectively. Assume that all lines are damaged and $m = 2$. Suppose $S^2 = [(\text{crew-1}) : 1, 3;\ (\text{crew-2}) : 2, 4]$. The energization and completion times of the four lines are: (i) $C_1^2 = E_1^2 = 10$, (ii) $C_2^2 = 40$, $E_2^2 = 40$, (iii) $C_3^2 = 30$, $E_3^2 = 40$, and (iv) $C_4^2 = E_4^2 = 70$. Notice that even though line 3 $(c - d)$ is completed at time $t = 30$, it can only be energized at time $t = 40$ since that's when repairs on line 2 $(b - c)$ are completed. In fact, repairs on two lines $(a - b$ and $c - d)$ have been completed before time $t = E_2^2 = 40$, but only one $(a - b)$ has been energized. The precedence graph for this example is $P = (a-b) \rightarrow (b-c) \rightarrow (c-d) \rightarrow (d-e)$. Sorting the energization times in S^2 in ascending order, the 1-crew sequence is: $S^1 = [1, 2, 3, 4]$, where we used P to break a tie between lines 2 and 3. It can be verified that the completion and energization times of the lines are identical in S^1.

Proposition 2.5. $H^{m,*} \geq H^{\infty,*}$

Proof. This is intuitive, since the harm is minimized when the number of repair crews is at least equal to the number of damaged lines. In the ∞-crew case, every job can be assigned to one crew. For any damaged line $j \in L^D$, $C_j^{\infty} = p_j$ and $E_j^{\infty} = \max_{i \preceq j} C_i^{\infty} = \max_{i \preceq j} p_i$. Also, $C_j^{m,*} \geq p_j = C_j^{\infty}$ and $E_j^{m,*} = \max_{i \preceq j} C_i^{m,*} \geq \max_{i \preceq j} p_i = E_j^{\infty}$. Therefore:

$$H^{m,*} = \sum_{j \in L^D} w_j E_j^{m,*} \geq \sum_{j \in L^D} w_j E_j^{\infty} = H^{\infty,*} \tag{2.37}$$

□

Proposition 2.6. *Let E_j^m be the energization time of line j after the conversion algorithm is applied to the optimal single crew repair schedule. Then, $\forall j \in L^D$, $E_j^m \leq \frac{1}{m} E_j^{1,*} + \frac{m-1}{m} E_j^{\infty,*}$.*

Proof. Let S_j^m and C_J^m denote respectively the start and energization times of some line $j \in L^D$ in the m-crew repair schedule, S^m, obtained by applying the conversion algorithm to the optimal 1-crew sequence, $S^{1,*}$. Also, let $\mathcal{I}_j$ denote the position of line j in $S^{1,*}$ and

$\{k : \mathcal{I}_k < \mathcal{I}_j\} := R$ denote the set of all lines completed before j in $S^{1,*}$. First, we claim that: $S_j^m \leq \frac{1}{m} \sum_{i \in R} p_i$. A proof can be constructed by following the approach taken in the proof of Proposition 2.2 and is therefore omitted. Now:

$$
\begin{aligned}
C_j^m &= S_j^m + p_j && (2.38)\\
&\leq \frac{1}{m} \sum_{i \in R} p_i + p_j && (2.39)\\
&= \frac{1}{m} \sum_{i \in R \cup j} p_i + \frac{m-1}{m} p_j && (2.40)\\
&= \frac{1}{m} C_j^{1,*} + \frac{m-1}{m} p_j && (2.41)
\end{aligned}
$$

and

$$
\begin{aligned}
E_j^m &= \max_{i \preceq_S j} C_i^m && (2.42)\\
&\leq \max_{i \preceq j} \frac{1}{m} C_i^{1,*} + \max_{i \preceq j} \frac{m-1}{m} p_i && (2.43)\\
&= \frac{1}{m} C_j^{1,*} + \frac{m-1}{m} \max_{i \preceq j} p_i && (2.44)\\
&= \frac{1}{m} E_j^{1,*} + \frac{m-1}{m} E_j^{\infty,*} && (2.45)
\end{aligned}
$$

□

Theorem 2.4. *The conversion algorithm is a* $\left(2 - \frac{1}{m}\right)$*-approximation.*

Proof.

$$H^m = \sum_{j \in L^D} w_j E_j^m \tag{2.46}$$

$$\leq \sum_{j \in L^D} w_j \left(\frac{1}{m} E_j^{1,*} + \frac{m-1}{m} E_j^{\infty} \right) \quad \cdots \text{ using Proposition 2.6} \tag{2.47}$$

$$= \frac{1}{m} \sum_{j \in L^D} w_j E_j^{1,*} + \frac{m-1}{m} \sum_{j \in L^D} w_j E_j^{\infty} \tag{2.48}$$

$$= \frac{1}{m} H^{1,*} + \frac{m-1}{m} H^{\infty,*} \tag{2.49}$$

$$\leq \frac{1}{m} \left(m H^{m,*} \right) + \frac{m-1}{m} H^{m,*} \quad \cdots \text{ using Propositions 2.4 - 2.5} \tag{2.50}$$

$$= \left(2 - \frac{1}{m} \right) H^{m,*} \tag{2.51}$$

□

2.6.4 A Dispatch Rule

We now develop a multi-crew dispatch rule from a slightly different perspective, and show that it is equivalent to the conversion algorithm. In the process, we define a parameter, $\rho(l)$, $\forall l \in L^D$, which can be interpreted as a 'component importance measure' (CIM) in the context of reliability engineering. This allows us to easily compare our conversion algorithm to standard utility practices. Towards that goal, we revisit the single crew repair problem, in conjunction with the algorithm proposed by Horn (1972).

Let S_l denote the set of all trees rooted at node l in P and $s_l^* \in S_l$ denote the minimal subtree which satisfies:

$$\rho(l) := \frac{\sum_{j \in N(s_l^*)} w_j}{\sum_{j \in N(s_l^*)} p_j} = \max_{s_l \in S_l} \left(\frac{\sum_{j \in N(s_l)} w_j}{\sum_{j \in N(s_l)} p_j} \right), \tag{2.52}$$

where $N(s_l)$ is the set of nodes in s_l. We define the ratio on the left-hand side of the equality in eqn. 2.52 to be the ρ-factor of line l, denoted by $\rho(l)$. We refer to the tree s_l^* as the minimal ρ-maximal tree rooted at l, which resembles the definitions discussed in Sidney (1975). With ρ-factors calculated for all damaged lines, the repair scheduling with single crew can be solved optimally, as stated in Algorithm 4 below, adopted from (Horn 1972). Note that ρ-factors are defined based on the soft precedence graph P, whereas the following dispatch rules are stated in terms of the original network G to be more in line with industry practices.

Algorithm 4 Algorithm for single crew repair scheduling in distribution networks

Whenever the crew is free, say at time t, select among the candidate lines the one with the highest ρ-factor. The candidate set comprises all the damaged lines, one of whose end points is within the set of energized nodes at time t.

It has been proven by Adolphson & Hu (1973) that Algorithms 2 and 4 are equivalent. The ρ-factors can be calculated in multiple ways: (1) following the method proposed by Horn (1972), (2) as a byproduct of Algorithm 2, and (3) using a more general method based on parametric minimum cuts in an associated directed precedence graph by Lawler (1978). Algorithm 4 can be extended straightforwardly to accommodate multiple crews. However, in this case, it could happen that the number of damaged lines that are connected to energized nodes is smaller than the number of available repair crews. To cope with this issue, we also consider the lines which are connected to the lines currently being repaired, as described in Algorithm 5 below.

Algorithm 5 Algorithm for multi-crew repair scheduling in distribution networks

Whenever a crew is free, say at time t, select among the remaining candidate lines the one with the highest ρ-factor. The candidate set consists of all the damaged lines that are connected to already energized nodes, as well as the lines that are being repaired at time t.

Theorem 2.5. *Algorithm 5 is equivalent to Algorithm 3 discussed in Section 2.6.*

Proof. As stated above, Algorithms 2 and 4 are both optimal algorithms and we assume that, without loss of generality, they produce the same optimal sequences. Then it suffices to show that Algorithm 5 converts the sequence generated by Algorithm 4 in the same way that Algorithm 3 does to Algorithm 2.

The proof is by induction on the order of lines being selected. In iteration 1, it is obvious that Algorithms 4 and 5 choose the same line for repair. Suppose this is also the case for iterations 2 to $t-1$, with the lines chosen for repair being l_1, l_2, l_3, $\cdots$, and l_{t-1} respectively. Then, in iteration t, the set of candidate lines for both algorithms is the set of immediate successors of the supernode $\{l_1, l_2, \cdots, l_{t-1}\}$. Both algorithms will choose the job with the largest ρ-factor in iteration t, thereby completing the induction process. □

2.6.5 Comparison with current industry practices

According to FirstEnergy Group's recommendation for its operating companies, repair crews will "address outages that restore the largest number of customers before moving to more isolated problems". This policy can be interpreted as a priority-based scheduling algorithm and fits within the scheme of the dispatch rule discussed above, the difference being that, instead of selecting the line with the largest ρ-factor, FirstEnergy chooses the one with the largest weight (which turns out to be the number of customers). Additionally, according to Step 5 of a recent report by Edison Electric Institute, crews should be dispatched to "repair lines that will return service to the largest number of customers in the least amount of time". This policy is analogous to Smith's ratio rule (Smith 1956) where jobs are sequenced in descending order of the ratios w_l/p_l, ensuring that jobs with a larger weight and a smaller repair time have a higher priority. The parameter, $\rho(l)$, can be viewed as a generalization of the ratio w_l/p_l and characterizes the repair priority of some damaged line l in terms of its own importance as well as the importance of its succeeding nodes in P. Stated differently, $\rho(l)$ can be interpreted as a broad component importance measure for line l. Intuitively, we expect a dispatch rule based on $\rho(l)$ to work better than current industry practice since it takes a more holistic view of the importance of a line and, additionally, has a proven theoretical performance bound. Simulation results presented later confirm that a dispatch rule based on our proposed ρ-factors indeed results in a better restoration trajectory compared to standard industry practices.

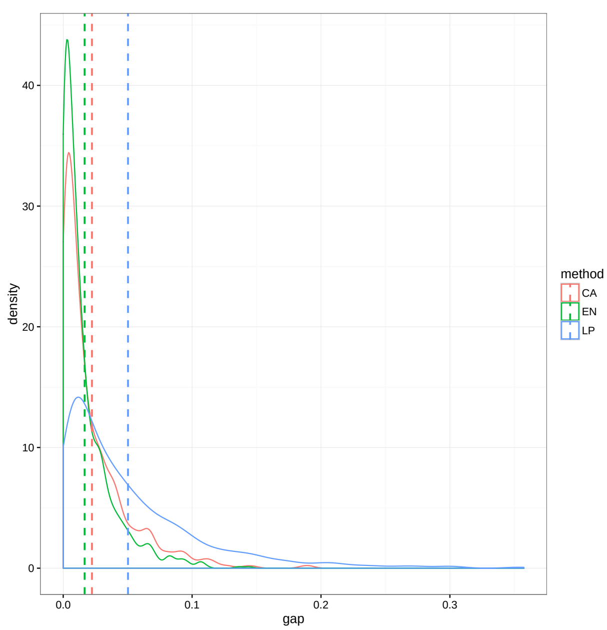
2.7 Case Studies

In this section, we apply our proposed methods to three IEEE standard test feeders of different sizes. We consider the worst case, where all lines are assumed to be damaged. In each case, the importance factor w of each node is a random number between 0 and 1, with the exception of a randomly selected extremely important node with $w = 5$. The repair times are uniformly distributed on integers from 1 to 10. We compare the performances of the three methods, with computational time being of critical concern since restoration activities, in the event of a disaster, typically need to be performed in real time or near real time. All experiments were performed on a desktop with a 3.10 GHz Intel Xeon processor and 16 GB RAM. The ILP formulation was solved using Julia for Mathematical Programming (JuMP) with Gurobi 6.0.

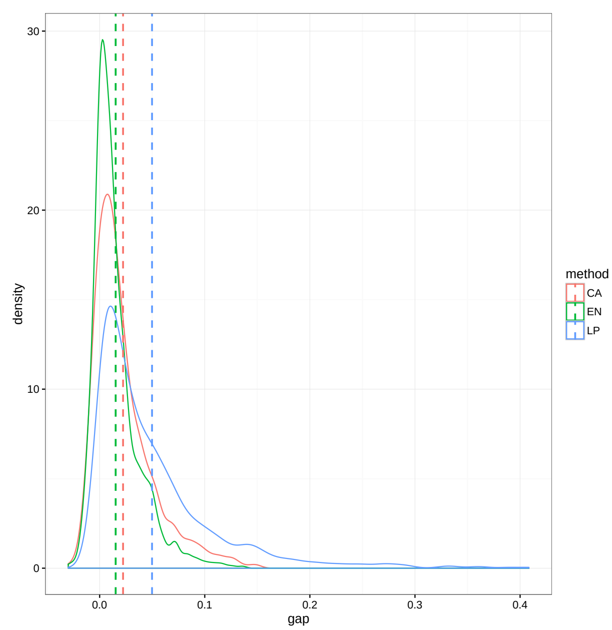
2.7.1 IEEE 13-Node Test Feeder

The first case study is performed on the IEEE 13 Node Test Feeder shown in Fig. 2.1, assuming that the number of repair crews is $m = 2$. Since this distribution network is small, an optimal solution could be obtained by solving the ILP model. We ran 1000 experiments in order to compare the performances of the two heuristic algorithms w.r.t the ILP formulation.

Fig. 2.3 shows the density plots of optimality gaps of LP-based list scheduling algorithm (LP) and the conversion algorithm (CA), along with the better solution from the two (EN). Fig. 2.3a shows the optimality gaps when all repair times are integers. The density plot in this case is cut off at 0 since the ILP solves the problem optimally. Non-integer repair times can be scaled up arbitrarily close to integer values, but at the cost of reduced computational efficiency of the ILP. Therefore, in the second case, we perturbed the integer valued repair times by ± 0.1, which represents a reasonable compromise between computational accuracy and efficiency. The optimality gaps in this case are shown in Fig. 2.3b. In this case, we solved the ILP using rounded off repair times, but the cost function was computed using the (sub-optimal) schedules provided by the ILP model and the actual non-integer repair times. This is why the heuristic algorithms sometimes outperform the ILP model, as is evident from Fig. 2.3b. In both cases, the two heuristic algorithms can solve most of the instances with an optimality gap of less than 10%. Comparing the two methods, we see that the conversion algorithm (CA) has a smaller mean optimality gap, a thinner tail, and a better worst case performance. However, this does not mean that the conversion algorithm is universally superior. In approximately 34% of the problem instances, we have found that the LP-based list scheduling algorithm yields a solution which is no worse than the one provided by the conversion algorithm.



(a) Integer-valued repair time



(b) Integer-±0.1-valued repair time

Figure 2.3: Density plot of optimality gap with means

2.7.2 IEEE 123-Node Test Feeder

Next, we ran our algorithms on one instance of the IEEE 123-Node Test Feeder (Kersting 2001) with $m = 5$. Since solving such problems to optimality using the ILP requires a prohibitively large computing time, we allocated a time budget of one hour. As shown in Table 2.1, both LP and HA were able to find a better solution than the ILP, at a fraction of the computing time.

	Harm	Time(s)
ILP	3.0788×10^3	3600
Conversion Algorithm	2.2751×10^3	$<$1s
Linear Relaxation	2.3127×10^3	24s

Table 2.1: Performance comparison for the IEEE 123-node test feeder

2.7.3 IEEE 8500-Node Test Feeder

Finally, we tested the two heuristic algorithms on one instance of the IEEE 8500-Node Test Feeder medium voltage subsystem (Arritt & Dugan 2010) containing roughly 2500 lines, with $m = 10$. We did not attempt to solve the ILP model in this case. As shown in Table 2.2, it took about more than 60 hours to solve its linear relaxation (which is reasonable since we used the ellipsoid method to solve the LP with exponentially many constraints) and the conversion algorithm actually solved the instance in two and a half minutes.

We also compared the performance of our proposed ρ-factor based dispatch rule to standard industry practices discussed in Section 2.6.5. We assign the same weights to nodes for all three dispatch rules. The plot of network functionality (fraction restored) as a function of time in Fig. 2.4 shows the comparison of functionality trajectories. While the time to full restoration is almost the same for all three approaches, it is clear that our proposed algorithm results in a greater network functionality at intermediate times. Specifically, an additional 10% (approximately) of the network is restored approximately halfway through the restoration process, compared to standard industry practices.

	Harm	Time
Conversion Algorithm	7.201×10^5	150.64 s
Linear Relaxation	7.440×10^5	60.58 h

Table 2.2: Performance comparison for the IEEE 8500-node test feeder

2.7.4 Discussion

From the three test cases above, we conclude that the MILP model, although useful for benchmarking purposes, is feasible in practice only for small networks due to the immense computational time required to solve it to optimality or even near optimality. Even though it can be slow for large problems as shown in the IEEE 8500-Node Test Feeder, the LP-based list scheduling algorithm can serve as a useful secondary tool for moderately sized problems. The conversion algorithm has by far the best overall performance in terms of performance guarantee, solution quality and computational time. Most importantly, we have demonstrated using the IEEE 8500-Node Test Feeder that our proposed dispatch rule, based on the conversion algorithm, can lead to a significantly smaller aggregate harm compared to dispatch rules currently adopted in practice. Stated equivalently, our proposed dispatch rule can improve the resilience of a typical distribution network significantly.

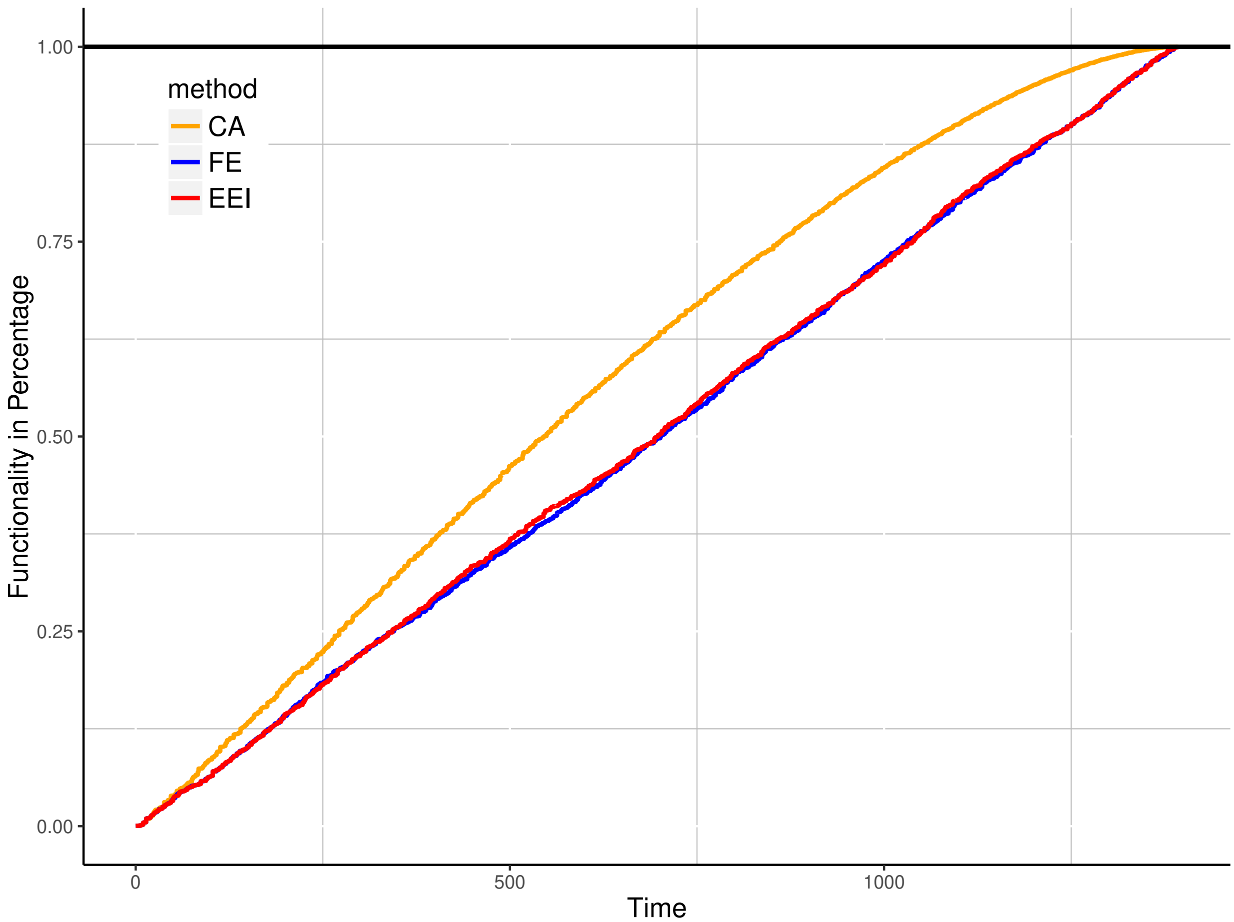


Figure 2.4: Comparison of restoration trajectories: CA stands for our proposed ρ-factor based dispatch rule, FE for FirstEnergy Group's dispatch rule, and EEI for Edison Electric Institute's dispatch rule.

Chapter 3

SCHEDULING POST-DISASTER REPAIRS IN PARTIALLY AUTOMATED RADIAL DISTRIBUTION NETWORKS

3.1 Motivation and Problem Formulation

Distribution automation deployment has been the major strategy of smart grid investment and grid modernization. In Chapter 2, we implicitly consider the case of "fully automated" distribution systems where, unlike the usual definition, the level of automation is only determined by the number of switches. However, we did not include the introduction of distribution automation and reasoning of its impact, in order to focus on the problem of repair scheduling.

3.1.1 Definitions and utilities of switches

According to Office for Eletricity Delivery and Energy Reliability, U.S. Department of Energy (2016), distribution automation (DA) uses digital sensors and switches with advanced control and communication technologies to automate feeder switching; voltage and equipment health monitoring; and outage, voltage, and reactive power management. And it is pointed out as one of the major findings in the report that DA technologies and systems could improve distribution system resilience to extreme weather events, mostly because of the ability to isolate and locate the fault. Such capability is made possible by the utilization of remote controlled switches (RCS) and communication networks. However, it is always too ideal to expect the secondary systems working as normal during and immediately after disasters.

The first thing distribution system operators should do in this case would be to open the (automatic) feeder switch to avoid the danger of electricution by the loose ends in the downstream. Then after finishing the repairs of some damaged components and confirming there is no safety issue, the upperstream switch could be closed, energizing some customers. Therefore, the benefit of RCS is completely undermined by the potential risk of safety and the RCS's is no more than a manual switch. As a result, the level of DA when facing disasters is determined by the number of switches, no matter remote controlled or manual ones.

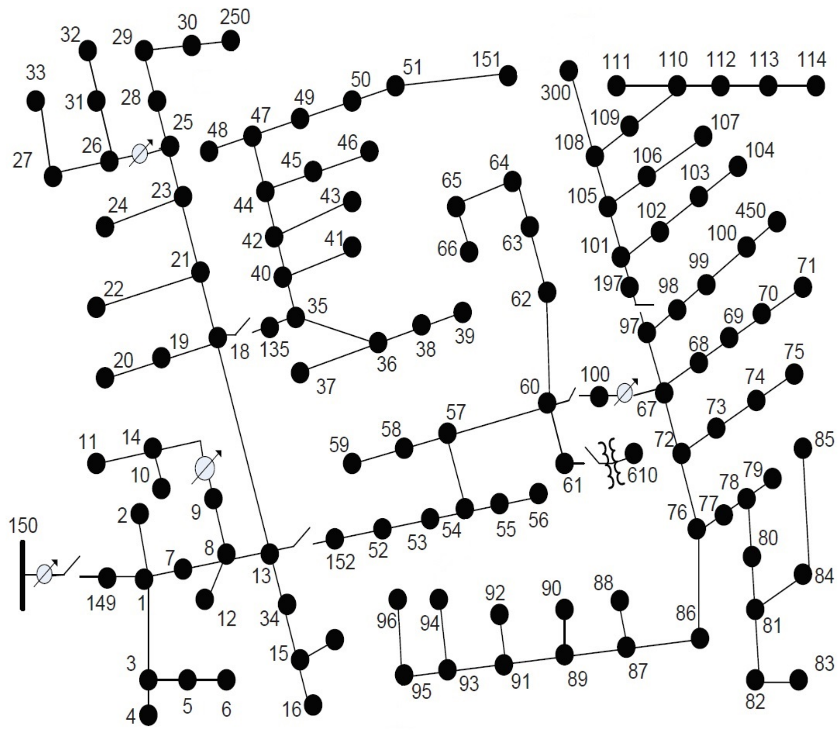
Also by Office for Eletricity Delivery and Energy Reliability, U.S. Department of Energy (2016), the American Recovery and Reinvestment Act (ARRA) of 2009 provided DOE with $7.9 billion to invest in SGIG projects, with more than a quarter on deployment of DA. Such an effort upgrades 6,500 distribution circuits out of more than 200,000 in U.S. Despite the trend towards fully automated distribution networks, it is, as of right now, still practical to consider the case of partial automated distribution networks. With a limited number of switches in the radial distribution networks, the systems are divided into several parts. See a modified IEEE 123 node test feeder in Figure 3.1 as an example. The nodes and lines in the distribution network are divided into 7 sub-network by 6 switches.

3.1.2 Distribution networks modeling

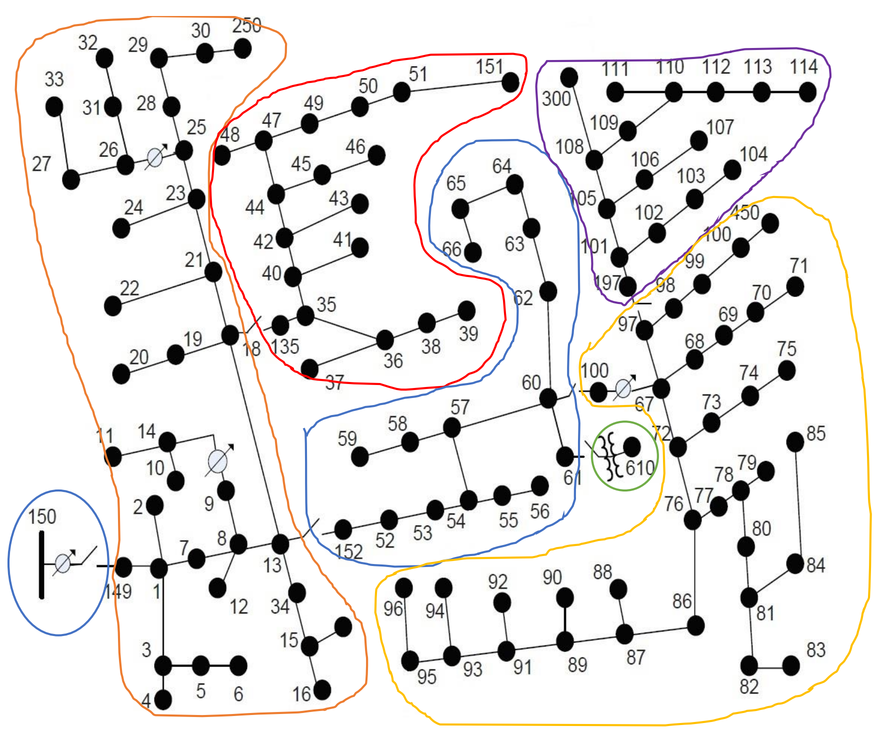
A distribution network can be modeled by a graph G with a set of nodes N and a set of edges L. Let $S \subset N$ represent the set of source nodes which are initially energized and $D = N \setminus S$ represent the set of sink nodes where consumers are located. An edge in G represents a distribution feeder or some other connecting component. We assume that the network topology G is radial, which is a valid assumption for many electricity distribution networks.

3.1.3 Damage modeling

Let L^D denote the sets of damaged edges. Without loss of generality, we assume that there is only one source node in G. If an edge is damaged, all downstream nodes lose power due to lack of electrical connectivity. Each damaged edge $l \in L^D$ has a (potentially) unique repair time p_l. At the operational stage, we assume perfect knowledge of the set L^D and the corresponding repair times.



(a) Before partitioning



(b) After partitioning

Figure 3.1: IEEE 123 node test feeder

3.1.4 Distribution power flow modeling

Instead of a rigorous power flow model, we model network connectivity using a simple network flow model, i.e., as long as a sink node is connected to the source, we assume that all the loads connected to this node can be supplied without violating any security constraint. For simplicity, we treat the three-phase distribution network as if it were a single-phase system. Our analysis could be extended to a three-phase system using a multi-commodity flow model, as in (Yamangil et al. 2015*a*).

In a fully automated system, a node can be energized after there is an energization path from root to it, and therefore the switch could be closed. On the contrary, in this case, the switch cannot be closed until all the lines in the corresponding sub-network are repaired.

Assume m identical repair teams, we redefine the set of damaged edges L^D as the job set $\mathcal{J} = \{J_1, \cdots, J_n\}$, where $J_k = \{j_1, \cdots, j_{n_k}\}$ is the composite job $k \in [1, n]$. The composite job J_k denotes the set of damaged lines that lies within the same sub-network. As a result, each job j could have its own potentially unique completion time C_j while all the job in the same sub-network J has the same energization time E_J. Similarly, we could also derive the relationship between the two.

Proposition 3.1. *Given a schedule with m identical repair teams, let E_J be energization time of group $J \in \mathcal{J}$ and C_j be the completion time of $j \in J$. Then it always satisfies that,*

$$E_J = \max_{J' \preceq J} \max_{j \in J'} C_j \tag{3.1}$$

Proof. This proposition can be shown in a similar manner as Proposition 2.1. □

3.2 LP-based List Scheduling Algorithm

We start with the completion time vector linear relaxation, which is based on the following theorem by Schulz et al. (1996).

Theorem 3.1 ((Schulz et al. 1996)). *The supermodular polyhedron* $Q = \{C \in \mathbb{R}^N : \sum_{j \in A} p_j C_j \geq f(A) \quad \forall A \subset N\}$ *is the convex hull of the completion time vectors in* m *parallel machine scheduling, where*

$$f(A) = \frac{1}{2m}\left(\sum_{j \in A} p_j\right)^2 + \frac{1}{2}\sum_{j \in A} p_j^2 \tag{3.2}$$

Based on the convex hull, we develop the following linear relaxation for this problem,

$$\underset{\mathbf{C},\mathbf{E}}{\text{minimize}} \quad \sum_{J \in \mathcal{J}} \Omega_J E_J \tag{3.3a}$$

$$\text{subject to} \quad C_j \geq p_j, \ \forall j \in \mathcal{J} \tag{3.3b}$$

$$E_J \geq C_j, \ \forall j \in J \in \mathcal{J} \tag{3.3c}$$

$$E_J \geq E_{J'}, \ \forall J' \to J \tag{3.3d}$$

$$\sum_{j \in A} p_j C_j \geq f(A), \ \forall A \subset \mathcal{J} \tag{3.3e}$$

Eqn. (3.3b) constrains the completion time of any damaged line to be lower bounded by its repair time. Eqn. (3.3c) ensures that the group of lines cannot be energized until all lines in the group have been repaired. Eqn. (3.3d) models the soft group precedence constraints. And eqn. (3.3e) characterizes the scheduling polyhedron. And a list scheduling algorithm based on LP midpoints, similar to that in (Queyranne & Schulz 2006), is proposed as follows.

Algorithm 6

Let C^{LP} *denote any feasible solution to the constraints (3.3b) - (3.3e). Define the LP mid points as* $M_j^{LP} := C_j^{LP} - p_j/2, \ \forall j \in \mathcal{J}$. *Create a job priority list by sorting* M_j^{LP}*'s in an ascending order. Whenever a crew is free, assign to it the next job from the priority list. The first* m *jobs in the list are assigned arbitrarily to the* m *crews.*

Proposition 3.2. *Given a schedule with* m *identical repair teams, let* R_j *be the set consisting* *of the lines that starts earlier than j on any machine. Then it always satisfies that,*

$$S_j^H \leq \frac{1}{m} \sum_{i \in R_j} p_i \tag{3.4}$$

Proof. The proof can be trivial. Split the set R into m subsets, corresponding to the schedules of the m crews, i.e., $R = \bigcup_{k=1}^{m} R^k$. Since job j is assigned to the first idle crew and repairs commence immediately, we have:

$$S_j^H = \min \left\{ \sum_{i \in R^k} p_i : k = 1, 2, \dots m \right\} \tag{3.5}$$

$$\leq \frac{1}{m} \sum_{k=1}^{m} \sum_{i \in R^k} p_i = \frac{1}{m} \sum_{i \in R} p_i \,, \tag{3.6}$$

where the inequality follows from the fact that the minimum of a set of positive numbers is upper bounded by the mean. □

Theorem 3.2. *Algorithm 6 is a 2-approximation.*

Proof. The result follows trivially as long as it holds that,

$$E_J^H \leq 2E_J^{LP}, \forall J \in \mathcal{J} \tag{3.7}$$

Define $\mathbf{M^{LP}} := \left[M_j^{LP} : j = 1, 2, \dots, |\mathcal{J}| \right]$. Let $\tilde{\mathbf{M}}^{\mathbf{LP}}$ denote $\mathbf{M^{LP}}$ sorted in ascending order and $\tilde{\mathcal{I}}_j$ denote the position of line $j \in \mathcal{J}$ in $\tilde{\mathbf{M}}^{\mathbf{LP}}$. By Algorithm 6, R_j is equivalent to $\left\{ k : \tilde{\mathcal{I}}_k < \tilde{\mathcal{I}}_j \right\}$.

Note that $M_j^{LP} = E_j^{LP} - p_j/2$, we rewrite eqn. (3.3e) as follows:

$$\sum_{j \in A} p_j M_j^{LP} \geq \frac{1}{2m} \left(\sum_{j \in A} p_j \right)^2 \tag{3.8}$$

Now, for every $j \in \mathcal{J}$, letting $A = R_j$, we have:

$$\left(\sum_{i \in R_j} p_i \right) M_j^{LP} \;\geq\; \sum_{i \in R_j} p_i M_i^{LP} \;\geq\; \frac{1}{2m} \left(\sum_{i \in R_j} p_i \right)^2 \tag{3.9}$$

where the first inequality follows from the fact that $M_j^{LP} \geq M_i^{LP}$ for any $i \in R$. Combining Proposition 3.2 and eqn. (3.9), it follows that $S_j^H \leq 2M_j^{LP}$. Consequently,

$$C_j^H = S_j^H + p_j \leq 2M_j^{LP} + p_j = 2C_j^{LP} \tag{3.10}$$

Then by Proposition 3.1, for all $J \in \mathcal{J}$,

$$E_J^H = \max_{J' \preceq J} \max_{j \in J'} \quad C_j^H \tag{3.11}$$

$$\leq 2\max_{J' \preceq J} \max_{j \in J'} \quad C_j^{LP} \tag{3.12}$$

$$= 2E_J^{LP} \tag{3.13}$$

And this completes the proof. □

Remark 3.1. *There can be exponentially many constraint (3.3e) in LP relaxation. The separation problem for these inequalities can be solved in polynomial time using the ellipsoid method (Schulz et al. 1996). However, numerical study on IEEE 8500 node test feeders in Chapter 2 shows this algorithm could be really slow for network with severe damages.*

This problem, compared with the one considering a fully automated distribution network, should be slightly easier to solve in the sense that the solution space is narrowed with lines forming composite jobs by switches. However, the number of constraint (3.3e) is not reduced.

3.3 A Conversion Algorithm

In this section, we first take a step back and analyze the repair sequencing as it would be done with a single crew because this provides important insights into the general structure of the multi-crew scheduling problem. And then we will convert the one-crew sequence into a multi-crew schedule with bounded performance.

3.3.1 The case with one repair team

We first recognize that the case with one repair team is equivalent to an existing scheduling problem $1 \mid outtree \mid \sum_j w_j C_j$.

Lemma 3.1. *Single crew repair and restoration scheduling in distribution networks with limited number of switches is equivalent to* $1 \mid outtree \mid \sum w_j C_j$*, where the outtree precedences are given in the soft precedence constraint graph* P*.*

Proof. We plan to show this in three steps.
Step 1, interchanging positions within the same group does not affect the scheduling objective. This is obvious since all the nodes in the group cannot be energized until after the last line is repaired. It does not matter which that last line is.
Step 2, the lines within the same group should be sequenced without interruption. It can be shown by swapping technique. Suppose that group $J = \{j_1, j_2, \cdots, j_n\}$ was interrupted by some line j' in a optimal sequence, i.e., the optimal sequence contains $\{j_1, \cdots, j_k, j', j_{k+1}, \cdots, j_n\}$. We will consider two cases. If j' is the last job of its own group J', then $\{j', j_1, \cdots, j_n\}$ has an objective no larger than the original one since E_J along with all other energization times remains the same and $E_{J'}$ decreases. If j' is the last job of its own group J', then $\{j_1, \cdots, j_n, j'\}$ has an objective no larger than the original one since E_J decreases and $E_{J'}$ and all others does not change.

This indicates that the problem reduced to a single-crew problem of composite jobs. The processing time is the sum of repair times of each line within the group and the same for the weight, i.e.,

$$P_J = \sum_{j \in J} p_j; \;\; \Omega_J = \sum_{j \in J} w_j. \tag{3.14}$$

Step 3, now this problem is reduced to the problem in fully automated distribution networks and Proposition 2.3 in Chapter 2 will complete the proof. □

3.3.2 A conversion algorithm and its performance bound

An algorithm for converting the optimal single crew sequence to a multiple crew schedule is given in Algorithm 7. We now prove that it is a $\left(2 - \frac{1}{m}\right)$ approximation algorithm. We start with two lemmas that provide lower bounds on the minimal harm for an m-crew schedule, in terms of the minimal harms for single crew and ∞-crew schedules. Let $H^{1,*}$, $H^{m,*}$ and $H^{\infty,*}$ denote the minimal harms when the number of repair crews is 1, some arbitrary m ($2 \le m < \infty$), and ∞ respectively.

Algorithm 7 Algorithm for converting the optimal single crew schedule to an m-crew schedule

Treat the optimal single crew repair sequence as a priority list, and, whenever a crew is free, assign to it the next job from the list. The first m jobs in the single crew repair sequence are assigned arbitrarily to the m crews.

Lemma 3.2. $H^{m,*} \ge \frac{1}{m} H^{1,*}$

Proof. Given an arbitrary m-crew schedule S^m with harm H^m, we first construct a 1-crew repair sequence, S^1. We do so by sorting the energization times of the groups of the damaged lines in S^m in ascending order and assigning to S^1 the corresponding sorted sequence of groups with arbitrary sequences within each group. Ties, if any, are broken according to (soft) precedence constraints or arbitrarily if there is none. Let C_J denote the completion of group J, aka, the completion of the last job in J. By construction, for any two groups of damaged lines J and J' with precedence constraint $J' \prec J$, the completion time of J' must be strictly smaller than that J in the constructed S^1, i.e., $C^1_{J'} < C^1_J$. Additionally, $C^1_J = E^1_J$ for any $J \in \mathcal{J}$ because the completion and energization times of lines are identical since the sequence respects the soft precedence constraints in P.

Next, we claim that $E^1_J \le mE^m_J$ for any fixed $J \in \mathcal{J}$, where E^1_i and E^m_i are the energization times of line i in S^1 and S^m respectively. In order to prove it, we first observe that:

$$E^1_J = C^1_J = \sum_{\{J':\ E^m_{J'} \le E^m_J\}} \sum_{j \in J'} p_j \le \sum_{\{J':\ C^m_{J'} \le E^m_J\}} \sum_{j \in J'} p_j \,, \tag{3.15}$$

where the second equality follows from the manner we constructed S^1 from S^m and the inequality follows from the fact that $C^m_J \le E^m_J \Rightarrow \{J :\ E^m_{J'} \le E^m_J\} \subseteq \{J :\ C^m_{J'} \le E^m_J\}$ for any m-crew schedule. Next, we split the set $\{j : j \in J', \forall J' \text{ satisfies } C^m_{J'} \leq E^m_J\} := R_J$ (say) into m subsets, corresponding to the schedules of the m crews in S^m, i.e., $R_J = \bigcup_{k=1}^m R^k_J$, where R^k_J is a subset of the jobs in R_J that appear in the k^{th} crew's schedule. It is obvious that the sum of the repair times of the lines in each R^k_J can be no greater than E^m_J. Therefore,

$$E^1_J \leq \sum_{j \in R_J} p_j = \sum_{k=1}^{m} \left(\sum_{j \in R^k_J} p_j \right) \leq mE^m_J \tag{3.16}$$

Proceeding with the optimal m-crew schedule $S^{m,*}$ instead of an arbitrary one, it is easy to see that $E^1_J \leq mE^{m,*}_J$. The lemma then follows straightforwardly.

$$H^{m,*} = \sum_{J \in \mathcal{J}} \Omega_J E^{m,*}_J \geq \sum_{J \in \mathcal{J}} \Omega_J \frac{1}{m} E^1_J \tag{3.17}$$

$$= \frac{1}{m} H^1 \geq \frac{1}{m} H^{1,*} \tag{3.18}$$

□

Lemma 3.3. $H^{m,*} \geq H^{\infty,*}$

Proof. This is intuitive and obvious, since the harm is minimized when the number of repair crews is at least equal to the number of damaged lines.

In the ∞-crew case, every job can be assigned to one crew. For any damaged line $j \in J \in \mathcal{J}$, $C^\infty_j = p_j$ and $E^\infty_J = \max_{J' \preceq J} \max_{j \preceq J'} p_j$. Also, $C^{m,*}_j \geq p_j = C^\infty_j$. As a result, $E^{m,*}_J = \max_{J' \preceq J} \max_{j \preceq J'} C^{m,*}_j \geq \max_{J' \preceq J} \max_{j \preceq J'} p_j = E^\infty_j$. Therefore:

$$H^{m,*} = \sum_{J \in \mathcal{J}} \Omega_J E^{m,*}_J \geq \sum_{J \in \mathcal{J}} \Omega_J E^\infty_J = H^{\infty,*} \tag{3.19}$$

□

Theorem 3.3. *Let E^m_J be the energization time of line j after the conversion algorithm is applied to the optimal single crew repair schedule. Then, $\forall J \in \mathcal{J}$, $E^m_J \leq \frac{1}{m} E^{1,*}_J + \frac{m-1}{m} E^{\infty,*}_J$.*

Proof. Let S_j^m and C_j^m denote respectively the start and energization times of some line $j \in \mathcal{J}$ in the m-crew repair schedule, S^m, obtained by applying Algorithm 7 to the optimal 1-crew sequence, $S^{1,*}$. Also, let $\mathcal{I}_j$ denote the position of line j in $S^{1,*}$ and then $R_j = \{k : \mathcal{I}_k < \mathcal{I}_j\}$. Now by Proposition 3.2:

$$\begin{aligned} C_j^m &= S_j^m + p_j && (3.20)\\ &\leq \frac{1}{m} \sum_{i \in R_j} p_i + p_j && (3.21)\\ &= \frac{1}{m} \sum_{i \in R_j \cup \{j\}} p_i + \frac{m-1}{m} p_j && (3.22)\\ &= \frac{1}{m} C_j^{1,*} + \frac{m-1}{m} p_j && (3.23) \end{aligned}$$

and by Proposition 3.1,

$$\begin{aligned} E_J^m &= \max_{J' \preceq J} \max_{j \in J'} C_j && (3.24)\\ &\leq \max_{J' \preceq J} \max_{j \in J'} \frac{1}{m} C_j^{1,*} + \max_{J' \preceq J} \max_{j \in J'} \frac{m-1}{m} p_j && (3.25)\\ &= \frac{1}{m} C_J^{1,*} + \frac{m-1}{m} \max_{J' \preceq J} \max_{j \in J'} p_j && (3.26)\\ &= \frac{1}{m} E_J^{1,*} + \frac{m-1}{m} E_J^{\infty} && (3.27) \end{aligned}$$

□

Theorem 3.4. *The conversion algorithm is a* $\left(2 - \frac{1}{m}\right)$*-approximation.*

Proof.

$$
\begin{aligned}
H^m &= \sum_{J \in \mathcal{J}} w_J E_J^m && (3.28)\\
&\leq \sum_{J \in \mathcal{J}} w_J \left(\frac{1}{m} E_J^{1,*} + \frac{m-1}{m} E_J^{\infty} \right) && (3.29)\\
&= \frac{1}{m} \sum_{J \in \mathcal{J}} w_J E_J^{1,*} + \frac{m-1}{m} \sum_{J \in \mathcal{J}} w_J E_J^{\infty} && (3.30)\\
&= \frac{1}{m} H^{1,*} + \frac{m-1}{m} H^{\infty,*} && (3.31)\\
&\leq \frac{1}{m} (m H^{m,*}) + \frac{m-1}{m} H^{m,*} && (3.32)\\
&= \left(2 - \frac{1}{m} \right) H^{m,*} && (3.33)
\end{aligned}
$$

□

Remark 3.2. *Algorithm 7 in some sense imposes a group precedence constraint defined as follows.*

If there is a group precedence constraint $J_{k_1} \prec J_{k_2}$ *for any* $k_1, k_2 \in \{1, \cdots n\}$, *it must satisfy that for any* $j_1 \in J_{k_1}$ *and* $j_2 \in J_{k_2}$, $S_{j_1} \leq S_{j_2}$.

However, such a constraint is not always satisfied in the optimal solution. For example, in Fig. (3.2), where 1, $1'$ *and* $1''$ *belongs to a group* 1 *and* 2 *is a standalone job, the schedule at the bottom will outperforms if* 2 *is the child of group* 1 *and* $\Omega_2 \gg \Omega_1$.

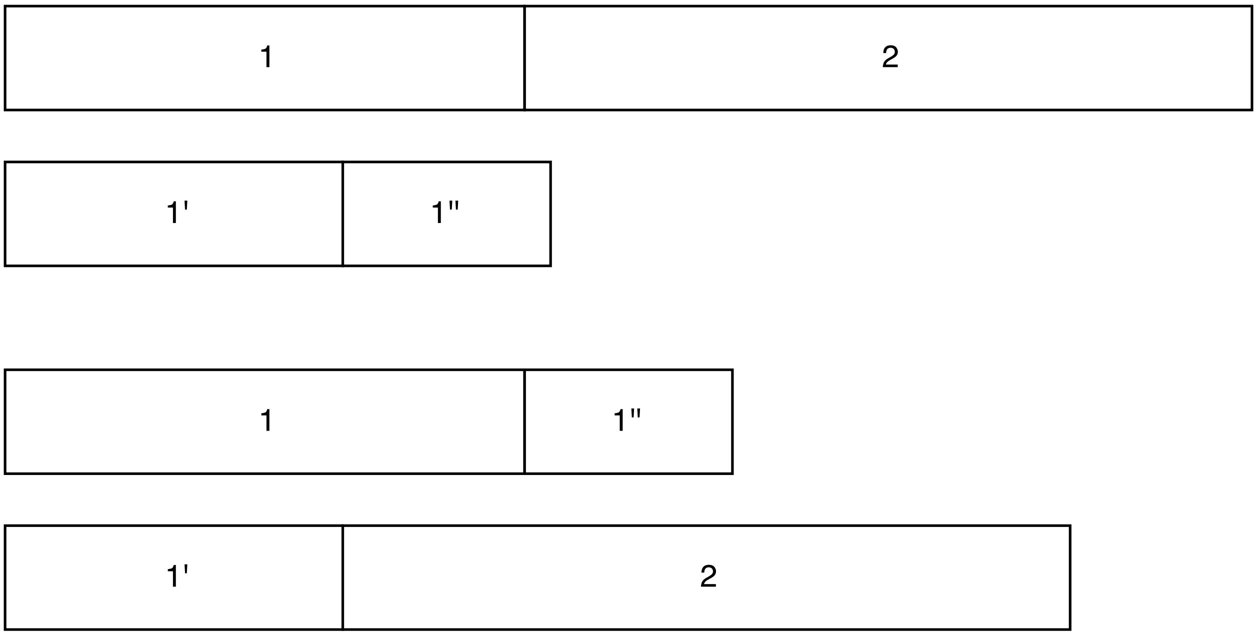


Figure 3.2: Case when group precedence is not optimal

Remark 3.3. *As mentioned in proof of Lemma 3.1, interchanging positions within the same group does not affect the scheduling objective in sequencing repairs with one repair team. And by Algorithm 7, interchanging positions within the same group does not affect the performance guarantee. In other words, Algorithm 7 does not specify a schedule but rather a family of schedules.*

Remark 3.4. *Recall that Graham's* List *scheduling algorithm with arbitrary priority list is a* $2 - \frac{1}{m}$ *approximation. Therefore,* $2 - \frac{1}{m}$ *is the tightest bound possible for Algorithm 7.*

Remark 3.5. *Although the performance bound could not be improved by investigating the schedules within the groups, the objectives might be improved for some cases.*

Chapter 4

HANDLING UNCERTAINTIES IN POST-DISASTER REPAIR SCHEDULING

4.1 Stochastic Scheduling Policy

In deterministic or offline scheduling, all problem data is known beforehand and a scheduling result assigns a job to a machine at a specified time. However, in many practical situations, we might not have complete information about problem data including processing times of the jobs, the number of machines or even the number of jobs to be scheduled. Stochastic scheduling is one of the relaxed models, in which the processing time p_j of job j is not known with certainty until job j finishes. To put it mathematically, the processing time of job j is a random variable $\mathcal{P}_j$ and its realization p_j is known upon completion. The information we have is the distributions of $\mathcal{P}_j$, but in some approximation algorithms only first and second order moments are needed. $\mathcal{P}_j$'s are pairwise independent. The resulting start times, completion times and function times are also random variables $\mathcal{S}_j$, $\mathcal{C}_j$ and $\mathcal{F}_j$. As typical in stochastic optimization, one cannot expect a scheduling policy minimize the objective for any realization of processing times, but rather aims to minimize the objective in expectation. Notice that, due to heavy notation of the expected sign $\mathbb{E}$, function time in this context is equivalent to energization time previously defined and they might be used interchangably. This would also allow us to use this model in a broader setting outside of the power systems.

Moreover, the result of scheduling with uncertainty processing times will not be the same as deterministic scheduling, since no specific time can be determined for a job to start. So they call the dynamic allocation of jobs on machines with incomplete information a scheduling policy (Möhring et al. 1999). The theoretical notion and characterization of such a policy is proposed by Radermacher (1981), Möhring et al. (1984, 1985) and also see (Uetz 2001) for a detailed review. To put it simple, a scheduling policy can make decisions at a certain time t, but only based on the observation up to t and a priori information. And it must not anticipate any information about the future. At this point, we only focus on static list scheduling policy, where a priority list of jobs is obtained ahead of time and will not change during the process of scheduling. With no precedence constraints, the list scheduling policy could be as simple as assigning the idle machine the top of the list. We present 2 ways of coming up with the list, corresponding to the 2 list scheduling algorithms in the Chapter 2.

4.2 LP-based List Scheduling Policy

The valid inequalities (2.15) are extended to stochastic parallel machine scheduling by Möhring et al. (1999),

Theorem 4.1 (Möhring et al. (1999)). *Let* Π *be any policy for stochastic parallel machine scheduling. Then inequalities below (4.1) are valid for the corresponding vector of expected completion times* $\mathbb{E}[\mathcal{C}^{\Pi}]$.

$$\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]\mathbb{E}[\mathcal{C}_j^{\Pi}] \geq \frac{1}{2m}\left(\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]\right)^2 + \frac{1}{2}\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]^2 - \frac{m-1}{2m}\sum_{j\in A}\mathrm{Var}[\mathcal{P}_j] \quad \forall A \subset N \tag{4.1}$$

And with an additional assumption on the second moments of all processing time distributions, the above result can be further simplified.

Corollary 4.2 (Möhring et al. (1999)). *Let* Π *be any policy for stochastic parallel machine scheduling. If* $\mathrm{Var}[\mathcal{P}_j]/\mathbb{E}[\mathcal{P}_j]^2 \leq \Delta$, *then inequalities below (4.2) are valid for the corresponding vector of expected completion times* $E[\mathcal{C}^{\Pi}]$.

$$\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]\mathbb{E}[\mathcal{C}_j^{\Pi}] \geq \frac{1}{2m}\left(\left(\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]\right)^2 + \sum_{j\in A}\mathbb{E}[\mathcal{P}_j]^2\right) - \frac{(m-1)(\Delta-1)}{2m}\left(\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]^2\right) \quad \forall A \subset N \tag{4.2}$$

Define a set function $f : 2^L \to \mathbb{R}$:

$$f(A) := \frac{1}{2m}\left(\left(\sum_{j\in A}\mathrm{E}[\mathcal{P}_j]\right)^2 + \sum_{j\in A}\mathrm{E}[\mathcal{P}_j]^2\right) - \frac{(m-1)(\Delta-1)}{2m}\left(\sum_{j\in A}\mathbb{E}[\mathcal{P}_j]^2\right) \tag{4.3}$$

In Section 2.3.1, we introduce the soft precedence constraints based on the energization time vector. For each realization, soft precedence constraints are satisfied. Therefore, soft precedence constraints in the expected sense will always be satisfied and turns out to be a relaxation. Combined with the Corollary 4.2, the LP relaxation can be written as:

$$\begin{aligned}
\underset{F}{\text{minimize}} \quad & \sum_{j\in L} w_j F_j && \text{(4.4a)}\\
\text{subject to} \quad & F_j \geq \mathbb{E}[\mathcal{P}_j],\ j\in L && \text{(4.4b)}\\
& F_j \geq F_i,\ (i,j)\in\mathcal{P} && \text{(4.4c)}\\
& \sum_{j\in A}\mathbb{E}[\mathcal{P}_j]F_j \geq f(A) \quad \forall A\subset L && \text{(4.4d)}
\end{aligned}$$

where variable F_j represents the expectation of function time and $\mathcal{P}$ stands for the set of soft precedence constraints. Although there is an exponential number of constraints, the separation problem for these inequalities can be solved in polynomial time using the ellipsoid method (Schulz et al. 1996).

Let F^{LP} denote any feasible solution to the constraint set (4.4b) - (4.4d) of this LP. Now we use the list scheduling algorithm with the list defined by sorting the jobs in non-decreasing order of F_j^{LP}'s. Assume without loss of generality that $F_1^{LP} \leq \cdots \leq F_n^{LP}$.

Proposition 4.1. *Let $\mathcal{S}^{\Pi}$ and $\mathcal{C}^{\Pi}$ denote the random vector of starting times and completion times, respectively, of the scheduling constructed by the list scheduling algorithm. Then*

$$\mathbb{E}[\mathcal{F}_j^{\Pi}] \leq \left(2 - \frac{1}{m} + \max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_j^{LP} \tag{4.5}$$

Proof. Consider any job j. From the algorithm, job $1, \cdots, j-1$ are scheduled in order with no idle time in between, which means all machines are busy before the last job starts. Then for any realization p of $\mathcal{P}$ and its scheduling result by policy Π,

$$S_j \leq \frac{1}{m}\sum_{i=1}^{j-1} p_i \tag{4.6}$$

Since it holds for every realization, then

$$\mathbb{E}[\mathcal{S}_j^\Pi] \leq \frac{1}{m}\sum_{i=1}^{j-1} \mathbb{E}[\mathcal{P}_i] \tag{4.7}$$

Fix j and let A be the set of $\{1, 2, \cdots, j\}$. Combined with the fact that $F_1^{LP} \leq F_2^{LP} \leq \cdots \leq F_n^{LP}$,

$$(\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i])F_j^{LP} \geq \sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]F_i^{LP} \geq \frac{1}{2m}\left(\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]\right)^2 + \frac{m-\Delta(m-1)}{2m}\left(\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]^2\right) \tag{4.8}$$

Divide both sides by $\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]$,

$$F_j^{LP} \geq \frac{1}{2m}\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i] + \frac{m-\Delta(m-1)}{2m}\frac{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]^2}{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]} \tag{4.9}$$

When $\Delta \leq m/(m-1)$, the second term on the right-hand side of (4.9) is nonnegative, so

$$F_j^{LP} \geq \frac{1}{2m}\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i] \tag{4.10}$$

When $\Delta > m/(m-1)$ on the other hand, since

$$\frac{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]^2}{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]} \leq \max_{i=1,\cdots,j} \mathbb{E}[\mathcal{P}_i] \leq F_j^{LP} \tag{4.11}$$

and therefore,

$$F_j^{LP} \geq \frac{1}{2m}\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i] + \frac{m-\Delta(m-1)}{2m}\frac{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]^2}{\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i]} \tag{4.12}$$

$$\geq \frac{1}{2m}\sum_{i=1}^{j} \mathbb{E}[\mathcal{P}_i] + \frac{m-\Delta(m-1)}{2m}F_j^{LP} \tag{4.13}$$

Combining (4.10) and (4.13),

$$\frac{1}{m}\sum_{i=1}^{j}\mathbb{E}[\mathcal{P}_i] \leq \left(1+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_j^{LP} \tag{4.14}$$

Then the expected completion time under policy Π, given by the sum of the expected starting time and the expected repair time, could be also bounded:

$$\mathbb{E}[\mathcal{C}_j^{\Pi}] = \mathbb{E}[\mathcal{S}_j^{\Pi}] + \mathbb{E}[\mathcal{P}_j] \leq \frac{1}{m}\sum_{i=1}^{j}\mathbb{E}[\mathcal{P}_i] + (1-\frac{1}{m})\mathbb{E}[\mathcal{P}_j] \tag{4.15}$$

$$\leq \left(1+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_j^{LP} + (1-\frac{1}{m})F_j^{LP} \tag{4.16}$$

$$\leq \left(2-\frac{1}{m}+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_j^{LP} \tag{4.17}$$

Finally, based on the definition of the soft precedence constraints in the expected sense,

$$\mathbb{E}[\mathcal{F}_j^{\Pi}] = \max_{i\preceq_S j}\mathbb{E}[\mathcal{C}_i^{\Pi}] \leq \max_{i\preceq_S j}\left(2-\frac{1}{m}+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_i^{LP} \tag{4.18}$$

$$= \left(2-\frac{1}{m}+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)F_j^{LP} \tag{4.19}$$

where the last equality follows trivially from the definition of a soft precedence constraint in the expected sense. □

Theorem 4.3. *The scheduling policy using LP relaxations is a* $\left(2-\frac{1}{m}+\max\left\{1, \frac{m-1}{m}\Delta\right\}\right)$*-approximation.*

Proof. Proposition 4.1 and the fact that linear program (4.4) is a relaxation of the stochastic scheduling problem concludes the proof. □

Chapter 5

DISTRIBUTION SYSTEMS HARDENING AGAINST NATURAL DISASTERS

5.1 Introduction

Natural disasters have caused major damage to electricity distribution networks and deprived homes and businesses of electricity for prolonged periods, for example Hurricane Sandy in November 2012 (NERC 2014), the Christchurch Earthquake in February 2011 (Kwasinski et al. 2014) and the June 2012 Mid-Atlantic and Midwest Derecho (Infrastructure Security and Energy Restoration, Office of Electricity Delivery and Energy Reliability, U.S. Department of Energy 2012). Estimates of the annual cost of power outages caused by severe weather between 2003 and 2012 range from $18 billion to $33 billion on average(Executive Office of the President 2013). Physical damage to grid components must be repaired before power can be restored (The GridWise Alliance 2013, NERC 2014). On the operational side, approaches have been proposed for scheduling the available repair crews in order to minimize the cumulative duration of customer interruption, which reduces the harm done to the affected community (Nurre et al. 2012, Coffrin & Van Hentenryck 2014). On the planning side, Kwasinski et al. (2014) reported that facilities that had been upgraded or hardened in Christchurch, at a cost of $5 million, remained serviceable immediately after the September 2010 earthquake and saved approximately $30 to $50 million in subsequent repairs. Hardening minimizes the potential damages caused by disruptions, thereby facilitating restoration and recovery efforts, and the time it takes for the infrastructure system to resume operation (Omer 2013). However, as indicated in (Rollins 2007), the difficulty of hardening does not lie in the design or construction of a hardened system, rather in the ability to quantify the expected performance improvement so that rational decisions can be made regarding increased cost versus potential future benefit.

Recall that we have defined a measure of resilience, harm H, in Chapter 1,

$$H = \sum_n w_n T_n, \tag{5.1}$$

where w_n can be interpreted as the contribution of node n to the overall loss in functionality of the system or the importance of node n and T_n is the time to restore node n. In Chapter 2, we approximate this quantity with

$$H = \sum_n w_n E_n \tag{5.2}$$

by relaxing the operational voltage constraints and load flow models. *The objective for operational problems is to minimize the measure $\sum_n w_n T_n$, given a specific disaster scenario, while the objective for planning problems is to minimize $\sum_n w_n T_n$ in an expected sense, where the expectation is over all possible disaster scenarios in consideration.*

5.1.1 Literature review

Defending critical infrastructures at the transmission level has been a major research focus over the past decade (Brown et al. 2006, Bier et al. 2007, Yuan et al. 2014). In general, this research adopted the setting of Stackelberg game and formulated the problem with a tri-level defender-attacker-defender model. Such a model assumes that the attacker has perfect knowledge of how the defender will optimally operate the system after the attack and the attacker manipulates the system to its best advantage.

In recent years, several researchers have investigated different methods for distribution systems hardening, but most focus solely on the robustness, i.e., worst-case load shedding at the onset of disaster. Of note, a resilient distribution network planning problem (RDNP) was proposed by Yuan et al. (2016) to coordinate the hardening and distributed generation resource allocation. A tri-level defender-attacker-defender model is studied, in which the defender (hardening planner) selects a network hardening plan in the first stage, the attacker (natural disaster) disrupts the system with an interdiction budget, and finally, the defender (the distribution system operator) reacts by controlling DGs and switches in order to minimize the shed load. This model is improved by Ma et al. (2016) by considering the investment cost and by eliminating the assumption that enhanced components should remain intact during any disaster scenario. Another direction of research enforces chance constraints on the loss of critical loads and normal loads respectively (Yamangil et al. 2015*b*, Nagarajan et al. 2016). A two-stage stochastic program and heuristic solution of hardening strategy were proposed by Romero et al. (2015), specifically for earthquake hazards, under the assumption that the repair times for similar types of components follow an uniform distribution, which simplifies the problem to a certain extent.

5.1.2 Our approach

To the best of our knowledge, this paper is the first to consider the restoration process in conjunction with hardening. Our approach can be seen as a two-stage stochastic problem. The first stage selects from the set of potential hardening choices and determines the extent of hardening to maximize the expected resilience measure R, while the second stage solves the operational problem in each possible scenario by optimizing the sequence of repairs given the hardening results. In the operational problem, we consider scheduling post-disaster repairs in distribution network with parallel repair crews. This issue will be discussed in more detail in Section 5.6. Since an ideal formulation of the problem is hard to solve and also turns out impractical, we developed a deterministic single crew approximation with a heuristic approach to solve the hardening problem. Wang et al. (2015) make a distinction between hardening activities and resiliency activities which are focused on the effectiveness of humans post-disaster. By using only one repair crew in the operational problem, we can also focus on the effects of network structure and components, and reduce the reliance on resourcefulness (i.e. the number of repair crews available).

The rest of the paper is organized as follows. In Section 5.2, we briefly review the operational problem of scheduling post-disaster repairs in distribution networks with multiple repair crews and discuss an MILP model for solving the single crew repair sequencing problem. In Section 5.3, we formulate the problem of distribution system hardening against natural disasters and model it as a stochastic optimization problem, followed by a deterministic reformulation (Section 5.4) and single crew approximation (Section 5.5). In Section 5.6, we motivate why we believe it is important to consider the restoration process (operational phase) in the hardening problem (planning phase) and develop the so called 'restoration process aware hardening problem'. Two solution methods, an MILP formulation and an iterative heuristic algorithm, are also discussed in this section. The performance of these methods is validated by various case studies on various standard IEEE test feeders in Section 5.7.

5.2 An MILP approach for Optimal Post-disaster Sequencing

With only one repair crew, the damaged components must be repaired one by one, so there can be L^D decisions to make, one at each time stage. The duration of each stage depends on the repair time of the component. We use two sets of binary decision variables. The first set of decision variables is denoted by $\{x_l^t\}$, where $x_l^t = 1$ if edge l is repaired at time stage t and is equal to 0 otherwise. The second set of decision variables is denoted by $\{u_i^t\}$, where $u_i^t = 1$ if node i is energized at the end of time stage t and is equal to 0 otherwise. Let T denote the restoration time horizon and h^t denote the harm till time stage t. The MILP model for minimizing the aggregate harm is shown below:

$$\min_{x,u} \quad \sum_{t=1}^{T} h^t \tag{5.3a}$$

$$\text{s.t.} \quad u_i^0 = 1,\ \forall i \in S \tag{5.3b}$$

$$\sum_{t=1}^{T} u_i^t = 1,\ \forall i \in D \tag{5.3c}$$

$$\sum_{l \in L^D} x_l^t = 1,\ \forall t \in [1, T] \tag{5.3d}$$

$$\sum_{\tau=0}^{t-1} u_i^\tau + u_{t(l)}^t - 2x_l^t \geq 0, l \in L^D, i \in Ne(t(l)), \forall t \tag{5.3e}$$

$$d^0 = 0 \tag{5.3f}$$

$$d^t \geq d^{t-1} + p_l \times x_l^t,\ \forall t \in [1, T],\ \forall l \in L^D \tag{5.3g}$$

$$h^t \geq w_j u_j^t d^t,\ \forall j \in D,\ \forall t \in [1, T] \tag{5.3h}$$

The first set of constraints binds the two sets of decision variables. Constraints (5.3b) and (5.3c) specify that all source nodes be energized initially and all sink nodes be energized by time T. Constraint (5.3d) requires that only one damaged edge be chosen for repair at any time stage. Constraint (5.3e) requires that when an edge $l \in L^D$ is chosen for repair at time stage t, i.e., $x_l^t = 1$, both $u_{t(l)}^t$ and $\sum_{\tau=0}^{t-1} u_i^\tau$ must be equal to 1. In other words, if the tail node of edge l, $t(l)$, is to be energized at time stage t, at least one of its neighbors in the damaged component graph G', denoted by $Ne(t(l))$, must have been energized at some previous time stage. This constraint follows directly from the outtree precedences in Lemma 2.1.

The second set of constraints connects the aggregate harm with the decision variables. The intermediate variable d^t models the aggregate restoration time just prior to time stage t. Constraint (5.3f) initializes the aggregate restoration time to 0 while constraint (5.3g) requires that the difference $d^t - d^{t-1}$, for some t, be at least the repair time of the edge being repaired at time t. Finally, constraint (5.3h) models the t^{th} stage harm if l is the edge being repaired at time stage t, whose tail node is j. Note that constraint (5.3h) can be easily linearized using the big-M method, details of which are omitted.

5.3 The Hardening Problem: Formulation

5.3.1 Damage modeling

As mentioned above, damages are modeled by repair time vectors associated with network components. Since no *a priori* exact information about the damages is available at the planning stage, we model the repair times as a random vector $\vec{\mathcal{P}}$. The uncertainties are twofold: the possible scenarios of natural disasters that planners want to take into account and the uncertain damages to components caused by a specific disaster. The distribution of $\vec{\mathcal{P}}$ can be a mixture of a Bernoulli distribution which represents the probability of damage and a (possibly) continuous distribution of repair time, such as the exponential (Patton 1979) or log-normal distribution (Billinton & Wojczynski 1985). Mixed distributions, usually do not admit a closed-form expression of their distribution functions. In our work, we do not assume any knowledge of the distribution function, except for knowledge of the first moment $\mathbb{E}[\vec{\mathcal{P}}]$.

Some planners tend to use the sample average approximation (SAA) methods (Kleywegt et al. 2002) by considering a limited set of component damage scenarios, which are either defined by users or drawn from a probabilistic model, as by Yamangil et al. (2015*b*), Nagarajan et al. (2016). It is known that SAA methods converge to the optimal solution as the sample size goes to infinity. However, SAA methods require that the selected scenarios be typical and right on target, or the sample averaging needs to be performed over a large number of cases.

5.3.2 Hardening options and costs

In practice, multiple hardening actions are usually available for each network component. For example, hardening an edge can involve some combination of vegetation management, pole reinforcement, undergrounding, enhanced pole guying (Ma et al. 2016). Typically, the goal of hardening a component is to lower the probability of its failure in the event of a disaster. However, since we are interested in maximizing the resilience of the system, or equivalently, minimizing the aggregate harm, simply lowering the probability of failure of a component is not sufficient. Since the aggregate harm is a function of the restoration times of the nodes, which in turn depend on the repair times of the damaged components (and the repair schedule), hardening a component can only be beneficial if it leads to a corresponding reduction in the repair time of that component.

In this paper, we assume that there is a finite set of hardening strategies for each edge l, which we denote by K_l. Each such strategy can be some combination of several disjoint hardening actions. We require that the hardening process select one strategy from the set K_l. Let $\vec{p} = \{p_l\}$, where p_l is the 'expected repair time' of component l before hardening, $\Delta\vec{p} = \{\Delta p_{lk}\}$, where Δp_{lk} is the 'expected reduction in the repair time' of component l due to hardening strategy $k \in K_l$, and c_{lk} be the cost of implementing hardening strategy k on edge l. We make the following assumption on the relationship between c_{lk} and Δp_{lk}:

Assumption 5.1. *For any two hardening strategies* $(k_1, k_2) \in K_l$*, if* $\Delta p_{lk_1} < \Delta p_{lk_2}$*, then* $c_{lk_1} < c_{lk_2}$ *and vice versa.*

Generally, the more a component is hardened, the greater is the cost of hardening, but so is the reduction in repair times. The reasoning behind Assumption 5.1 is similar to that of Proposition 1 in Sinha & Zoltners (1979). If there exists two hardening strategies $(k_1, k_2) \in K_l$ which violate the assumption, i.e., $\Delta p_{lk_1} > \Delta p_{lk_2}$ is true while $c_{lk_1} < c_{lk_2}$, strategy k_2 cannot be part of the optimal hardening solution.

5.3.3 A stochastic programming model

Definition 5.1. *Given a repair time vector* $\vec{p}$*, the min-harm (or equivalently, max-resilience) function, denoted by* $f^m(\cdot)$*, is the mapping* $\vec{p} \overset{f^m(\cdot)}{\longmapsto} H^{m,*}$*, where* $H^{m,*}$ *is the harm when repairs are scheduled optimally with* m *repair crews.*

Let C denote the capital budget available for hardening, $\mathcal{P}$ denote the repair time after hardening (modeled as a random vector to account for different disaster scenarios), and y_{lk} be a binary variable which is equal to 1 if hardening strategy k is chosen for edge l and

0 otherwise. A stochastic optimization model for minimizing the expected aggregate harm assuming m repair crews is shown below:

$$\min_{\{\Delta p_l\},\{y_{lk}\}} \quad \mathbb{E}[f^m(\vec{\mathcal{P}})] \tag{5.4a}$$

$$\text{s.t.} \quad \vec{p} - \Delta\vec{p} = \mathbb{E}[\vec{\mathcal{P}}] \tag{5.4b}$$

$$\sum_{k \in K_l} y_{lk} \leq 1, \ \forall l \in L^D \tag{5.4c}$$

$$\sum_{l \in L^D} \sum_{k \in K_l} c_{lk} y_{lk} \leq C \tag{5.4d}$$

$$\Delta p_l = \sum_{k \in K_l} \Delta p_{lk} y_{lk}, \ \forall l \in L^D \tag{5.4e}$$

$$y_{lk} \in \{0, 1\}, \ \forall l \in L^D, \ \forall k \in K_l \tag{5.4f}$$

where the expectation in eqn (5.4a) is over all disaster scenarios (the assumption being, different disaster scenarios cause different types/scales of damage and therefore lead to different repair times). While the notation L^D denotes the set of actual damaged edges in the context of the post-disaster scheduling problem (operational phase), we interpret it as the set of all edges which could potentially be damaged in the event of a disaster, the worst case operational scenario, in the context of the hardening problem. Eqn. (5.4b) is the mean-enforcing constraint (which requires that we have knowledge of the first moment of $\vec{\mathcal{P}}$), eqns. (5.4c) and (5.4f) force at most one hardening strategy to be chosen per edge from the set K_l, eqn. (5.4d) enforces the budget constraint, and eqn. (5.4e) models the (possible) reduction in repair time of each edge l due to hardening. Observe that the set of constraints (5.4c), (5.4d) and (5.4f) mimics a 0-1 knapsack constraint since we are essentially choosing a subset of hardening strategies from the set of all hardening strategies over all edges, subject to a budget constraint.

5.4 *Deterministic robust reformulation by Jensen's inequality*

Unfortunately, the aforementioned stochastic program is extremely difficult to solve, even with perfect knowledge of the statistical distribution of $\vec{\mathcal{P}}$. This is due to two reasons.

First, it is almost impossible to know beforehand the explicit form of $f^m(\cdot)$, even when the operational problem is solvable in polynomial time for $m = 1$. Second, evaluation of the objective function requires knowledge of the distribution function of $\vec{\mathcal{P}}$, while at the same time, this distribution function depends upon the decision variable (see eqns. 5.4a and 5.4b). This circular dependency effectively rules out the applicability of SAA methods. While metaheuristics such as simulated annealing could be used to solve the above problem to (near) optimality, doing so might require an inordinate amount of computation time. We therefore propose a deterministic robust reformulation in Section 5.4 which is more computationally tractable. We begin this section by showing that the min-harm function $f^m(\vec{p})$ is concave.

Theorem 5.1. *The min-harm function $f^m(\vec{p})$ is concave.*

Proof. Let $\vec{p_i}$ and $\vec{p_j}$ be two different repair time vectors and $f^m_{\vec{p_i}}(\vec{p_j})$ denote the harm evaluated by the optimal schedule corresponding to $\vec{p_i}$ when the actual repair time vector is $\vec{p_j}$. Obviously, $f^m(\vec{p_j}) = f^m_{\vec{p_j}}(\vec{p_j})$. For some $\vec{p_0} \neq \vec{p}$, we have:

$$f^m_{\vec{p_0}}(\vec{p}) - f^m_{\vec{p_0}}(\vec{p} - \Delta\vec{p}) = \sum_{l \in L^D} \Delta p_l \sum_{j \in R_l} w_j, \tag{5.5}$$

where R_l denotes the set of jobs assigned to the same crew as l, scheduled no earlier than l in the optimal schedule corresponding to $\vec{p_0}$. This shows that $f^m_{\vec{p_0}}(\vec{p})$ is a linear function of the p_l's in $\vec{p_0}$. And since $f^m(\vec{p})$ is the optimal schedule,

$$f^m(\vec{p}) = \min_{\vec{p_0}} \; f^m_{\vec{p_0}}(\vec{p}), \tag{5.6}$$

which implies that $f^m(\vec{p})$ is the point-wise minimum of a set of affine functions and is therefore concave. □

Since $f^m(\vec{p})$ is concave, Jensen's inequality Jensen (1906) holds and the objective function (5.4a) can be naturally upper bounded as follows:

$$\mathbb{E}[f^m(\vec{\mathcal{P}})] \leq f^m(E[\vec{\mathcal{P}}]) \tag{5.7}$$

The preceding discussion motivates the following deterministic robust reformulation (note that constraint (5.4b) has been wrapped into the objective function):

$$\begin{aligned} \min_{\{\Delta p_l\},\{y_{lk}\}} \quad & f^m(\vec{p} - \Delta\vec{p}) \\ \text{s.t.} \quad & (5.4c) \sim (5.4f) \end{aligned} \tag{5.8}$$

As will be apparent from Section 5.6, the above model is a key development which allows for an integrated treatment of the restoration process and the hardening problem.

We conclude this section with a note on the worst case impact on the objective function caused by the upper bounding by Jensen's inequality. Assume that the support of $\vec{\mathcal{P}}$ is bounded, i.e., $\vec{\mathcal{P}} \in [\vec{0}, \vec{p}_{max}]$. Then, it follows from Theorem 1 in Simic (2008) that:

$$f^m\left(\mathbb{E}[\vec{\mathcal{P}}]\right) - \mathbb{E}\left[f^m(\vec{\mathcal{P}})\right] \leq f^m\left(\vec{p}_{max}\right) - 2f^m\left(\frac{\vec{p}_{max}}{2}\right) \tag{5.9}$$

5.5 *Single crew approximation*

While the stochastic model and its deterministic reformulation discussed above are applicable for any value of m, for the rest of the paper, we make the assumption that $m = 1$. That is, the hardening decisions, which are made at the planning stage, are based on an assumption of single crew repair sequencing at the operational stage. The main motivation for making the single crew assumption is that it is practically impossible to know at the planning stage the actual number of repair crews that will be available in the event of a disaster. While hardening decisions based on an assumption of m_1 repair crews are most likely not the optimal decisions if the number of crews is actually m_2, a single crew assumption allows us to factor in the restoration process in these hardening decisions, without requiring a precise *a priori* knowledge of m or a joint probability distribution on the type/magnitude/scale of the disaster event and m.

We now provide a theoretical upper bound on the aggregate harm during the operational stage, applicable for any arbitrary value of m, even when hardening decisions have been made based on $m = 1$. Consider two hardening strategies, A and B, with corresponding expected reduction in repair time vectors, $\Delta\vec{p}_A$ and $\Delta\vec{p}_B$. Suppose strategy A is obtained from the minimization of the objective function (5.8) with $m = 1$ and B is an arbitrary hardening strategy. For a strategy S, we also define $H_S^{m,*} := f^m(\vec{p} - \Delta\vec{p}_S)$, the deterministic optimal harm (i.e., the objective function function in eqn. 5.8) at the operational stage with m repair crews and H_S^m denote the harm for an m-crew schedule computed using Algorithm 3 with repair time vector $\vec{p} - \Delta\vec{p}_S$. Then:

$$
\begin{aligned}
H_B^{m,*} &\geq \frac{1}{m} H_B^{1,*} && (5.10)\\
&\geq \frac{1}{m} H_A^{1,*} && (5.11)\\
&\geq H_A^m - \left(\frac{m-1}{m}\right) H_A^\infty && (5.12)\\
&\geq H_A^m - \left(\frac{m-1}{m}\right) H^\infty && (5.13)
\end{aligned}
$$

where the first inequality follows from Proposition 2.4, the second inequality follows from the fact that hardening strategy A is by definition optimal when $m = 1$, the third inequality follows from eqn. 2.49 in the proof of Theorem 2.4, and the fourth inequality follows from the fact that the aggregate harm defined by any m after hardening is upper bounded by the aggregate harm before hardening. Rearranging terms, we have:

$$
H_A^m \leq H_B^{m,*} + \left(\frac{m-1}{m}\right) H^\infty \qquad (5.14)
$$

Since B represents any hardening strategy,

$$
\begin{aligned}
H_A^m &\leq \min_B \left\{ H_B^{m,*} + \left(\frac{m-1}{m}\right) H^\infty \right\}\\
&= H_{OPT}^{m,*} + \left(\frac{m-1}{m}\right) H^\infty, && (5.15)
\end{aligned}
$$

where $H_{OPT}^{m,*}$ represents the deterministic optimal harm when a network has been hardened by minimizing objective function (5.8), with perfect knowledge of m.

The implication of eqn. (5.15) is that, while hardening strategy A may not be optimal for some chosen $m > 1$, the approximation gap between the harm when an m-crew schedule (obtained by applying Algorithm 3) is used during the operational stage and the harm corresponding to an optimal hardening strategy for that specific value of m is at most $\left(\frac{m-1}{m}\right) H^{\infty}$. In practice, the value of H^{∞} can be determined straightforwardly during the planning stage. Clearly, the smaller H^{∞} is, the better the single crew approximation is and an exact or probabilistic *a priori* knowledge of m corresponding to different disaster events may not even be necessary if H^{∞} is small enough. Note that the H^{∞} term on the r.h.s of eqn. (5.15) could be way smaller than $H_{OPT}^{m,*}$. This is likely to be so when the hardening budget is limited since the benefits of an infinite number of repair crews will outweigh the benefits of hardening.

We wish to emphasize that our single crew approximation during the planning stage does not prevent the network operator from deploying multiple crews during the operational stage for post-disaster restoration. In fact, a network which has been designed/hardened with an eye on the restoration process, albeit with one repair crew, will ensure a smaller aggregate harm (or improved resilience) during the restoration process post-disaster when additional repair crews might be available, as opposed to a network which has been designed/hardened with no consideration given to the restoration process. Simulation results discussed in Section 5.7.3 confirm this.

5.6 Restoration Process Aware Hardening Problem

Usually, the restoration problem and the hardening problem are treated separately because the former is an operational problem while the latter is a planning problem. However, we argue that the two problems should not be treated in isolation because hardening can affect the repair times, which in turn, can influence the restoration times through the sequencing process and thereby the aggregate harm or resilience. The model that we formulate is similar to single machine scheduling with controllable processing times, which dates back to the 1980s (Nowicki & Zdrzałka 1990). See Section 2 in (Shioura et al. 2016) for a review of recent advances. Our problem is more complicated in the sense that the effect of hardening decisions (analogous to 'costs of compression amount' in the context of single machine scheduling with controllable processing times) are not just linear, instead they are embedded in the sequencing problem.

In this section, we discuss two solution approaches for the so-called 'restoration process aware hardening problem' (RPAHP), first an MILP formulation, followed by a heuristic algorithm framework inspired by a continuous convex relaxation, considering $m = 1$.

5.6.1 MILP formulation

In Section 5.2, we developed an MILP model for optimizing the repair schedule with one repair crew, while in Section 5.5, we developed a deterministic single crew approximation of the hardening problem , both with the same objective, minimization of the aggregate harm. These two models can be easily incorporated into an integrated MILP formulation, as shown below:

$$\min_{\vec{x},\vec{u},\Delta\vec{p},\vec{y}} \quad \sum_{t=1}^{T} h^t \tag{5.16}$$

$$\text{s.t} \quad (5.3b) \sim (5.3f)$$

$$d^t \geq d^{t-1} + (p_l - \Delta p_l) \times x_l^t, \; \forall t \in [1, T], \; \forall l \in L^D \tag{5.17}$$

$$(5.4c) \sim (5.4f)$$

Observe that the impact of hardening, Δp_l, is incorporated into constraint (5.17). The product of Δp_l and x_l^t on the r.h.s of eqn. (5.17) can be easily linearized using the big-M method, details of which are omitted.

5.6.2 A continuous convex relaxation

For notational brevity, we define $f(\cdot) := f^1(\cdot)$. As stated previously, the min-harm function $f^m(\cdot)$ is concave piecewise affine and so is $f(\cdot)$. In general, concave minimization problems are $\mathcal{NP}$-hard Garey et al. (1976). In our case, there are at most $n!$ affine pieces, corresponding to $n!$ number of affine possible sequences, where $n = |L^D|$ is the number of damaged edges.

The RPAHP involves two types of decision variables, the sequencing variables (the x's and u's) and the hardening variables (the Δp's). Given the hardening variables, it is straightforward to see that the joint optimization problem reduces to the single crew sequencing problem, which can be solved optimally in polynomial time as stated previously in Section 2.6.2.

Now let us consider the case where the sequencing variables are fixed. Let

$$\Omega_l := \sum_{j \in R_l} w_j \tag{5.18}$$

where R_l is the set of some edges $l \in L^D$ and all its successors in the given sequence. The quantity Ω_l represents the reduction in aggregate harm per unit decrease in p_l. The objective function for the hardening problem can now be recast as $f(\vec{p}) = \sum_{l \in L^D} \Omega_l \, p_l$, which implies:

$$f(\vec{p} - \Delta\vec{p}) = \sum_{l \in L^D} \Omega_l p_l - \sum_{l \in L^D} \Omega_l \Delta p_l \tag{5.19}$$

Since the first term on the r.h.s of eqn. (5.19) is a constant, instead of minimizing $f(\vec{p} - \Delta\vec{p})$, an equivalent formulation is:

$$\max_{\vec{y}} \quad \sum_{l \in L^D} \sum_{k \in K_l} \Omega_l \Delta p_{lk} y_{lk} \tag{5.20a}$$

$$\text{s.t.} \quad \sum_{k \in K_l} y_{lk} \leq 1 \ \ , \forall l \in L^D \tag{5.20b}$$

$$\sum_{l \in L^D} \sum_{k \in K_l} c_{lk} y_{lk} \leq C \tag{5.20c}$$

$$y_{lk} \in \{0, 1\}, \ \ \forall l \in L^D, \ \ \forall k \in K_l \tag{5.20d}$$

This model is similar to that of the multiple choice knapsack problem (Sinha & Zoltners 1979), where $\Omega_l \Delta p_{lk}$'s are the value coefficients and c_{lk}'s are the cost coefficients. Since the multiple choice knapsack is known to be $\mathcal{NP}$-hard, we propose an algorithm based on convex envelopes and LP relaxation, similar to (Kameshwaran & Narahari 2009).

Definition 5.2 (Convex envelope (Horst & Tuy 2013)). *Let $M \subset \mathcal{R}^n$ be convex and compact and let $g : M \to \mathcal{R}$ be lower continuous on M. A function $\hat{g} : M \to \mathcal{R}$ is called the convex envelope of f on M if it satisfies:*

- *$\hat{g}(x)$ is convex on M,*
- *$\hat{g}(x) \leq g(x)$ for all $x \in M$,*
- *there is no function $h : M \to \mathcal{R}$ satisfying (1), (2) and $g(x_0) < h(x_0)$ for some point $x_0 \in M$.*

Intuitively, the convex envelope is the best underestimating convex function of the original function. Details of a polynomial time algorithm for computing the convex envelope of a piecewise linear function can be found in (Kameshwaran & Narahari 2009).

Given a discrete function of c_{lk} vs. Δp_{lk} for some edge l and a set of all hardening actions $k \in K_l$, we first connect the neighboring points, starting from the origin, to construct a continuous piecewise linear cost function $C_l(\Delta p_l)$, where Δp_l is the relaxed continuous decision variable. It follows from Assumption 5.1 that C_l is a strictly increasing function. Let $\hat{C}_l$ denote the convex envelope of C_l and $\hat{K}_l = \{1, 2, \cdots, |\hat{K}_l|\}$ denote the set of breakpoints/knots on the convex envelope (excluding the origin) corresponding to the hardening strategies in consideration, indexed in ascending order of Δp_{lk}. The linear relaxation of (5.20) based on the convex envelope approximations, which we denote as **(LP)**, can then be formulated as:

$$\max_{\Delta \vec{p}} \quad \sum_{l \in L^D} \Omega_l \Delta p_l \tag{5.21a}$$

$$\text{s.t.} \quad Q_l \geq \max_{k \in \hat{K}_l} \left[\mu_{lk} \left(\Delta p_l - \alpha_{lk} \right) + b_{lk} \right], \forall l \in L^D \tag{5.21b}$$

$$\sum_{l \in L^D} Q_l \leq C \tag{5.21c}$$

$$0 \leq \Delta p_l \leq \Delta p_{l,|\hat{K}_l|}, \forall l \in L^D \tag{5.21d}$$

where μ_{lk} and b_{lk} are the slope and intercept of the k^{th} piece of $\hat{C}_l$, α_{lk} is the lower breakpoint of the k^{th} piece of $\hat{C}_l$, and Q_l is an intermediate decision variable which accounts for the budget spent on edge l. This formulation is similar to the conventional continuous knapsack problem, and it turns out that the optimal values of Δp_l are always from the set

$\{0, \text{some } \beta_{lk}, (\alpha_{lk}, \beta_{lk})\}$, where β_{lk} is the upper breakpoint of the k^{th} piece of $\hat{C}_l$. Furthermore, at most one Δp_l can have an intermediate value in the range $(\alpha_{lk}, \beta_{lk})$ in the optimal solution. For some l and $k > 1$, the intercept parameter, b_{lk}, is: $b_{lk} = \hat{C}_l(\alpha_{lk}) = \hat{C}_l\left(\beta_{l(k-1)}\right)$. For $k = 1$, $\alpha_{lk} = \hat{C}_l(\alpha_{lk}) = b_{lk} = 0$.

The preceding LP relaxation (5.21) can also be solved optimally using a greedy algorithm by first sorting the ratios $\left\{\frac{\Omega_l}{\mu_{lk}}\right\}$ in a descending order, and then choosing the components (and the degree of hardening) based on that sorted list iteratively, until the budget is exhausted. Ties, if any, during the selection process, are broken arbitrarily. We use a Δp variable for each edge l and each segment k of $\hat{C}_l$. All these Δp_{lk} variables are initialized to 0. Once a selection is made from the sorted list at any iteration T, say (l_T, k_T), we set $\Delta p_{l_T k_T}$ equal to the maximum value possible within the range $[\alpha_{l_T k_T}, \beta_{l_T k_T}]$ such that the 'cumulative budget' at the end of iteration T does not exceed C. Typically, this maximum value will be at the upper breakpoint $\beta_{l_T k_T}$, unless, doing so results in a budget violation. In that case, a proper value within the range $(\alpha_{l_T k_T}, \beta_{l_T k_T})$ is chosen such that the budget is met exactly. At the end of every iteration, we evaluate the expression of budget spent:

$$\Lambda = \sum_{l \in L^D} \hat{C}_l\left(\max_{k \in \hat{K}_l}\{\Delta p_{lk}\}\right), \tag{5.22}$$

which represents the cumulative budget consumed till the current iteration. The algorithm terminates when $\Lambda = C$. Upon termination, the optimal Δp_l values can be obtained from the Δp_{lk} values as follows:

$$\Delta p_l = \max_{k \in \hat{K}_l}\{\Delta p_{lk}\}. \tag{5.23}$$

We now provide an example which helps illustrate the operation of the algorithm.

Consider a scenario where two edges are to be repaired, $l = 1, 2$, and the hardening cost functions for the two edges are as shown in Fig. 5.1. Suppose $C = 10$ and $\Omega_1 = \Omega_2 = 1$.

The selections made by the greedy algorithm at each step are as follows:

- *Step 1*: Breaking ties arbitrarily, choose $(l = 2, k = 1)$, set $\Delta p_{21} = 1$, cumulative hardening cost $= C_1(1) = 1$.

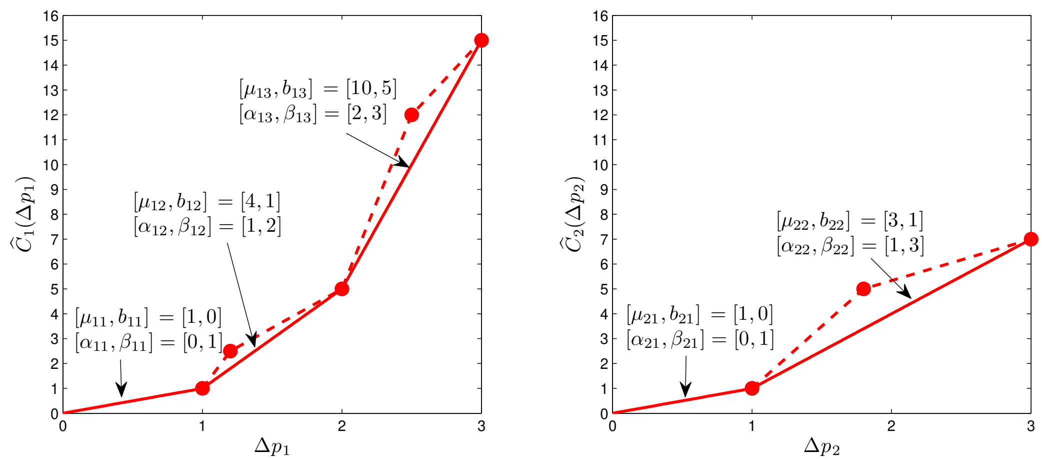


Figure 5.1: Hardening cost functions for illustrating the greedy algorithm used to solve the continupus knapsack-like problem. The solid circles represent the actual discrete hardening strategies and costs, the dashed lines represent the piecewise linear constructions, while the solid lines represent the convex envelope approximations.

- *Step 2*: Choose $(l = 1, k = 1)$, set $\Delta p_{11} = 1$, cumulative hardening cost $= C_1(1) + C_2(1) = 1 + 1 = 2$.

- *Step 3*: Choose $(l = 2, k = 2)$, set $\Delta p_{22} = 3$, cumulative hardening cost $= C_1(1) + C_2(\max[1, 3]) = 1 + 7 = 8$.

- *Step 4*: Choose $(l = 1, k = 2)$, set $\Delta p_{12} = 1.5$, cumulative hardening cost $= C_1(\max[1, 1.5]) + C_2(\max[1, 3]) = 3 + 7 = 10$. Note that, unlike the previous 3 steps, we can only afford 1.5 units of hardening corresponding to $(l = 1, k = 2)$ so that the budget is not violated.

The LP solutions are therefore the points $(1.5, 3)$ and $(3, 7)$ for edges 1 and 2 respectively.

5.6.3 An iterative heuristic algorithm

We now discuss an iterative heuristic algorithm for solving the RPAHP. First, we note that the solutions obtained from the greedy algorithm used to solve the convex relaxation formulation (5.21) may need to be *rounded down* to the nearest lower breakpoints on the convex envelopes so that the hardening strategy is feasible for each edge. In the example, this represents an actually available 'under-budget' hardening strategy closest to the point $(1.5, 3.0)$ chosen by the greedy algorithm. Observe that the point $(1.2, 2.5)$ represents a feasible hardening strategy, even though it is not on the convex envelope. However, no rounding is necessary for line 2 since the point selected by the greedy algorithm, $(3.0, 7.0)$, does correspond to an actual hardening strategy. By rounding down, wherever necessary, we ensure that the budget constraint will not be violated due to the rounding process. However, after completion of the rounding process, we may find that a portion of the budget has been left unspent. We therefore incorporate a *backfill* heuristic which iteratively solves LP relaxations of the form (5.21) with the unspent budget from the previous iteration and the remaining available hardening options, along with updated convex envelopes, followed by a rounding down to a feasible hardening strategy. The backfill process terminates whenever the budget has been spent exactly, or, when no further enhancement is possible on any edge without exceeding the budget. Fig. 5.2 provides a flowchart of the iterative heuristic algorithm for solving the restoration process aware hardening problem. In the context of the example provided in the previous sub-section, rounding down the LP solution for edge 1 to the point $(1, 1)$ creates an unspent budget of 2 units, which becomes the new budget for the second iteration. During the second iteration, the points $(0, 0)$ (no hardening is a feasible option in iteration 1), $(2, 5)$, $(2.5, 12)$ and $(3, 15)$ in the left panel of Fig. 5.1 are no longer in consideration and the convex envelope is recomputed over the set of points $(1, 1)$ and $(1.2, 2.5)$, with the former being the new origin. Since the optimal repair schedule depends on the repair times, we update the schedule after every iteration t with the new repair time vector, $\vec{p}(t + 1) \leftarrow \vec{p}(t) - \Delta\vec{p}(t)$. The backfill process terminates whenever the budget has been spent exactly, or, when no further enhancement is possible on any edge without exceeding the budget.

Summarizing what we have so far, we now describe a general framework of a multi-run heuristic algorithm for solving the RPAHP, as shown in Fig. 5.2. Broadly speaking, the approach involves three major stages. In the first stage, we compute the single crew optimal sequence, given $\vec{p}$, the expected repair time vector before hardening. In the second stage, we use the optimal repair sequence obtained from the first stage and solve the LP relaxation (5.21) using the convex envelopes of the hardening cost functions, followed by rounding, which yields a set of feasible hardening decisions. In the third stage, we implement a backfill procedure by re-solving the LP relaxation (5.21) with updated information, as described in the previous paragraph. We provide three options in Fig. 5.2 which differ in how often the repair sequence is updated based on some hardening decisions. Option 3, which is the most aggressive, updates the repair sequence *after every iteration of the greedy algorithm used for solving the LP relaxation* (5.21). To avoid clutter, we have opted to show the feedback arrow in Option 3 going directly to the 'blue greedy algorithm box', instead of expanding the details of it. Option 1, which is the most conservative, does not update the repair sequence at all and uses the initial Ω_l's until termination. Option 2 represents a middle ground and updates the repair sequence *after completion of the greedy algorithm used for solving the LP relaxation* (5.21). Implementation details of these three options are shown in Algorithm 8.

We close this section by pointing out that each iteration of the above algorithm can also be interpreted from the perspective of the convex-concave procedure discussed in Lipp & Boyd (2016). Essentially, the min-harm function $f(\cdot)$ is being convexified at some $\vec{p}$ and the resulting linearized problem solved to give the near-optimal hardening results.

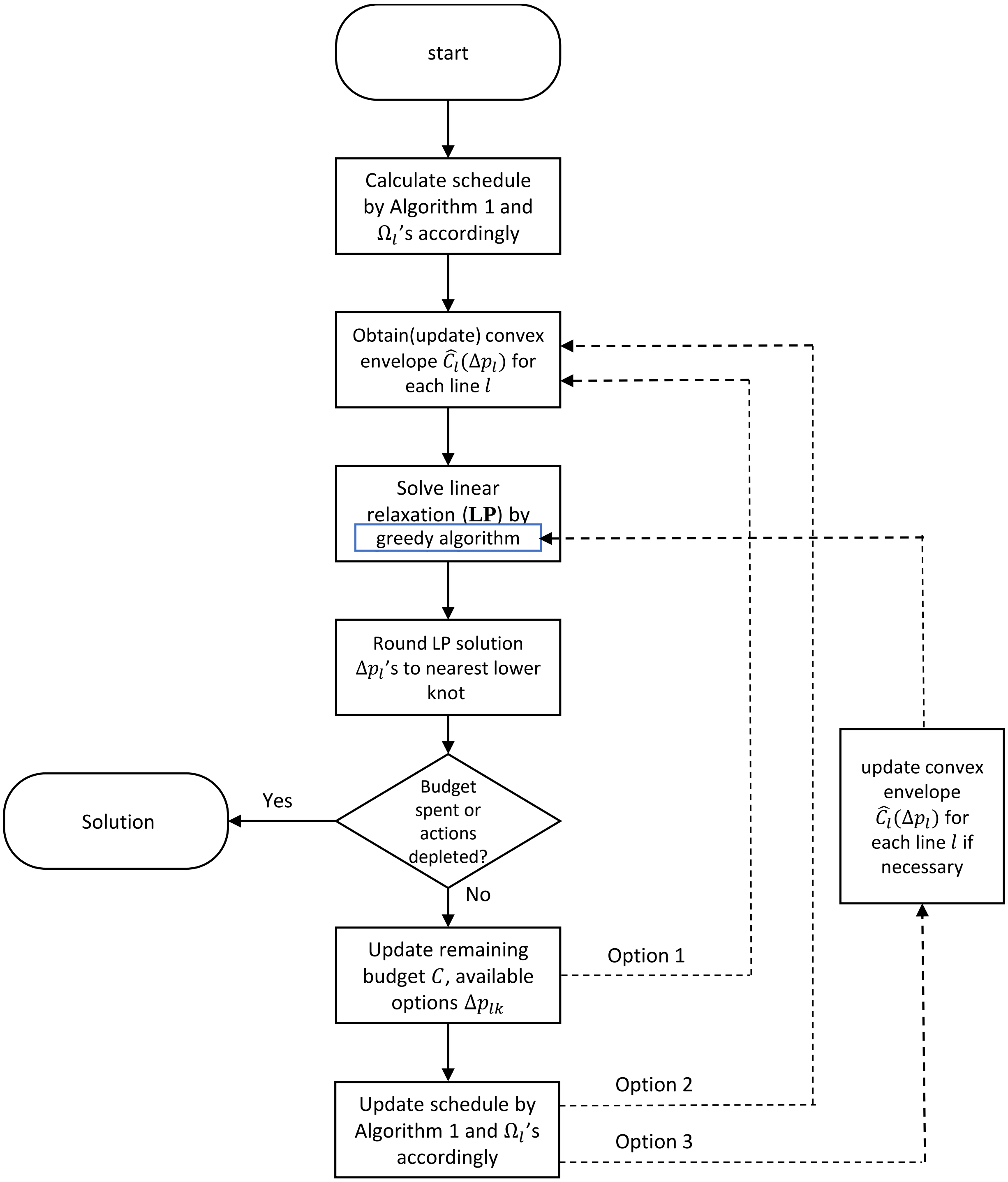


Figure 5.2: Flowchart of an iterative heuristic algorithm for solving the RPAHP.

Algorithm 8 Algorithms for restoration process aware distribution systems hardening.

Compute the optimal sequence given the expected repair time $\vec{p}$, using Algorithm 2;

Calculate the weights Ω_l according to eqn. (5.18);

Obtain the convex envelopes of costs $\hat{C}_l(\Delta p_l)$ with $\hat{K}_l$ pieces, along with the coefficients $\mu_{lk}, b_{lk}, \alpha_{lk}$ and β_{lk}, for each edge $l \in L^D$;

$H \leftarrow \emptyset$;

$k_l \leftarrow 1, \ \forall l \in L^D$;

while true **do**

  find $l \in L^D \setminus H$ with largest $\frac{\Omega_l}{\mu_{l,k_l}}$;

  let $\Delta p_l = \beta_{lk_l}$ and calculate the current cost $\Lambda = \sum_{l \in L^D} \hat{C}_l(\Delta p_l)$;

  if $\Lambda = C$ **then**

    break;

  else if $\Lambda > C$ **then**

    $\Delta p_l = \beta_{l,k_l-1}$;

    Option 2 & 3: Update the optimal sequence given the current expected repair time $\vec{p} - \Delta\vec{p}$ and then update Ω's.

    Update the convex envelope of cost $\hat{C}_l(\Delta p_l)$ for edge l and then update the coefficients $\mu_{lk}, b_{lk}, \alpha_{lk}$ and β_{lk}, for each edge $l \in L^D$;

  else if $k_l = |\hat{K}_l|$ **then**

    $\Delta p_l = \beta_{l,k_l-1}$;

    $H \leftarrow \{H, l\}$;

  else

    $k_l = k_l + 1$;

    Option 3: Update the optimal sequence given the current expected repair time $\vec{p} - \Delta\vec{p}$ and then update Ω's.

    continue;

  end if

  if $|H| = |L^D|$ **then**

    break;

  end if

end while

5.7 Case studies

5.7.1 IEEE 13 node test feeder

We first tested the MILP and the heuristic approach discussed in the previous section on the IEEE 13 node test feeder with randomly generated C_l's and two different budgets. Values of $\mathbb{E}[f(\cdot)]$ in this section were computed using Monte Carlo simulations assuming an independent geometric distribution for each p_l. With a budget of $C = 20$, hardening actions did not result in different repair schedules and both the MILP and heuristic approaches yielded identical results. With a budget of $C = 60$, even though the hardening actions suggested by the MILP and heuristic approaches differ for two edges, as shown in Table 5.1, the objective values obtained from the greedy algorithm, both for $\mathbb{E}[f(\cdot)]$ and its upper bound $f(\mathbb{E}[\cdot])$, are very close to those provided by the MILP formulation. In fact, the ratios of the $f(\mathbb{E}[\cdot])$ measure from the greedy algorithm to the $\mathbb{E}[f(\cdot)]$ measure from the MILP algorithm are both approximately 1.04 for $C = 20$ and $C = 60$ (note that this ratio captures the worst case performance loss, including the effect of upper bounding the true objective function using Jensen's inequality).

Table 5.1: Comparison of hardening results on the IEEE 13 node test feeder with a budget of $C = 60$.

Edge l	Δp_l by MILP	Δp_l by heuristic
671-684	0.2	0.4
645-646	0	0.4
632-645	0.5	0.8
632-671	5.3	3.5
$f(\mathbb{E}[\cdot])$	14.411	$\boxed{14.520}$
$\mathbb{E}[f(\cdot)]$	$\boxed{13.987}$	14.146

We then varied the hardening budget from 0 to 50. Fig. 5.3 shows the aggregate harm as a function of the budget for $m = 1$. The MILP and the heuristic produced almost identical results when using the $f(\mathbb{E}[\cdot])$ measure so that their plots almost overlap. The plots corresponding to the $\mathbb{E}[f(\cdot)]$ measure are also very close, considering the errors introduced by Monte Carlo simulations. The gap between the true objective and the deterministic objective, $\mathbb{E}[f(\cdot)] - f(\mathbb{E}[\cdot])$, is fairly constant for both the MILP and the heuristic. As expected, the aggregate harm decreases (resilience increases) as the hardening budget increases.

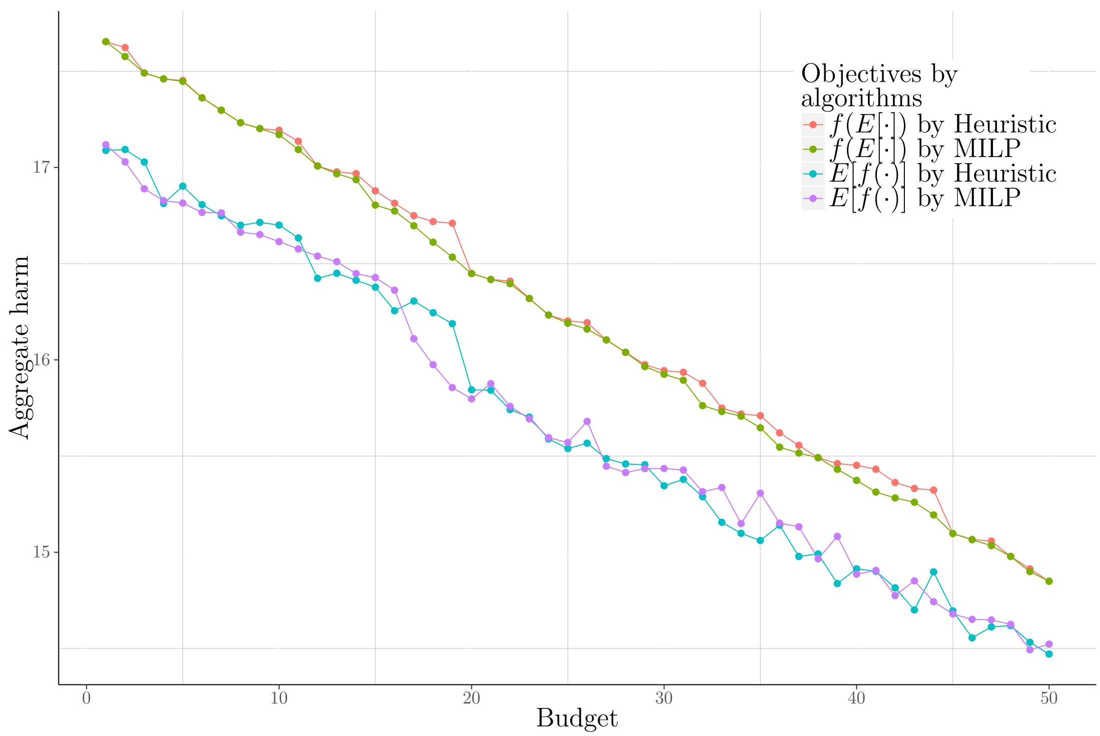
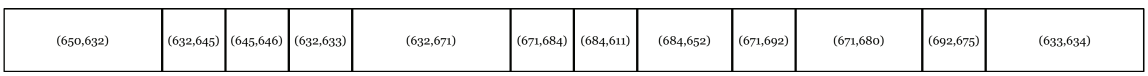
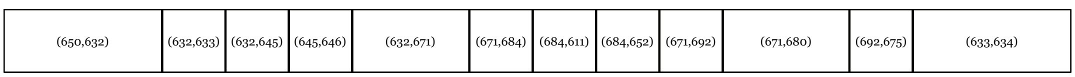


Figure 5.3: Aggregate harm vs. hardening budget for the IEEE 13 node test feeder.

Finally, in Fig. 5.4, we compare the pre-hardening sequencing decisions with the post-hardening decisions calculated using the MILP for the same case study as in Table 5.1. In these Gantt charts, each 'box' represents the repair of a line and the width of each 'box' is proportional to the repair time, appropriately scaled for better visualization. Observe that the two sequences differ by the relative locations of lines (632, 645), (645, 646) and (632, 633). This confirms the interaction between sequencing and hardening decisions. Finally, we want to emphasize that these sequencing decisions are abstract constructs that only serve to model operational decisions at the planning stage.



(a) Gantt chart of sequencing decisions before hardening



(b) Gantt chart of sequencing decisions after hardening (MILP)

Figure 5.4: Comparison of optimal single crew repair schedules before and after hardening.

5.7.2 IEEE 37 node test feeder

Next, we ran our algorithms on one instance of the IEEE 37 node test feeder (Kersting 2001). Since the running time of the MILP formulation increases exponentially with network size, we allocated a time budget of 10 hours. In contrast, the heuristic algorithm yielded a solution within seconds. Table 5.2 shows the edges for which the MILP and heuristic approach produced different hardening results, along with the objective function values.

Analogous to Fig. 5.3, Fig. 5.5 shows a plot of the aggregate harm vs. the hardening budget for $m = 1$. Due to the inordinate amount of time required to solve the MILP, we show results only for the heuristic. Unlike the 13 node feeder, the aggregate harm in this case exhibits a steep drop initially before gradually tapering off. This tapering off reflects the fact that our cost functions were so chosen that no line could be hardened enough to reduce its repair time to zero; i.e., for every line l, we ensured that $p_l - \Delta p_l > 0$.

Table 5.2: Comparison of reduction in repair times, Δp_l's, due to hardening on the IEEE 37 node test feeder with a budget of $C = 200$.

edge l	MILP(10 hours)	Option 1	Option 2	Option 3
(744, 729))	0.8	0	0	0
(702, 703)	0.2	0.2	0.2	0.3
(708, 733)	0	0.3	0.3	0.3
(702, 705)	2.1	0.4	0.4	2.1
(734, 737)	0	0	0.2	0.2
(708, 732)	0	0.2	0.2	0
(734, 710)	0.1	1.4	1.7	0.1
$f(\mathbb{E}[\cdot])$	843.08	842.04	843.84	837.93
$\mathbb{E}[f(\cdot)]$	672.21	667.35	667.42	666.09

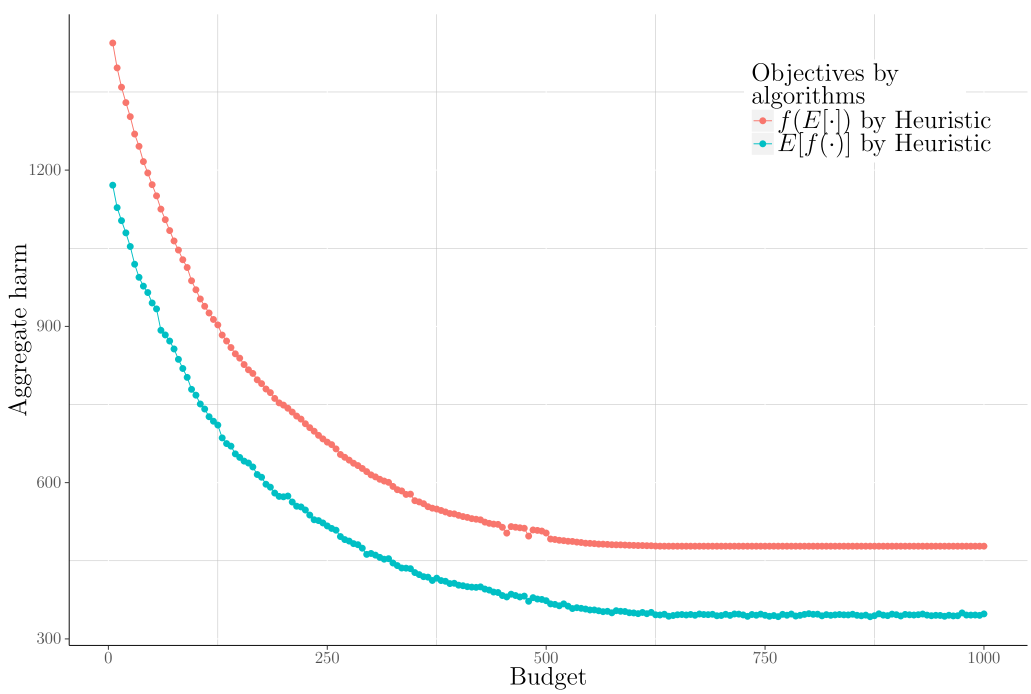


Figure 5.5: Aggregate harm vs. hardening budget for the IEEE 37 node test feeder.

In order to compare the performances of the three options within the heuristic framework, we then varied the hardening budget from 1 to 400. Fig. 5.6 summarizes these results. Intuitively, when the hardening budget is small, we expect the three options to behave similarly since reductions in repair times, if any, are likely to be small enough so as not to trigger a change in the repair schedule, rendering the 'update schedule' step in Fig. 5.2 moot. Similarly, when the hardening budget is large, all three options should behave similarly since most edges are likely to be hardened to the maximum degree possible at the end of the first run, and in this case, the 'update schedule' step would be inconsequential since the algorithm would tend to terminate after the first run. As can be observed from Fig. 5.6, the three options indeed behave similarly at either end of the budget spectrum, but produce somewhat different results for intermediate budgets (in the range $21 - 303$), although the differences are not appreciable. For a better understanding of the average performance of the three options, we conducted 200 trials with randomly generated hardening cost functions and a budget of $C = 282$. Option 1 turned out to be the best on 53 trials, option 2 on 69 trials, and option 3 on 158 trials. Note that the numbers do not add up to 200 since ties were counted while ranking the three options. All three options produced identical results on 19 trials. However, the largest difference that we observed between any two options was 4.7%. Consequently, we recommend Option 1 as the preferred option if ease of implementation and fastest computational performance are desired.

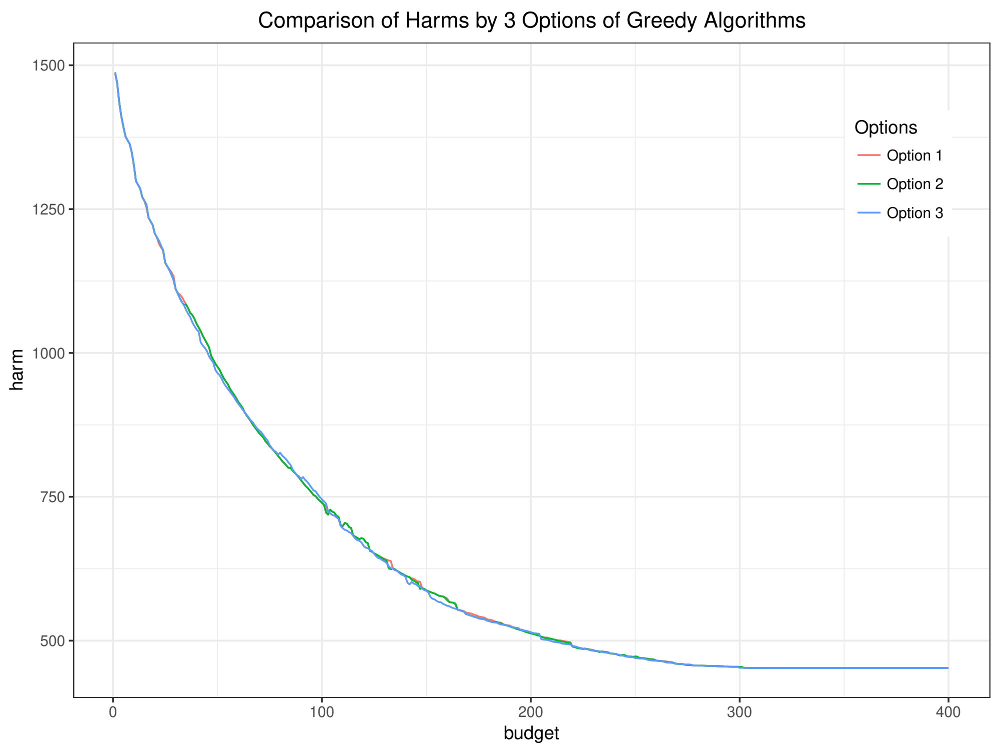


Figure 5.6: Comparison of harms by 3 options of the heuristic framework on the IEEE 37 node test feeder.

5.7.3 IEEE 8500 node test feeder

Finally, we tested the performance of the heuristic algorithm on one instance of the IEEE 8500 node test feeder medium voltage subsystem (Arritt & Dugan 2010) containing roughly 2500 edges. We did not attempt to solve the ILP model in this case, but the heuristic algorithm took just 9.36 seconds to solve this problem.

Analogous to Figs. 5.3 and 5.5, Fig. 5.7 shows a plot of the aggregate harm vs. the hardening budget for $m = 1$. Due to issues with computational time, we chose to plot the $f(\mathbb{E}[\cdot])$ measure only as a function of the budget using the heuristic algorithm.

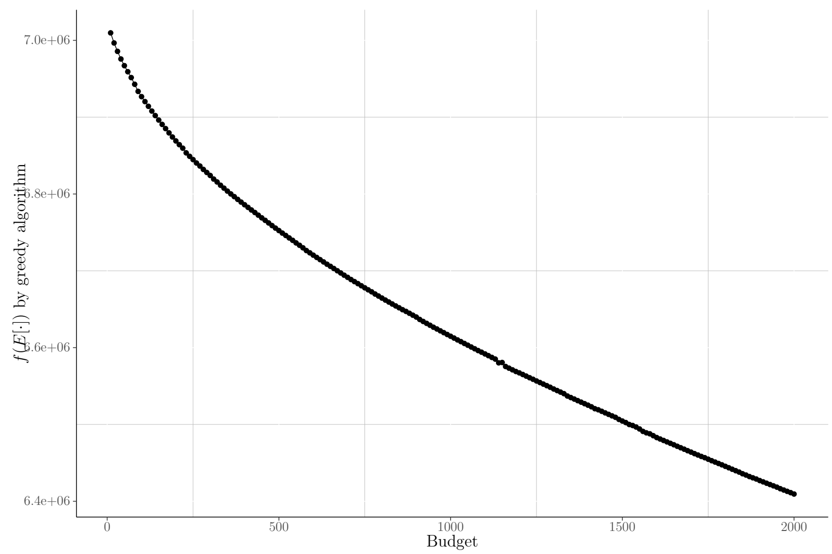


Figure 5.7: Aggregate harm vs. hardening budget for the IEEE 8500 node test feeder.

Even if hardening decisions at the planning stage are made based on single crew operational scheduling, the resilience of the system would still improve if multiple crews are deployed at the operational stage. To emphasize this aspect, Fig. 5.8 shows the normalized improvement in harm,

$$\beta := \frac{H^m - H_A^m}{H^m} \tag{5.24}$$

for different values of m. The reduction in the repair time vector due to hardening, $\Delta\vec{p}_A$, was obtained using the iterative heuristic algorithm with a budget of $C = 2000$. For $m > 1$, the aggregate harms before and after hardening, H^m and H_A^m, were determined from m-crew schedules obtained using Algorithm 3. The normalized improvement in harm shows a slight decrease (note the scale on the y-axis). This generally decreasing trend is understandable since the improvement in system resilience due to the availability of an increasing number of repair crews will gradually outweigh the improvement in resilience due to hardening with a limited budget. Nevertheless, even for $m = 50$, we can observe that the normalized improvement in harm due to hardening remains above 8%, even though the hardening decisions were made considering $m = 1$.

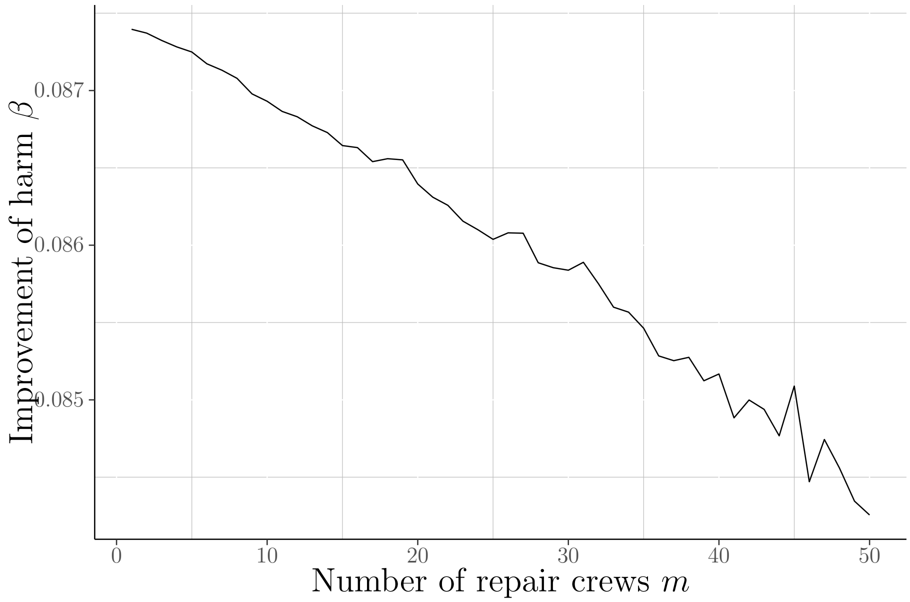


Figure 5.8: Normalized improvement in harm, β (see eqn. 5.24), as a function of the number of repair crews, m, for the IEEE 8500 node test feeder.

5.8 Conclusions

In this paper, we investigated the problem of strategically hardening a distribution network to be resilient against natural disasters. Motivated by research on resilient infrastructure systems in civil engineering, we proposed an equivalent definition of resilience with a clear physical interpretation. This allows us to integrate the post disaster restoration process and the planning stage component hardening decision process into one problem, which, we argued, is necessary since both aspects ultimately contribute to system resilience. This is a major departure from most current research where the two aspects of resilience are treated separately. We first modeled the restoration problem as an MILP and the hardening problem as a stochastic program, which was reformulated using Jensen's inequality and approximated by single crew for computational tractability. Finally, we unified the sequencing and hardening aspects and proposed an integrated MILP model as well as an iterative heuristic algorithm. The expected component repair times are used to generate an optimal single crew repair sequence, based on which hardening decisions are made sequentially in a greedy manner. Simulations on IEEE standard test feeders show that the heuristic approach provides near-optimal solutions efficiently even for large networks.

Chapter 6

CONCLUSIONS AND FUTURE WORK

As we mentioned in Chapter 1, our ultimate goal is to develop useful tools for industry. The work in this dissertation still needs to be refined further to fit in the need of practice. But hopefully this dissertation will provide some basic insights for future work. Therefore, in this chapter, we will focus on some potential research directions as we feel like there are not many to conclude and this is just a start.

6.1 Post-disaster repair scheduling in a reconfigurable distribution network

In our previous research, we assume the distribution network is radial, a valid assumption for many distribution networks especially those in rural areas. To completely bring the damaged system back online in this kind of systems, all damaged components should be repaired. However, there exist some distribution system in some metropolitan areas that is built as a meshed network while operated radial. This gives rise to a natural problem called distribution network reconfiguration (Sarfi et al. 1994), which aims to find the best system topology under some criteria. The ability of reconfiguration not only provides additional robustness of the system but also provides alternative ways of energizing the system. In light of this, we consider the problem of joint optimization of scheduling and reconfiguration. Current setting is, to find a reconfiguration and stick with it during the restoration period.
The complexity of this problem has not been analyzed in literatures but Nurre & Sharkey (2014) have provided the complexity results for a class of similar problems called INDS. In particular, they showed that $1 \mid MF \mid$ Cumulative and $P_2 \mid MF \mid$ Cumulative where all weights are equal are both strongly $\mathcal{NP}$-hard. Similarly, by reducing the well-known strongly $\mathcal{NP}$-hard problem Set Cover to an instance (A) of this joint optimization of scheduling and

reconfiguration, it is possible to show that the problem is indeed strongly $\mathcal{NP}$-hard and cannot be approximated to within a ratio of $\log n$ unless $\mathcal{NP} \subset \text{TIME}(n^{O(\log\log n)})$, where n is the number of nodes in this network. However, the instance (A) would be a general network built on a bi-partite graph, which is nowhere close to a real-life distribution networks. Therefore, such an analysis could be too conservative in practice.

We are considering two solving techniques: 1) performing some spanning tree selecting strategy and scheduling the repairs within the tree above; this approach resembles that by Nojima & Kameda (1992), Wang & Cui (2012); 2) find out a tree decomposition (Halin 1976, Robertson & Seymour 1984) of the loopy network and schedule the clusters of repairs, which are connected in a radial network.

We implemented 3 different spanning tree selecting strategies, Shortest Path Tree (SPT), Minimum Spanning Tree (MST) and (Node-) Weighted Shortest Path Tree (WSPT) and tested the optimality gap on 1000 instances of a modified IEEE 13 node test feeder with 3 additional lines as shown in Figure 6.1.

A comparison of bar plots of optimality gaps is shown in Figure 6.2, where MST seems to be the best strategy in the simulations. There is still no further theoretical analysis regarding this kind of strategies.

6.2 Joint optimization of scheduling and routing

When the repair team are sent to repair a sequence of components, they do have to travel in a transportation network. Co-optimizing the repairs and routes will be the most practical solution to the most fundamental problem of restoring the post-disaster system. Previous work on this topic relies on Mixed Integer Programming (MIP) and therefore cannot be applied to the case of severe damages.

Ideally, this work would require: 1) a test case of distribution system coupled with a transportation network; 2) a not-so-slow MILP formulation or a near-optimal relaxation as the benchmark of optimality; 3) an approximation algorithm that works as the problem scales up and that hopefully has a bounded performance.

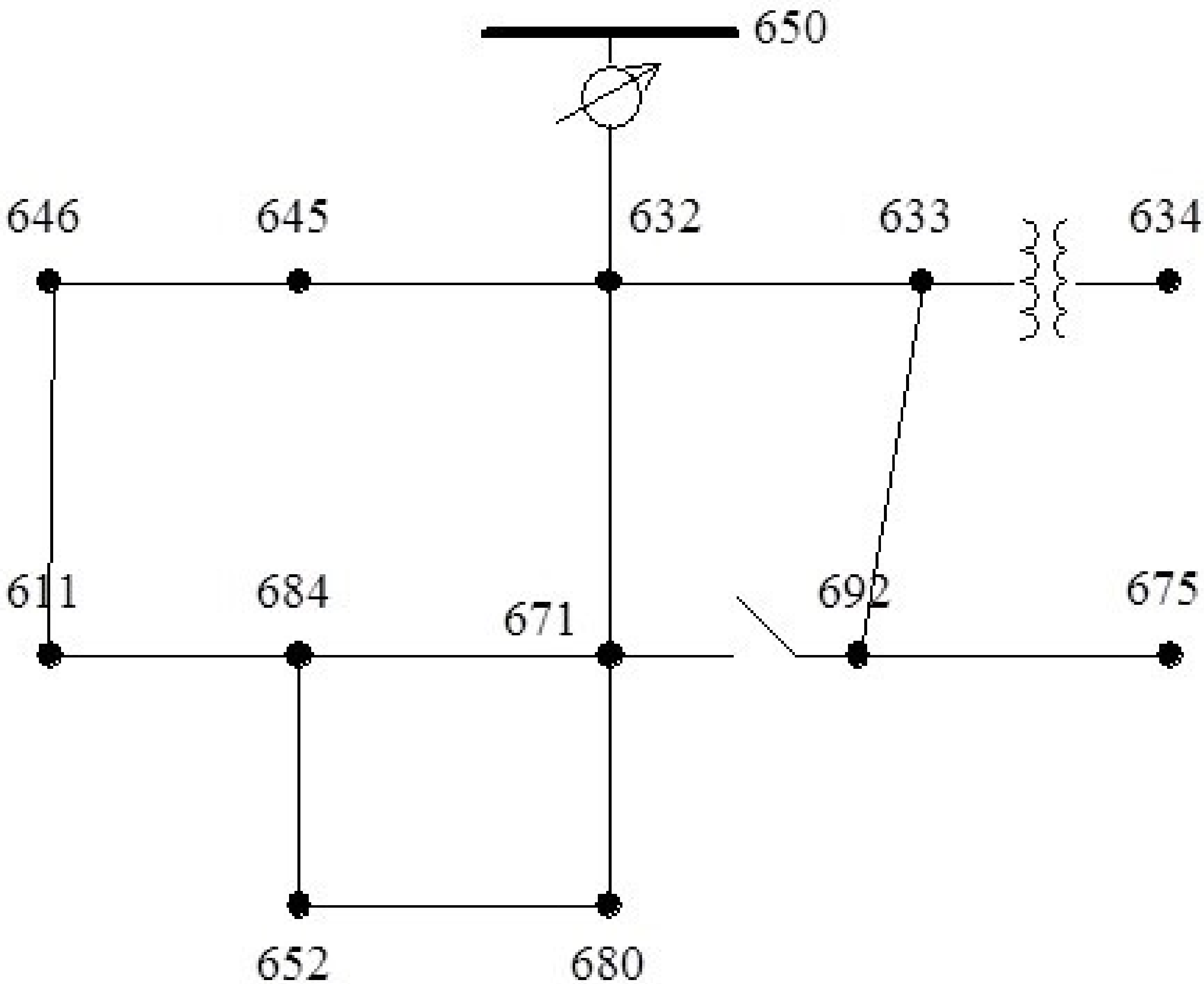


Figure 6.1: A Modified IEEE 13 node test feeder with 2 additional lines.

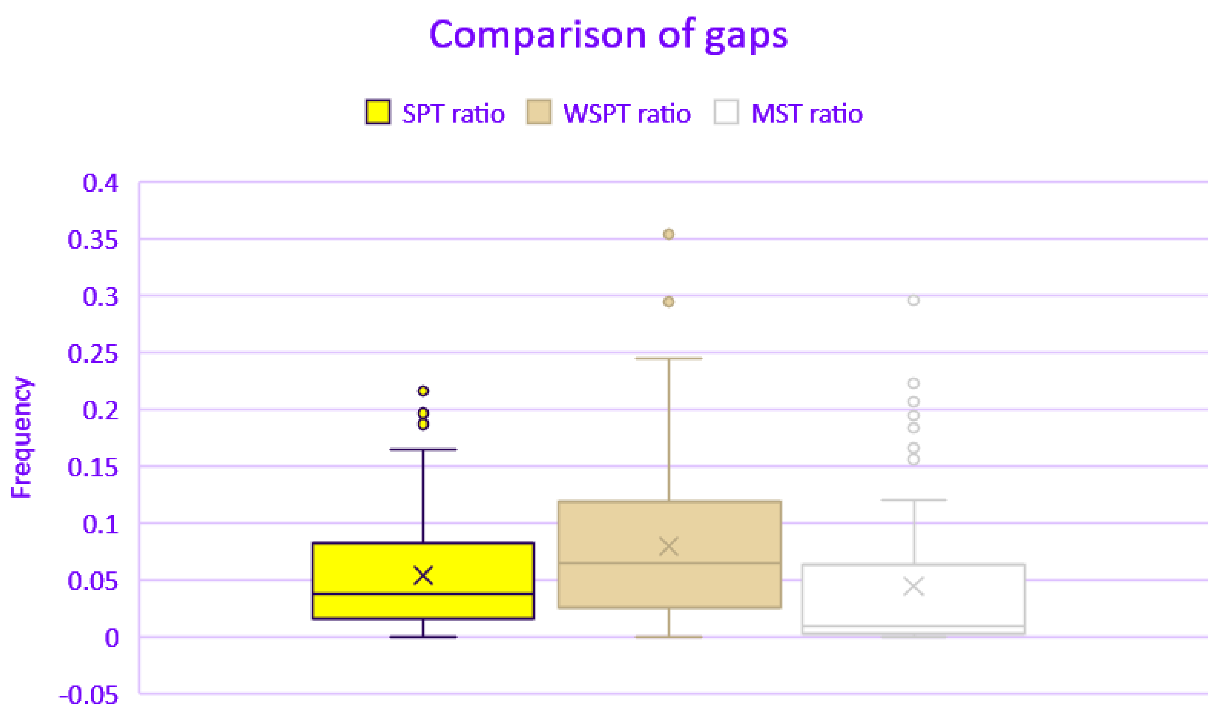


Figure 6.2: Numerical Results of 3 spanning tree selecting strategies

BIBLIOGRAPHY

Adibi, M. M. & Fink, L. H. (1994), 'Power system restoration planning', *IEEE Transactions on Power Systems* **9**(1), 22–28.

Adibi, M. M. & Fink, L. H. (2006), 'Overcoming restoration challenges associated with major power system disturbances - restoration from cascading failures', *IEEE Power and Energy Magazine* **4**(5), 68–77.

Adolphson, D. & Hu, T. C. (1973), 'Optimal linear ordering', *SIAM Journal on Applied Mathematics* **25**(3), 403–423.

Amin, M. (2001), 'Toward self-healing energy infrastructure systems', *IEEE Computer Applications in Power* **14**(1), 20–28.

Arif, A., Wang, Z., Wang, J. & Chen, C. (2017), 'Power distribution system outage management with co-optimization of repairs, reconfiguration, and dg dispatch', *IEEE Transactions on Smart Grid* pp. 1–1.

Arritt, R. F. & Dugan, R. C. (2010), The IEEE 8500-node test feeder, *in* 'IEEE PES Transmission and Distribution Conference and Exposition', pp. 1–6.

Barthelemy, M. (2004), 'Betweenness centrality in large complex networks', *The European Physical Journal B-Condensed Matter and Complex Systems* **38**(2), 163–168.

Berkeley III, A. R., Wallace, M. & COO, C. (2010), 'A framework for establishing critical infrastructure resilience goals', https://www.dhs.gov/xlibrary/assets/niac/niac-a-framework-for-establishing-critical-infrastructure-resilience-goals-2010-10-19.pdf. [Online; accessed 24-April-2018].

Bier, V. M., Gratz, E. R., Haphuriwat, N. J., Magua, W. & Wierzbicki, K. R. (2007), 'Methodology for identifying near-optimal interdiction strategies for a power transmission system', *Reliability Engineering & System Safety* **92**(9), 1155–1161.

Billinton, R. & Wojczynski, E. (1985), 'Distributional variation of distribution system reliability indices', *IEEE Transactions on Power Apparatus and Systems* (11), 3151–3160.

Breiding, R. (2015), Determining the Damage State of the Electrical Distribution System Following an Earthquake, PhD thesis.

Brown, G., Carlyle, M., Salmerón, J. & Wood, K. (2006), 'Defending critical infrastructure', *Interfaces* **36**(6), 530–544.

Brown, R. E., Gupta, S., Christie, R. D., Venkata, S. S. & Fletcher, R. (1997), 'Distribution system reliability assessment: momentary interruptions and storms', *IEEE Transactions on Power Delivery* **12**(4), 1569–1575.

Brucker, P. (2007), *Scheduling algorithms*, Vol. 3, Springer.

Bruneau, M., Chang, S. E., Eguchi, R. T., Lee, G. C., O'Rourke, T. D., Reinhorn, A. M., Shinozuka, M., Tierney, K., Wallace, W. A. & von Winterfeldt, D. (2003), 'A framework to quantitatively assess and enhance the seismic resilience of communities', *Earthquake spectra* **19**(4), 733–752.

Çağnan, Z., Davidson, R. A. & Guikema, S. D. (2006), 'Post-earthquake restoration planning for Los Angeles Electric Power', *Earthquake Spectra* **22**(3), 589–608.

Chanda, S. & Srivastava, A. K. (2016), 'Defining and enabling resiliency of electric distribution systems with multiple microgrids', *IEEE Transactions on Smart Grid* **7**(6), 2859–2868.

Chekuri, C. & Khanna, S. (2004), Approximation algorithms for minimizing average weighted completion time, *in* 'Handbook of Scheduling: Algorithms, Models, and Performance Analysis', CRC Press.

Chen, C., Wang, J., Qiu, F. & Zhao, D. (2015), 'Resilient distribution system by microgrids formation after natural disasters', *Smart Grid, IEEE Transactions on* **PP**(99), 1–1.

Coffrin, C. & Van Hentenryck, P. (2014), Transmission system restoration: Co-optimization of repairs, load pickups, and generation dispatch, *in* 'Power Systems Computation Conference (PSCC), 2014', IEEE, pp. 1–8.

Cohen, R. & Havlin, S. (2010), *Complex networks: structure, robustness and function*, Cambridge University Press.

Edison Electric Institute (n.d.), 'Understanding the Electric Power Industry's Response and Restoration Process', http://www.eei.org/issuesandpolicy/electricreliability/mutualassistance/Documents/MA_101FINAL.pdf. [Online; accessed 11-March-2018].

Electric Power Research Institute (2013), 'Enhancing Distribution Resiliency: Opportunities for Applying Innovative Technologies', pp. 1–20.

Electric Reliability Council of Texas (2017), 'Electric Reliability Council of Texas, Inc.'s report concerning impacts of Hurricane/Tropical Storm Harvey', http://www.ercot.com/content/wcm/lists/114741/27706_-_2017.08.30_-__Harvey_Report_to_PUC.pdf. [Online; accessed 20-May-2018].

Executive Office of the President (2013), 'Economic benefits of increasing electric grid resilience to weather outages', https://www.energy.gov/downloads/economic-benefits-increasing-electric-grid-resilience-weather-outages. [Online; accessed 11-March-2018].

FirstEnergy Group (n.d.), 'Storm Restoration Process', https://www.firstenergycorp.com/content/customer/help/outages/storm_restorationprocess.html. [Online; accessed 11-March-2018].

Garey, M. R., Johnson, D. S. & Stockmeyer, L. (1976), 'Some simplified NP-complete graph problems', *Theoretical computer science* **1**(3), 237–267.

Graham, R. L., Lawler, E. L., Lenstra, J. K. & Kan, A. R. (1979), 'Optimization and approximation in deterministic sequencing and scheduling: a survey', *Annals of discrete mathematics* **5**, 287–326.

Halin, R. (1976), 'S-functions for graphs', *Journal of geometry* **8**(1-2), 171–186.

Hines, P. & Blumsack, S. (2008), A centrality measure for electrical networks, *in* 'Hawaii International Conference on System Sciences, Proceedings of the 41st Annual', IEEE, pp. 185–185.

Horn, W. (1972), 'Single-machine job sequencing with treelike precedence ordering and linear delay penalties', *SIAM Journal on Applied Mathematics* **23**(2), 189–202.

Horst, R. & Tuy, H. (2013), *Global optimization: Deterministic approaches*, Springer Science & Business Media.

Hou, Y., Liu, C.-C., Sun, K., Zhang, P., Liu, S. & Mizumura, D. (2011), Computation of milestones for decision support during system restoration, *in* '2011 IEEE Power and Energy Society General Meeting', IEEE, pp. 1–10.

IEEE Guide for Electric Power Distribution Reliability Indices - Redline (2012), *IEEE Std 1366-2012 (Revision of IEEE Std 1366-2003) - Redline* pp. 1–92.

Infrastructure Security and Energy Restoration, Office of Electricity Delivery and Energy Reliability, U.S. Department of Energy (2012), 'A review of power outages and restoration following the June 2012 Derecho', http://energy.gov/sites/prod/files/Derecho%202012_%20Review_0.pdf.

Jensen, J. L. W. V. (1906), 'Sur les fonctions convexes et les inégalités entre les valeurs moyennes', *Acta mathematica* **30**(1), 175–193.

Kameshwaran, S. & Narahari, Y. (2009), 'Nonconvex piecewise linear knapsack problems', *European Journal of Operational Research* **192**(1), 56–68.

Kenny Mercado, Senior Vice President, Electric Operations (2017), 'ERCOT Board of Directors - CenterPoint Energy's Response to Hurricane Harvey', http://www.ercot.com/content/wcm/key_documents_lists/103998/5.3.2_CenterPoint_Energy_s_Response_to_Hurricane_Harvey_REVISED_10.12.17.pdf. [Online; accessed 29-June-2018].

Kersting, W. H. (2001), Radial distribution test feeders, *in* '2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194)', Vol. 2, pp. 908–912 vol.2.

Kleywegt, A. J., Shapiro, A. & Homem-de Mello, T. (2002), 'The sample average approximation method for stochastic discrete optimization', *SIAM Journal on Optimization* **12**(2), 479–502.

Kwasinski, A., Eidinger, J., Tang, A. & Tudo-Bornarel, C. (2014), 'Performance of electric power systems in the 2010 - 2011 Christchurch, New Zealand, Earthquake Sequence', *Earthquake Spectra* **30**(1), 205–230.

Lawler, E. L. (1978), 'Sequencing jobs to minimize total weighted completion time subject to precedence constraints', *Annals of Discrete Mathematics* **2**, 75–90.

Lipp, T. & Boyd, S. (2016), 'Variations and extension of the convex–concave procedure', *Optimization and Engineering* **17**(2), 263–287.

Ma, S., Chen, B. & Wang, Z. (2016), 'Resilience enhancement strategy for distribution systems under extreme weather events', *IEEE Transactions on Smart Grid* .

Ma, T.-K., Liu, C.-C., Tsai, M.-S., Rogers, R., Muchlinski, S. & Dodge, J. (1992), 'Operational experience and maintenance of online expert system for customer restoration and fault testing', *IEEE Transactions on Power Systems* **7**(2), 835–842.

Miles, S. B. (2011), The role of critical infrastructure in community resilience to disasters, *in* 'Structures Congress 2011', pp. 1985–1995.

Mili, L. & Center, N. V. (2011), Taxonomy of the characteristics of power system operating states, *in* '2nd NSF-VT Resilient and Sustainable Critical Infrastructures (RESIN) Workshop, Tucson, AZ, Jan', pp. 13–15.

Möhring, R. H., Radermacher, F. J. & Weiss, G. (1984), 'Stochastic scheduling problems i — general strategies', *Zeitschrift für Operations Research* **28**(7), 193–260.
URL: *http://dx.doi.org/10.1007/BF01919323*

Möhring, R. H., Schulz, A. S. & Uetz, M. (1999), 'Approximation in stochastic scheduling: the power of LP-based priority policies', *Journal of the ACM (JACM)* **46**(6), 924–942.

Möhring, R., Radermacher, F. & Weiss, G. (1985), 'Stochastic scheduling problems ii — set strategies', *Zeitschrift für Operations Research* **29**(3), 65–104.

Nagarajan, H., Yamangil, E., Bent, R., Van Hentenryck, P. & Backhaus, S. (2016), Optimal resilient transmission grid design, *in* 'Power Systems Computation Conference (PSCC), 2016', IEEE, pp. 1–7.

National Infrastructure Advisory Council (2010), 'A framework for establishing critical infrastructure resilience goals: Final report and recommendations', https://www.dhs.gov/publication/niac-framework-establishing-resilience-goals-final-report. [Online; accessed 25-March-2018].

NERC (2014), 'Hurricane sandy event analysis report', http://www.nerc.com/pa/rrm/ea/Oct2012HurricanSandyEvntAnlyssRprtDL/Hurricane_Sandy_EAR_20140312_Final.pdf. [Online; accessed 11-March-2018].

Nojima, N. & Kameda, H. (1992), Optimal strategy by use of tree structure for post-earthquake restoration of lifeline network systems, *in* 'Proceedings of the 10th World Conference on Earthquake Engineering', pp. 5541–5546.

Nowicki, E. & Zdrzałka, S. (1990), 'A survey of results for sequencing problems with controllable processing times', *Discrete Applied Mathematics* **26**(2-3), 271–287.

Nurre, S. G., Cavdaroglu, B., Mitchell, J. E., Sharkey, T. C. & Wallace, W. A. (2012), 'Restoring infrastructure systems: An integrated network design and scheduling (INDS) problem', *European Journal of Operational Research* **223**(3), 794–806.

Nurre, S. G. & Sharkey, T. C. (2014), 'Integrated network design and scheduling problems with parallel identical machines: Complexity results and dispatching rules', *Networks* **63**(4), 306–326.
URL: *http://dx.doi.org/10.1002/net.21547*

Obama, B. (2013), 'Preparing the united states for the impacts of climate change', *Executive Order* **13653**, 66819–66824.

Office for Eletricity Delivery and Energy Reliability, U.S. Department of Energy (2016), 'Distribution Automation: Results from the smart grid investment grant program', https://www.energy.gov/sites/prod/files/2016/11/f34/Distribution%20Automation%20Summary%20Report_09-29-16.pdf. [Online; accessed 7-July-2018].

Omer, M. (2013), *The Resilience of Networked Infrastructure Systems: Analysis and Measurement*, Vol. 3, World Scientific.

on Critical Infrastructure Protection, P. C. (1997), 'Critical Foundations: Protecting America's Infrastructures', https://fas.org/sgp/library/pccip.pdf. [Online; accessed 24-April-2018].

O'Rourke, T. D. (2007), 'Critical infrastructure, interdependencies, and resilience', *The Bridge* **37**(1), 22.

Ouyang, M. & Dueñas-Osorio, L. (2014), 'Multi-dimensional hurricane resilience assessment of electric power systems', *Structural Safety* **48**, 15–24.

Ouyang, M. & Fang, Y. (2017), 'A mathematical framework to optimize critical infrastructure resilience against intentional attacks', *Computer-Aided Civil and Infrastructure Engineering* .

Panteli, M., Mancarella, P., Trakas, D. N., Kyriakides, E. & Hatziargyriou, N. D. (2017), 'Metrics and quantification of operational and infrastructure resilience in power systems', *IEEE Transactions on Power Systems* **32**(6), 4732–4742.

Patton, A. (1979), Probability distribution of transmission and distribution reliability performance indices, *in* 'Reliability Conference for Electric Power Industry', pp. 120–122.

Pérez-Guerrero, R., Heydt, G. T., Jack, N. J., Keel, B. K. & Castelhano, A. R. (2008), 'Optimal restoration of distribution systems using dynamic programming', *IEEE Transactions on Power Delivery* **23**(3), 1589–1596.

Pinedo, M. L. (2012), *Scheduling: theory, algorithms, and systems*, Springer Science & Business Media.

Prada, J. F. (1999), 'The value of reliability in power systems-pricing operating reserves', http://web.mit.edu/energylab/www/pubs/el99-005wp.pdf.

Queyranne, M. (1993), 'Structure of a simple scheduling polyhedron', *Mathematical Programming* **58**(1-3), 263–285.

Queyranne, M. & Schulz, A. S. (2006), 'Approximation bounds for a general class of precedence constrained parallel machine scheduling problems', *SIAM Journal on Computing* **35**(5), 1241–1253.

Radermacher, F. (1981), 'Cost-dependent essential systems of es-strategies for stochastic scheduling problems', *Methods of Operations Research* **42**, 17–31.

Reed, D. A., Kapur, K. C. & Christie, R. D. (2009), 'Methodology for Assessing the Resilience of Networked Infrastructure', *IEEE Systems Journal* **3**(2), 174–180.

Reed, D. A., Powell, M. D. & Westerman, J. M. (2010), 'Energy supply system performance for hurricane katrina', *Journal of Energy Engineering* **136**(4), 95–102.

Reinhorn, A. M., Bruneau, M. & Cimellaro, G. P. (2010), 'Framework for analytical quantification of disaster resilience', *Engineering Structures* **32**(11), 3639–3649.

Rinaldi, S. M., Peerenboom, J. P. & Kelly, T. K. (2001), 'Identifying, understanding, and analyzing critical infrastructure interdependencies', *IEEE Control Systems* **21**(6), 11–25.

Robertson, N. & Seymour, P. D. (1984), 'Graph minors. iii. planar tree-width', *Journal of Combinatorial Theory, Series B* **36**(1), 49–64.

Rollins, M. (2007), 'The hardening of utilitiy lines-implications for utility pole design and use', http://woodpoles.org/portals/2/documents/TB_HardeningUtilityLines.pdf. [Online; accessed 25-March-2018].

Romero, N., Nozick, L. K., Dobson, I., Xu, N. & Jones, D. A. (2015), 'Seismic retrofit for electric power systems', *Earthquake Spectra* **31**(2), 1157–1176.

Sarfi, R. J., Salama, M. & Chikhani, A. (1994), 'A survey of the state of the art in distribution system reconfiguration for system loss reduction', *Electric Power Systems Research* **31**(1), 61–70.

Schulz, A. S. et al. (1996), Polytopes and scheduling, PhD thesis, Citeseer.

Sharkey, T. C., Cavdaroglu, B., Nguyen, H., Holman, J., Mitchell, J. E. & Wallace, W. A. (2015), 'Interdependent network restoration: On the value of information-sharing', *European Journal of Operational Research* **244**(1), 309–321.

Shioura, A., Shakhlevich, N. V. & Strusevich, V. A. (2016), 'Application of submodular optimization to single machine scheduling with controllable processing times subject to release dates and deadlines', *INFORMS Journal on Computing* **28**(1), 148–161.

Sidney, J. B. (1975), 'Decomposition algorithms for single-machine sequencing with precedence relations and deferral costs', *Operations Research* **23**(2), 283–298.

Simic, S. (2008), 'On a global upper bound for Jensen's inequality', *Journal of Mathematical Analysis and Applications* **343**(1), 414–419.

Sinha, P. & Zoltners, A. A. (1979), 'The multiple-choice knapsack problem', *Operations Research* **27**(3), 503–515.

Smith, W. E. (1956), 'Various optimizers for single-stage production', *Naval Research Logistics Quarterly* **3**(1-2), 59–66.

Strbac, G., Kirschen, D. & Moreno, R. (2016), 'Reliability standards for the operation and planning of future electricity networks', *Foundations and Trends® in Electric Energy Systems* **1**(3), 143–219.
URL: *http://dx.doi.org/10.1561/3100000001*

Sun, W., Liu, C.-C. & Zhang, L. (2011), 'Optimal generator start-up strategy for bulk power system restoration', *IEEE Transactions on Power Systems* **26**(3), 1357–1366.

The City of New York (2013), 'A stronger, more resilient New York', http://www.nyc.gov/html/sirr/html/report/report.shtml. [Online; accessed 24-April-2018].

The GridWise Alliance (2013), 'Improving electric grid reliability and resilience: Lessons learned from superstorm sandy and other extreme events', http://www.gridwise.org/documents/ImprovingElectricGridReliabilityandResilience_6_6_13webFINAL.pdf. [Online; accessed 11-March-2018].

Ton, D. T. & Wang, W. P. (2015), 'A more resilient grid: The US Department of Energy joins with stakeholders in an R&D plan', *IEEE Power and Energy Magazine* **13**(3), 26–34.

Toune, S., Fudo, H., Genji, T., Fukuyama, Y. & Nakanishi, Y. (2002), 'Comparative study of modern heuristic algorithms to service restoration in distribution systems', *IEEE Transactions on Power Delivery* **17**(1), 173–181.

Uetz, M. (2001), Algorithms for deterministic and stochastic scheduling, PhD thesis.

U.S. Energy Information Administation (2017), 'Hurricane Harvey caused electric system outages and affected wind generation in Texas', https://www.eia.gov/todayinenergy/detail.php?id=32892. [Online; accessed 29-June-2018].

Van Hentenryck, P. & Coffrin, C. (2015), 'Transmission system repair and restoration', *Mathematical Programming* **151**(1), 347–373.

Vugrin, E. D., Warren, D. E. & Ehlen, M. A. (2011), 'A resilience assessment framework for infrastructure and economic systems: Quantitative and qualitative resilience analysis of petrochemical supply chains to a hurricane', *Process Safety Progress* **30**(3), 280–290.

Wang, Y., Chen, C., Wang, J. & Baldick, R. (2015), 'Research on resilience of power systems under natural disasters - a review', *Power Systems, IEEE Transactions on* **PP**(99), 1–10.

Wang, Z. & Cui, Z. (2012), 'Combination of parallel machine scheduling and vertex cover', *Theoretical Computer Science* **460**, 10–15.

Wang, Z., Wang, J. & Chen, C. (2016), 'A three-phase microgrid restoration model considering unbalanced operation of distributed generation', *IEEE Transactions on Smart Grid* **PP**(99), 1–1.

Willis, H. H. & Loa, K. (2015), 'Measuring the resilience of energy distribution systems', *RAND Corporation, Santa Monica, Calif* .

Xu, N., Guikema, S. D., Davidson, R. A., Nozick, L. K., Çağnan, Z. & Vaziri, K. (2007), 'Optimizing scheduling of post-earthquake electric power restoration tasks', *Earthquake engineering & structural dynamics* **36**(2), 265–284.

Yamangil, E., Bent, R. & Backhaus, S. (2015*a*), 'Designing resilient electrical distribution grids', *Proceedings of the 29th Conference on Artificial Intelligence, Austin, Texas* .

Yamangil, E., Bent, R. & Backhaus, S. (2015*b*), Resilient upgrade of electrical distribution grids, *in* 'Twenty-Ninth AAAI Conference on Artificial Intelligence'.

Yuan, W., Wang, J., Qiu, F., Chen, C., Kang, C. & Zeng, B. (2016), 'Robust optimization-based resilient distribution network planning against natural disasters', *IEEE Transactions on Smart Grid* **7**(6), 2817–2826.

Yuan, W., Zhao, L. & Zeng, B. (2014), 'Optimal power grid protection through a defender–attacker–defender model', *Reliability Engineering & System Safety* **121**, 83–89.

VITA

Yushi Tan received a B.S. degree from Department of Electrical Engineering at Tsinghua University, Beijing, China in June, 2013. He has been working on a Doctorate degree in Electrical Engineering and a Master degree in Statistics at the University of Washington since September of 2013 and 2015 respectively. He was a student employee at Electric Power Research Institute (EPRI) during the summer of 2014 and a research aide at Argonne National Lab during the summer of 2015.