HIGH-PERFORMANCE COMPUTING FOR SMART GRID ANALYSIS AND OPTIMIZATION

BY

YI LIANG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Deming Chen, Chair
Professor Martin D. F. Wong
Assistant Professor Jian Huang
Assistant Professor Hao Zhu

ABSTRACT

The smart grid leverages a variety of advanced technologies, including smart metering, smart equipment, communication and control technologies, renewable energy sources, and machine learning, to improve the efficiency and reliability of existing electric power systems. The efficiency and reliability of power systems are of considerably importance to economic and environmental health in this new era. However, there are significant challenges for modernizing the power grids and accomplishing the vision of the smart grid. This dissertation presents a variety of optimization techniques that solve several key challenges in the smart grid and improve its efficiency and reliability.

Optimal power flow (OPF) plays an important role in power system operation. The emerging smart grid aims to create an automated energy delivery system that enables two-way flows of electricity and information. As a result, it will be desirable if OPF can be solved in real time in order to allow the implementation of the time-sensitive applications such as real-time pricing. We develop a novel method, the fast OPF algorithm, to accelerate the computation of alternating current optimal power flow (ACOPF). We first construct and solve an equivalent OPF problem for an equivalent reduced system. Then, a distributed algorithm is developed to retrieve the optimal solution for the original power system. Experimental results show that for a large power system, our method achieves 7.01X speedup over ACOPF with only 1.72% error, and is 75.7% more accurate than the DCOPF solution. The experimental results demonstrate the unique strength of the proposed technique for fast, scalable, and accurate OPF computation.

With the integration of intermittent renewable energy sources and demand response in the smart grid, there is increasing uncertainty involved in the traditional OPF problem. Therefore, probabilistic optimal power flow (POPF) analysis is required to accomplish the electrical and economic operational goals. We propose a novel method, the ClusRed algorithm, to

accelerate the computation of POPF for large-scale smart grid through clustering and network reduction (NR). A cumulant-based method and Gram-Charlier expansion theory are used to efficiently obtain the statistics of system states. We also develop a more accurate linear mapping method to compute the unknown cumulants. ClusRed can speed up the computation by up to 4.57X and can improve accuracy by about 30% when Hessian matrix is ill-conditioned compared to the previous approach.

Aside from improving the efficiency and reliability of power grids through addressing OPF related problems, we also study geomagnetic disturbances (GMDs) and how to mitigate their threat to the reliability of power grids. Geomagnetically induced currents (GICs) introduced by GMDs can damage transformers, increase reactive power losses and cause reliability issues in power systems. Finding an optimal strategy to place blocking devices (BDs) at transformer neutrals is essential to mitigating the negative impact of GICs. We develop a branch and cut (BC) based method and demonstrate that the BC method can provide optimal solutions to OBP problems. Furthermore, to practically solve the OBP problem, it is also important to account for the potential impact of BD placement on neighboring interconnected systems, solve the case where per-transformer GIC constraint exists and take the time-varying nature of the geoelectric field into consideration. In addition, the combined complexity of solving the OBP problem on a large-scale system poses a big computational challenge. However, together with other existing methods, the BC method cannot address the above issues well due to its algorithmic limitations. We then develop a simulated annealing (SA) based algorithm that, for the first time, can achieve near-optimal solutions for OBP problems for the above scenarios at a reduced computational complexity. More importantly, the SA method provides a comprehensive framework that can be used to solve various OBP problems, with different objective functions and constraints. We demonstrate the effectiveness and efficiency of our BC and SA methods using power systems of various sizes.

In addition to natural disasters, in the era of internet of things, cybersecurity is of growing concern to power industries. Malicious cyberbehaviors and technologies that used to challenge security in areas unrelated to power systems, such as information integrity or privacy, have suddenly started to endanger the safety of large-scale smart grids. In particular, short-term load forecasting (STLF) is one of many aspects that are subject to these attacks.

STLF systems have demonstrated high accuracy and have been widely employed for commercial use. However, classic load forecasting systems, which are based on statistical methods, are vulnerable to training data poisoning. We build and implement a first-of-its-kind data poisoning strategy that is effective at corrupting the forecasting model even in the presence of outlier detection. Our method applies to several forecasting models, including the most widely adapted and best-performing ones, such as multiple linear regression (MLR) and neural network (NN) models. Starting with the MLR model, we develop a novel closed-form solution that enables us to quickly estimate the new MLR model after a round of data poisoning without retraining. We then employ line search and simulated annealing to find the poisoning attack solution. Furthermore, we use the MLR attacking solution to generate a numerical solution for other models, such as NN. The effectiveness of our algorithm has been demonstrated on the Global Energy Forecasting Competition (GEFCom2012) data set with the presence of outlier detection.

To my parents, for their endless love and support.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my adviser Prof. Deming Chen for providing excellent guidance, invaluable feedback, and care during my PhD. I have been extremely lucky to have a supervisor who encouraged me to be bold and mentored me to be persistent. This dissertation would not be possible without your advice and wisdom. I would like to thank Prof. Wong, my dissertation committee member, for generously providing his friendly advice and expertise, and for his inspiring lectures at the University of Illinois at Urbana-Champaign. I would like to thank Prof. Hao Zhu and Prof. Thomas Overbye for their insightful comments and constructive suggestions. Their invaluable opinions have significantly improved the quality of this dissertation.

I would like to thank Dr. Di He for the inspiring discussions and collaborations. It has been a great pleasure to work with him. I also want to express my grateful thanks to my friends in the lab for the meaningful discussions and their support, especially Dr. Daifeng Guo, Dr. Yao Chen, Mr. Zelei Sun, Dr. Chen-hsuan Lin, Ms. Wei Zuo and Mr. Xinheng Liu.

Finally, I give my deepest gratitude to my parents for their endless love and support. I want to thank them for raising me up, putting all their efforts in educating me and always keeping their faith in me.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Energy consumption has become an increasingly important issue worldwide. Climate change, rising energy prices and increasing power demand have reshaped the existing energy system. In this situation, electricity, which is generated by consuming more than 40% of the total energy [1], is the key to the overall efficiency of the energy system. The emerging smart grid is aimed at creating a reliable and efficient power system by taking advantage of information technologies, renewable energy, smart digital devices, machine learning and so on. However, it is hard and time-consuming to find the optimal solutions to many operational and planning problems. Therefore, this dissertation develops new high-performance computing algorithms that can improve the operational efficiency and reliability of smart grids by solving several smart grid optimization problems efficiently.

We have conducted several studies to improve the efficiency of solving optimal power flow (OPF) problems in large-scale power systems, such as deterministic optimal power flow (DOPF) and probabilistic optimal power flow (POPF). In addition, we also consider the threats of geomagnetic disturbances (GMDs) and solve the optimal blocking device placement (OBP) problems that can improve the reliability of power system. We propose a branch and cut (BC) based method that can find optimal solution and a simulated annealing (SA) based method that can solve the scenario of considering neighboring impact, per-transformer constraint and the time-varying feature of geoelectric field. We also studied short-term load forecasting (STLF) and developed a poisoning attack strategy that is effective in the presence of outlier detector. Future work will be discussed to advance this study and further improve the efficiency and reliability of the smart grid.

Deterministic alternating current optimal power flow (ACOPF) is widely used in power systems for making operational decisions [2–4]. It will be desirable if OPF can be solved in real time in order to allow the imple-

mentation of time-sensitive applications, such as real-time pricing. However, many existing methods, such as quadratic programming [5], genetic algorithms (GA) [6] and particle swarm optimization [7] are not computationally efficient and cannot be used in large-scale power systems for real-time operation [8–12]. Therefore, we propose a novel algorithm to accelerate the computation of ACOPF through power system network reduction (NR) [13]. We formulate the ACOPF problem based on an equivalent reduced system and interpret its solution. As a result, the detailed optimal dispatch for the original power system can be obtained afterwards using a distributed algorithm. Our experimental results demonstrated the unique strength of the proposed technique for fast, scalable, and accurate ACOPF computation.

With the integration of renewable energy, demand response in the smart grid, there are many considerable uncertainties in the system [14]. Probabilistic optimal power flow (OPF) analysis is required to accomplish the electrical and economic operational goals [15]. However, existing methods suffer from slow speed and cannot provide accurate solutions when the Hessian matrix is close to ill-conditioned [16–19]. Thus, we propose a network reduction based method to speed up the POPF computation. We also take advantage of a cumulant-based method and Gram-Charlier expansion theory [20], which are well-known for their efficiency in computing the probabilistic distribution functions. This work shows that the proposed method can solve POPF much faster with better accuracy.

The quasi-dc geomagnetically induced currents (GICs) introduced by geomagnetic disturbances (GMDs) have been concerning power grid operators for many years [21, 22]. Over the last several years, the power industry has seen more concentrated interest in this area. The US Federal Energy Regulation Commission (FERC) now requires US utilities to perform GMD vulnerability assessment, and prepare corrective action plans and mitigation actions [23]. GMDs, such as solar storms, can cause rapid geomagnetic field variation, which produces geoelectric field and GICs flowing through transmission lines. GICs tend to flow through the neutral connection of transformers and can cause half-cycle saturation of transformers. As a result, power systems may suffer from transformer overheating and severe reactive power losses. GICs can damage the bulk power system assets, typically associated with transformers. This may eventually lead to system reliability issues, such as misoperations of proactive relays, and voltage instability [24]. The

impacts of GIC, GIC modeling and analysis are studied in [25–30]. Many methods were proposed to mitigate the effects of GIC [31, 32]. One widely recognized solution is to install blocking capacitors or switchable resistors to the neutral connection of Wye-connected transformers [33, 34]. In order to find the best way to protect power grids, optimal blocking device placement (OBP) was studied to minimize the damage of GICs. Several methods were proposed to formulate and solve the BD placement problem [35–37]. However, existing methods could not guarantee optimality in the BD placement problem when minimizing GMDs' damages. We formulate this OBP problem as a mixed integer nonlinear programming (MINLP) problem and develop a branch and cut based method [38] to solve this OBP problem. Nonetheless, the BC method does not scale up well to large-scale power systems. Together with all other existing methods, there are significant limitations that stop them from being used for practical OBP problems with many practical constraints. For example, the impact of BD placement on neighboring interconnected systems needs to be studied and evaluated in OBP problem. It is also required to use a realistic time-varying geoelectric field (TVGF) in OBP problems. In addition, per-transformer GIC constraint needs to be included in the constraint of OBP problems. The community is in need of a new algorithm that addresses the issues mentioned above and can produce a high-quality solution in the given affordable computational time. We develop a simulated annealing (SA) [39, 40] based algorithm that can accelerate the solution process, produce a near-optimal solution in a relative short time, and solve the above scenarios for OBP problems.

In the era of the internet of things, cybersecurity is of growing concern to power industries [41]. As power systems benefit from stronger connectivity and advanced probabilistic modeling, they also become more vulnerable to attacks that target these aspects. Today, accurately conducting short-term load forecasting (STLF) is essential to power systems [42]. The power industry relies heavily on accurate predictions to increase efficiency, reduce waste and maintain stability. If the forecast is corrupted, not only could there be financial losses, in extreme but realistic cases the bulk power system assets could be damaged, resulting in safety hazards. We study STLF and develop a data poisoning algorithm that reduces forecasting accuracy without setting off an outlier detector. The efficiency and reliability of smart grids are facing significant challenges.

The contributions and results of this dissertation are summarized as follows:

- We propose a fast ACOPF analysis framework through power system network reduction to speed up the computation of ACOPF problems [43]. We demonstrate that this approach can achieve 1.32X to 7.01X speedup over full ACOPF while introducing just 0.54% error on average. Compared to the widely used DCOPF, we reduce the error by 77.6% on average. If ACOPF can converge to the optimal solution, our proposed method can find an optimal solution, which demonstrates the robustness of our algorithm. The proposed method can be used to solve ACOPF for large-scale power systems in many applications, such as operational reliability analysis and power market management.

- ClusRed [44] is proposed to accelerate the computation of probabilistic OPF for large-scale smart grids through network reduction (NR). A cumulant-based method and Gram-Charlier expansion theory are used to efficiently obtain the statistics of system states. We develop a more accurate linear mapping method to compute the unknown cumulants. Our method speeds up the computation by up to 4.57X and can improve accuracy by about 30% when the Hessian matrix is ill-conditioned compared to the previous approach.

- By taking advantage of the special structure of the OBP problems, a branch and cut method is proposed to solve the OBP problem and address the effects of GMDs [45]. The effectiveness and optimality of our method is demonstrated on a 6-bus system, a 20-bus system and a much larger realistically sized system. Our method can provide optimal solution to OBP problems and can also be used to solve variations of OBP problems with different objective functions.

- We develop a fast SA method that can produce high-quality near-optimal solutions for OBP problems. For the first time, we study the BD placement impact on interconnected systems and solve for the BBS scenario; we study the time-varying nature of geoelectric field and solve for the TVGF scenario. We also demonstrate that our SA method can be easily extended to solve OBP problems with various configurations.

- We develop a poisoning strategy that can corrupt energy load forecasting model even in the presence of outlier detection [46]. Our closed-form model estimation technique, line search and simulated annealing based methods have been shown to be effective in the presence of an outlier detector.

The rest of the dissertation is organized as follows: Chapter 2 presents our work on accelerating deterministic ACOPF. Chapter 3 presents ClusRed, our distributed POPF solver. We then present our BC and SA method for solving OBP problems in Chapter 4 and Chapter 5. Chapter 6 presents our poisoning attack strategy for STLF systems. We conclude this dissertation and discuss the future directions in Chapter 7.

CHAPTER 2

FAST OPTIMAL POWER FLOW ANALYSIS FOR LARGE-SCALE SMART GRID

2.1 Introduction

The power system in United States is one of the largest and most complex cyber-physical systems in the world [1]. To support its automation, power systems need to monitor, control, and secure the grid in real time for efficient and reliable operation. Nowadays, the emerging smart grid aims to enable two-way flows of information and electricity to create an automated and advanced energy system with different decision makers involved. Timely and accurate analysis and control of such a large system are vitally important for its operating reliability and efficiency. Inaccurate or slow analysis of the power system may result in uneconomic operation of the grid and potential environmental pollution [2].

OPF has been widely used in power system planning and operation in the last 50 years, and seeks to optimize an objective function by adjusting a set of control variables subject to certain physical, operational, and policy constraints. However, even today, full ACOPF has not been widely adopted in real-time operations for large-scale power systems because of the high computational requirement. In the smart grid paradigm, the problem size grows tremendously with the integration of renewable energy, energy storage, and demand response. In addition, a more detailed model is needed to support various emerging applications, which further aggravates the computational burden. With the advent of the wholesale electricity market, ACOPF computation is now part of the core pricing mechanism for electricity pricing and trading. For example, an ultimate goal of the independent system operator (ISO) is to solve the security-constrained ACOPF over large-scale power systems. Typically, this problem must be solved daily in 2 hours, hourly in 15 minutes, and every 5 minutes in 1 minute by the ISO [3]. Currently, the prob-

lem is solved through various levels of approximations based on application and time sensitivity [4].

Although a highly nonlinear full ACOPF would provide the most accurate control settings in power system operations, due to the high computational demands of ACOPF, DCOPF is widely used. However, since DCOPF uses a linear approximation of the power flow equations and the lossless DC power flow assumption (the so-called DC power flow assumption), it is not accurate, and the assumption of neglecting reactive power and power losses largely limits its application to real-world problems [47]. Currently, people use various approximation techniques and engineering judgments to explore reasonable solutions to the ACOPF problem [4]. However, today's inaccurate approximation may unnecessarily cost billions of dollars annually because of the use of inaccurate OPF solutions [2]. It may also result in environmental pollution from unnecessary emissions and wasted energy [2]. As a result, accelerating ACOPF computation while maintaining high accuracy is very important.

A wide variety of optimization techniques have been examined to solve the non-convex ACOPF problems, such as quadratic programming [5], linear programming [8] and the interior point method [9, 10]. Alternative approaches include genetic algorithms (GA) [6], evolutionary programming [11], steepest descent-based methods [12] and particle swarm optimization [7].

However, these methods are not computationally efficient, and cannot be used in large-scale power systems for real-time operation. A distributed algorithm for the ACOPF problem was proposed in [48], where the OPF problem for the original systems was decomposed into per-area instances. This approach assumes the decoupling between different regions, which is not true for a densely interconnected power system. It can also result in very large border regions, which slow down the convergence and may even cause the problem of non-convergence. In addition, the convergence is not guaranteed unless the objective function of the OPF problem is convex with respect to the border region variables, which is not always true in reality.

There are NR techniques to reduce the computational burden by finding an equivalent system. Some traditional reduction methods, such as the Ward equivalent technique [13], are usually performed by computing the admittance and eliminating unnecessary elements that are not in the study area. The reduced model may lose sparsity and may not yield the same power flow pattern as the original one. In addition, this technique is only used for

power flow analysis. Alternatively, sensitivity matrix-based methods, such as the power transfer distribution factor (PTDF)-based method, are used for NR [16, 17]. The method proposed in [16] preserves the same power flow pattern as that in the original system at the operation set point where the reduction is performed. This method has the operation set point dependence and yields significant error when the system operates at a different set point. In [17], another NR method was proposed to derive an equivalent system that does not depend on the set point. However, both [16] and [17] are proposed for power system long-term planning studies. The generation of the equivalent system is not fast enough, which is not suitable for the real-time power system operating purpose.

In this chapter, a new method based on NR is proposed to solve for ACOPF for the large-scale smart grid. The contributions are:

- We propose a novel method to partition the power network that can efficiently reduce the error brought by NR and a fast analytical approximation method to identify the parameters of the equivalent system without using DC power flow assumption.
- Instead of only considering the reduced equivalent system, we propose a distributed method to efficiently recover the detailed solution for the original system. Congestion and the transmission capacity of lines are considered in the algorithm to ensure the feasibility of the ACOPF solution.
- We provide an effective methodology for scalable computation of the ACOPF problems with high accuracy and speed.

The rest of this chapter is organized as follows. In Section 2.2, we give the necessary background and the ACOPF formulation. Section 2.3 describes the framework and the algorithm of the NR-based ACOPF solution method. We present the numerical results in Section 2.4 and the conclusions in Section 2.5.

2.2 Preliminaries

Over the past 50 years, the steady-state OPF problem was well formulated and many variations of ACOPF formulations were studied. In this chapter,

we begin with the background of power system analysis. We will introduce power flow analysis and ACOPF formulation.

2.2.1 Power Flow Analysis

The power flow equations constitute the steady-state model of the power system and are widely used to compute the system states once the injections and the withdrawals at each network node are specified.

We consider a power system with $N+1$ buses and L lines. We denote by $\mathcal{N} \triangleq \{0, 1, 2, \cdots, N\}$ the set of buses, with the bus 0 being the slack bus, and by $\mathcal{L} \triangleq \{\ell_1, \ell_2, \cdots \ell_L\}$ the set of transmission lines that connect the buses in the set $\mathcal{N}$. We associate with each line $\ell \in \mathcal{L}$ the ordered pair $\ell = (i, j)$. The series admittance of line ℓ is denoted by $g_\ell - jb_\ell$. Each bus i is characterized by the voltage phasor:

$$E_i = V_i e^{j\theta_i}, \tag{2.1}$$

where V_i is the nodal voltage magnitude and θ_i is the nodal voltage phase angle. The net injected complex power at each bus i is

$$S_{net_i} = P_{net_i} - jQ_{net_i}, \tag{2.2}$$

where the net power injection at each node i is $P_{net_i} = P_{g_i} - P_{L_i}$ and $Q_{net_i} = Q_{g_i} - Q_{L_i}$, where $P_{g_i}(Q_{g_i})$ is the real (reactive) power generated and $P_{L_i}(Q_{L_i})$ is the real (reactive) power consumed by the load at bus i. Equivalently, for each bus i, there are four real variables, P_{net_i}, Q_{net_i}, V_i, and θ_i. The power flow equations express the relationship that these variables must satisfy when the power system operates in the steady state. We denote by $\boldsymbol{Y}$ the $(N+1) \times (N+1)$ nodal admittance matrix, with Y_{ij} as the element in row $i+1$ and column $j+1$. We adopt the convention that $\boldsymbol{Y} = \boldsymbol{G} - j\boldsymbol{B}$, where $\boldsymbol{G}$ is the conductance matrix and $\boldsymbol{B}$ is the susceptance matrix. Then we have

$$\boldsymbol{I} = \boldsymbol{Y}\boldsymbol{E}, \tag{2.3}$$

where $\boldsymbol{I} = [\boldsymbol{I}_0, \boldsymbol{I}_1, \cdots, \boldsymbol{I}_N]^T$ is the vector of nodal current injection phasors, and $\boldsymbol{E} = [\boldsymbol{E}_0, \boldsymbol{E}_1, \cdots, \boldsymbol{E}_N]^T$ is the vector of nodal voltage phasors measured with respect to the ground node.

In power systems, we have three types of buses: (1) *slack bus* 0 with V_0 and θ_0 specified; (2) *P,V-bus* with P_{net_i} and V_i specified; and (3) *P,Q-bus* with P_{net_i} and Q_{net_i} specified. At each bus two of the four variables are known and the other two are unknown. At each bus i, the net complex power is given by

$$S_{net_i} = P_{net_i} - jQ_{net_i} = E_i^* \, I_i = E_i^* \sum_{k=0}^{N} Y_{ik} \, E_k. \tag{2.4}$$

Therefore the power balance equations at each bus can be formulated as follows by separating the real and imaginary part,

$$P_{net_i} = \sum_{k=0}^{N} V_i V_k [G_{ik} \cos\theta_{ik} - B_{ik} \sin\theta_{ik}], \tag{2.5}$$

$$Q_{net_i} = \sum_{k=0}^{N} V_i V_k [G_{ik} \sin\theta_{ik} + B_{ik} \cos\theta_{ik}], \tag{2.6}$$

where $i \in \mathcal{N}$, and $\theta_{ik} = \theta_i - \theta_k$ is the voltage angle difference between bus i and k. The complex power flow in the transmission line $\ell = (i, j)$ can be formulated as

$$S_{ij} = E_i^* I_{ij}. \tag{2.7}$$

The goal of power flow analysis is to solve the above nonlinear equations and obtain the voltage phasors and power flow in branches that represent the state of the system.

2.2.2 Optimal Power Flow

OPF is used to optimize the steady-state performance of a power system in terms of an objective function under certain equality and inequality constraints. With specified reference bus angle, line admittance, shunt capacitances, and P_{net_i} and Q_{net_i} at *P,Q-bus*, the ACOPF problem can be formulated as follows:

$$\begin{aligned} \min_{u} \quad & f(\boldsymbol{x}, \boldsymbol{u}) \\ s.t. \quad & g(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0} \\ & h(\boldsymbol{x}, \boldsymbol{u}) \leq \mathbf{0} \end{aligned} \quad , \tag{2.8}$$

where $\boldsymbol{u}$ is the vector of independent (or control) variables and $\boldsymbol{x}$ is the vector of dependent (or state) variables. Here,

$$\boldsymbol{u} = [P_m, V_m, t_\ell\,], \quad \textit{for } \forall \textit{ P,V-bus } m, \tag{2.9}$$

$$\boldsymbol{x} = [V_r, \theta_r, \theta_m\,], \quad \textit{for } \forall \textit{ P,V-bus } m \text{ and } \forall \textit{ P,Q-bus } r, \tag{2.10}$$

where t_ℓ is the vector of transformer tap settings. The equality constraints $g(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0}$ consist of nonlinear power balance equations in (2.5) and (2.6). The inequality constraints $h(\boldsymbol{x}, \boldsymbol{u}) \leq \mathbf{0}$ typically include

$$V_i^{min} \leq V_i \leq V_i^{max}, \tag{2.11a}$$

$$P_{g_i}^{min} \leq P_{g_i} \leq P_{g_i}^{max}, \tag{2.11b}$$

$$Q_{g_i}^{min} \leq Q_{g_i} \leq Q_{g_i}^{max}, \tag{2.11c}$$

$$S_{\ell_k} \leq S_{\ell_k}^{max}, \tag{2.11d}$$

$$t_{\ell_k}^{min} \leq t_{\ell_k} \leq t_{\ell_k}^{max}, \tag{2.11e}$$

for $\forall i \in \mathcal{N}$ and $\forall \ell_k \in \mathcal{L}$. Here, P_{g_i} and Q_{g_i} are the active power generation and reactive power generation of the generator at bus i. S_{ℓ_k} and t_{ℓ_k} are the power flow and the transformer tap setting on ℓ_k.

2.2.3 Applications

OPF is an efficient tool in power system operations and it has many applications. Below are two popular applications.

2.2.3.1 Minimization of Generation Cost

In the case of minimizing the generation cost, the objective function f is usually considered as the total active power generation cost:

$$f = \sum i \in \mathcal{N}_{\mathcal{G}} f^i(P_{g_i}), \tag{2.12a}$$

where $\mathcal{N}_{\mathcal{G}}$ ={i | bus i is connected to a generator}, and $f^i(P_{g_i})$ is the active power generation cost at bus i. $f^i(P_{g_i})$ is usually modeled by a quadratic function,

$$f^i(P_{g_i}) = a_i P_{g_i}^2 + b_i P_{g_i} + c_i, \tag{2.12b}$$

where a_i, b_i, c_i are the cost coefficients. If this problem can be solved accurately in real time, optimal control operations will be updated timely to achieve the lowest generation cost, and potentially a large amount of money can be saved.

2.2.3.2 Minimization of Line Loss

In this case, the objective function f is considered as the total loss on transmission lines [49]:

$$f = \sum_{(i,j)\in\mathcal{L}} G_{ij}(V_i^2 + V_j^2 - 2V_i V_j \cos(\theta_i - \theta_j)). \tag{2.13}$$

Solving this ACOPF problem in real time will enable timely adjustment of control settings to reduce the line loss, which can improve the economic efficiency of power system operation. In 2011, around 7% of the electricity generated was lost in the transmission lines in the U.S., which is worth about $3.23 billion. As a result, it is important to solve this ACOPF problem quickly and accurately.

2.3 Methods

In this chapter, we present the NR-based algorithm for solving the ACOPF for the large-scale smart grid. We will take the objective function of minimizing the active power generation cost, which is shown in (3.10), as an example to illustrate this method.

2.3.1 Approach Overview

The goal of this approach is to accelerate the computation of the ACOPF solution by reducing the number of variables in the ACOPF, thus reducing the size of the ACOPF problem. With NR, the size of $\boldsymbol{x}, \boldsymbol{u}$ and the admittance matrix $\boldsymbol{Y}$, are reduced down to the size of $\boldsymbol{x}^{eq}, \boldsymbol{u}^{eq}$, and $\boldsymbol{Y}^{eq}$ in the newly formulated ACOPF problem for the reduced equivalent system. We denote by $\mathcal{N}^{eq} \triangleq \{0, 1, 2, ..., N^{eq}\}$ the set of buses in the reduced system and by $\mathcal{L}^{eq} \triangleq \{\ell_1, \ell_2, \ell_3, ..., \ell_{L^{eq}}\}$ the set of transmission lines that connect the buses in set $\mathcal{N}^{eq}$. Similarly, power balance equations and line flow equations are formulated as

$$\begin{aligned} S^{eq}_{net_i} &= P_{net_i} - jQ_{net_i} = E^{eq*}_i \, I^{eq}_i \\ &= E^{eq*}_i \sum_{j=0}^{N^{eq}} Y^{eq}_{ij} E_j{}^{eq}, \forall i \in \mathcal{N}^{eq}, \end{aligned} \tag{2.14a}$$

$$S^{eq}_{ij} = E^{eq*}_i I^{eq}_{ij}, \;\; \forall (i,j) = \ell \in \mathcal{L}^{eq}. \tag{2.14b}$$

In order to keep the equivalence between the reduced ACOPF problem and the original ACOPF problem, the power injection pattern and the power flow pattern should be maintained. The goal of NR is to find an aggregation function that maps the variables in the original system to the variables in the reduced system, and that minimizes the mismatch between the original system and the reduced equivalent system. However, it is impossible to use existing methods to analytically or numerically find the exact aggregation function in real time for the large-scale smart grid. In this approach, the aggregation function is analytically approximated by linearizing the AC power balance equations. Therefore, we propose the NR-based ACOPF computation algorithm.

The overall algorithm flow of the proposed method is shown in Fig. 2.1. We first generate the similarity descriptors and the congestion indicators. By clustering, we group buses into subsystems. NR is performed to generate an equivalent reduced system. Then, the ACOPF problem is formulated and calculated for the equivalent reduced system. After checking the feasibility of the solution, detailed control settings are recovered by solving ACOPF for each subsystem.

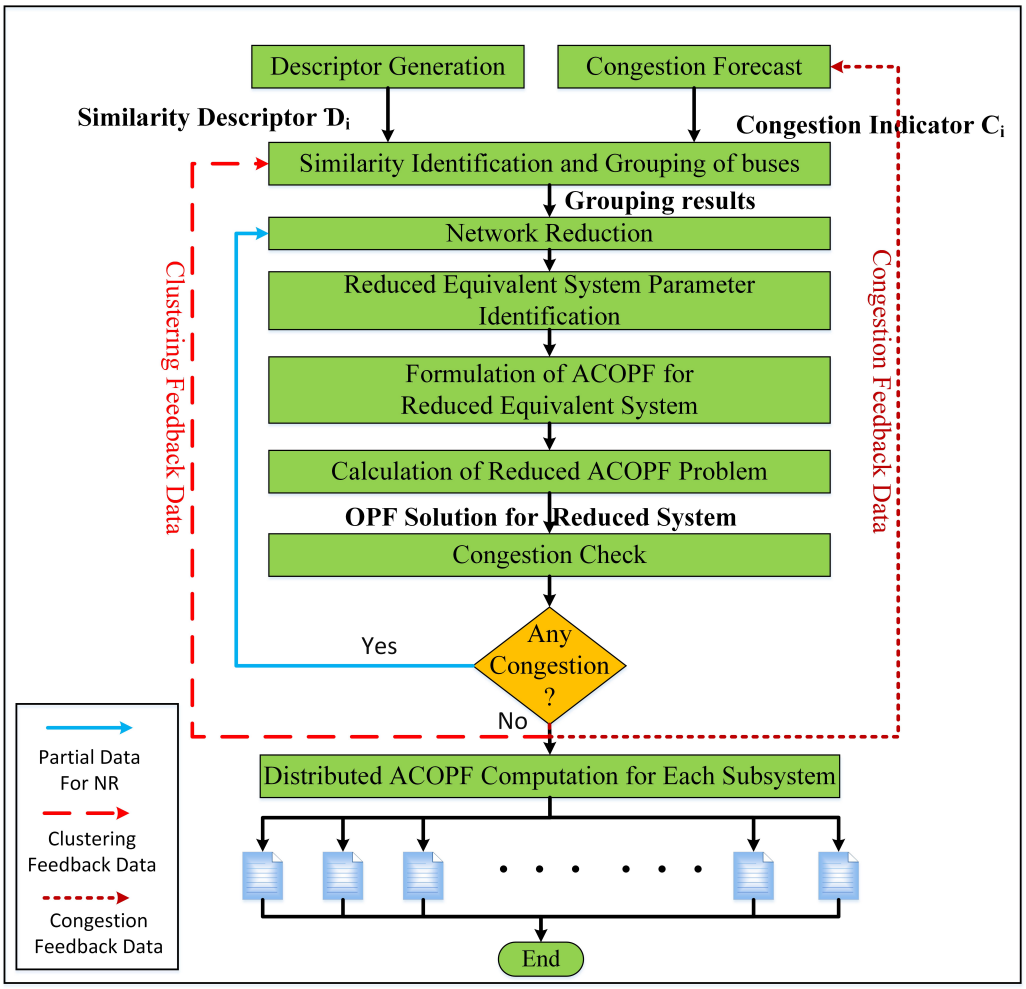


Figure 2.1: Fast ACOPF algorithm

2.3.2 Generation of Similarity Descriptor

In order to identify the similarity between buses and to group similar buses into one subsystem, a novel bus similarity descriptor containing voltage, load/generator model, and surrounding network topology information is introduced here. The traditional descriptor in power system applications only considers the bus itself and ignores the interactions with its adjacent buses in its local network. Such isolation of buses cannot fully reflect bus features. We assign each bus $i \in \mathcal{N}$ the similarity descriptor

$$\mathcal{D}_i = \tau_i \cdot (V_i, \theta_i, M_{G_i}, M_{L_i}, \Gamma_i), \tag{2.15}$$

where

$$\tau_i = (\tau_{V_i}, \tau_{\theta_i}, \tau_{M_{G_i}}, \tau_{M_{L_i}}, \tau_{\Gamma_i}) \tag{2.16}$$

is the weight vector; M_{G_i} and M_{L_i} represent the generator model and load model at bus i; Γ_i is defined by

$$\Gamma_i = (\sum_{j,(i,j)\in\mathcal{L}} \frac{V_i}{V_j \cos\theta_{ij}}, \sum_{j,(i,j)\in\mathcal{L}} \frac{V_i}{V_j \sin\theta_{ij}}), \tag{2.17}$$

which is the local topology descriptor. In this approach, we use the polynomial "ZIP" load model [50] to describe the load model. The generators are modeled as synchronous generators and inductor generators [51, 52]. All the parameters that describe the generator's and load's features are included in the descriptor. The similarity descriptor provides a measure of how "close" two buses are. It allows us to identify buses that can be merged together in the reduced network by using clustering algorithms. The generation of similarity descriptors is done offline and will be updated when required.

2.3.3 Congestion Forecast

In a power system, congestion occurs whenever the provision of transmission services required by the preferred generation/demand schedule exceeds the physical capability of the grid. In this dissertation, we only consider the restrictions imposed by the physical transmission capacity of the line. Congestion may increase the total generation cost because it may prevent cheap electricity generation from being dispatched. If we neglect congestion in the original system, it is possible that the calculated ACOPF solution after NR is not feasible. Therefore, it is very important to properly group buses in order to preserve the congestion profile. We propose a new method to ensure that the line flow constraints are not violated in the original system when applying the control settings derived from the ACOPF solution of the reduced network.

Congestion forecast is a heuristic method that predicts where congestion is going to occur. It takes power system field measurement data and the load forecast result $S_{L_i}^f$ as inputs. In addition, it takes the uncommitted transfer capability (UTC) of the successfully calculated ACOPF cases as feedback.

UTC in line $\ell = (i, j)$ is defined by

$$u_{(i,j)} \triangleq S^{max}_{(i,j)} - |S_{(i,j)}| \,, \quad \text{where } \ \ell = (i, j) \in \mathcal{L}. \tag{2.18}$$

In order to predict congestion, we assign a congestion indicator for bus $i \in \mathcal{N}$ based on the following heuristics:

(1) If the power generation capacity at bus i plus the total UTC of the transmission lines connected to bus i are larger than its demand, then bus i can either accommodate itself or import power from other generators. The reverse is also true.

(2) Motivated by OPF, power systems will force cheap generators to generate as much power as they can and export it to reduce the overall cost until some factors, such as loss or congestion on the lines, limit the benefit of increasing generation output. We use the derivative of the cost function with respect to the power generation $\lambda^f_{P_{g_i}} = \frac{\partial f}{\partial P_{g_i}}$ at current operational state to evaluate it.

(3) Based on different loading and generating conditions, the system will update the control settings to the new optimal control settings by solving the ACOPF problem. Depending on the system condition, UTC changes correspondingly. We denote u_ℓ as the original UTC and $\tilde{u}_\ell$ as UTC after applying new optimal control settings. By comparing these two UTCs, we find the lines that became congested and accordingly predict which lines are going to get congested.

Based on the above heuristics, we define

$$\phi^c_i = 1 - e^{\gamma_{\phi_i}(\sum_{(i,j)\in\mathcal{L}} u_{(i,j)} + S^{max}_{g_i} - S^f_{L_i})} \,, \tag{2.19}$$

$$\alpha^c_i = (\frac{\lambda^f_{P_{g_i}}}{\max_i \lambda^f_{P_{g_i}}})^{\gamma_{\alpha_i}}, \tag{2.20}$$

$$\beta^c_i = \min_{(i,j)\in\mathcal{L}} \beta_{ij} \,, \tag{2.21a}$$

$$\beta_{ij} = \begin{cases} 1, & u_{(i,j)} < \tilde{u}_{(i,j)} \\ (\frac{\tilde{u}_{(i,j)}}{u_{(i,j)}})^{\gamma_{\beta_i}}, & otherwise \end{cases} , \tag{2.21b}$$

where ϕ_i^c indicates the impact of supply and demand balance on congestion, and α_i^c reflects the impact of power generation cost on congestion. β_i^c indicates the possibility of lines connected to bus i getting congested after applying the optimal control settings. $\boldsymbol{\gamma_i} = (\gamma_{\phi_i}, \gamma_{\alpha_i}, \gamma_{\beta_i})$ is the weight vector. Note that $\gamma_{\phi_i}, \gamma_{\alpha_i} > 0$ and $0 < \gamma_{\beta_i} < 1$. We define the congestion indicator C_i:

$$C_i = \phi_i^c * \alpha_i^c * \beta_i^c \,. \tag{2.22}$$

It is obvious that $\phi_i^c, \alpha_i^c, \beta_i^c \in [0, 1]$, thus $C_i \in [0, 1]$. The congestion indicator C_i for bus i is assigned to be 1 when bus i is connected to lines that are susceptible to congestion. We tend to isolate bus i if C_i is close to 0 and group i into a subsystem if C_i is close to 1.

2.3.4 Similarity Identification and Grouping of Buses

In the similarity identification process, each point is represented by Ψ_i, which is defined by

$$\Psi_i = \begin{cases} 0, & C_i < \delta \\ \mathcal{D}_i, & C_i \geq \delta \end{cases}, \tag{2.23}$$

where δ is a threshold for congestion indicators. System operators select δ to meet their accuracy and performance requirements. A hot start K-means algorithm is used to cluster the buses [53]. The most recent historical clustering result is used as the starting point to improve the convergence speed of the K-means algorithm. Due to the special physical features of the slack bus and transformers, we isolate the slack bus and make sure that lines with transformers are not grouped into subsystems unless its tap ratio is close to 1. After the clustering process, the system is then divided into S subsystems. The set $\mathcal{N}$ is divided into S subsets, where $\mathcal{N}_k \in \mathcal{N}$ and $\mathcal{N}_k \cap \mathcal{N}_m = \emptyset$ for $\forall k, m \leq S$. Subsystem k contains all the buses in $\mathcal{N}_k$. Let $c_k \in \mathcal{N}_k$ denote the centroid bus in subsystem k. Value at bus c_k is the average value of the cluster.

2.3.5 Network Reduction and Reduced System Generation

The NR process follows the following strategy:

(a) Buses inside one subsystem are aggregated into one bus;

(b) Lines between two subsystems are aggregated into one line; and

(c) Lines inside one subsystem are ignored.

The power network parameters are approximated to maintain the same power injection pattern and power flow pattern as the original system. We propose a fast method to approximate the aggregation function.

2.3.5.1 Power Demand and Generation in the Equivalent System

Power demand $S_{L_k}^{eq}$ and power generation $S_{g_k}^{eq}$ at bus k in the equivalent system are calculated as follows:

$$S_{L_k}^{eq} = \sum_{i \in \mathcal{N}_k} S_{L_i} \quad \text{and} \quad S_{g_k}^{eq} = \sum_{i \in \mathcal{N}_k} S_{g_i} \,, \tag{2.24}$$

where S_{Li} and S_{g_i} are the power demand and power generation at bus i in the original system.

2.3.5.2 Bus Voltage in the Equivalent System

Since subsystem k is aggregated into bus k in the equivalent system, E_k^{eq} is approximated by the voltage of the centroid bus c_k in subsystem k.

$$E_k^{eq} = E_{c_k}. \tag{2.25}$$

2.3.5.3 Equivalent Line Admittance Approximation

Traditionally, many approaches identify the parameters of the equivalent system by calculating the sensitivity matrix. However, that kind of approach is computationally expensive, especially for large-scale power systems, as the calculation of the sensitivity matrix may take from minutes to hours to complete. In this approach, we approximate the parameter by the linearized

power balance equations. We merge the power balance equations (2.5) and (2.6) of the buses in one subsystem and generate power balance equations for the single equivalent bus. Similarly, we merge the line flow equations from (2.7) where line flows in the equivalent system are

$$S_{mn} = \sum_{r \in \mathcal{N}_m, v \in \mathcal{N}_n} S_{rv}. \tag{2.26}$$

In order to maintain the same power injection pattern and the same power flow pattern, by using Taylor's expansion, the line admittance matrix is approximated by

$$Y_{ij}^{eq} = \sum_{s \in \mathcal{N}_i, t \in \mathcal{N}_j} \left(\frac{V_s V_t \cos\theta_{st}}{V_{c_i} V_{c_j} \cos\theta_{c_i c_j}} G_{st} + j \frac{V_s V_t \sin\theta_{st}}{V_{c_i} V_{c_j} \sin\theta_{c_i c_j}} B_{st}\right). \tag{2.27}$$

2.3.6 Formulation and Calculation of the Reduced ACOPF Problem

In order to perform ACOPF computation, a new objective function and a set of new constraints after NR are generated based on (2.24)-(2.27). Several generators are aggregated into a single bus in the equivalent system. The cost function of an equivalent generator is greedily changed to a piecewise function:

$$\begin{aligned} f^k(P_{g_k}^{eq}) &= \min_{i \in \mathcal{N}_k} \sum_i f^i(P_{g_i}) \\ s.t.\ \ P_{g_k}^{eq} &= \sum_{i \in \mathcal{N}_k} P_{g_i}, \ \ \forall k \in \mathcal{N}^{eq}. \end{aligned} \tag{2.28}$$

The equality constraints, which are the power balance equations, are changed to

$$S_i^{eq} = P_i^{eq} - jQ_i^{eq} = E_i^{eq*} \sum_{k=0}^{N^{eq}} E_i^{eq} Y_{ik}^{eq}. \tag{2.29}$$

Inequality constraints are relaxed based on (2.24)-(2.27). For constraints on state variables and power generation limits, the minima of the lower bounds are used as the new lower bounds, and the maxima of the upper bounds are used as the new upper bounds. Line limits in the equivalent system are relaxed to the sum of the corresponding line limits in the original system. We find the optimal solution for the reduced equivalent system

by performing ACOPF analysis. Further computation is needed to find the optimal solution to the original system.

2.3.7 Congestion Check

Based on the ACOPF solution for the reduced system, the interchanged power between different subsystems is obtained. The feasibility of the ACOPF solution for the reduced system is then efficiently checked in parallel by performing power flow analysis for each subsystem while considering the original constraints. In order to consider power interchange, we add an additional equality constraint that models the power interchange activities between different subsystems to the constrained power flow analysis problem for each subsystem. If there is no solution for the constrained power flow analysis in a subsystem, it indicates that there exists a congested line in that subsystem.

As shown in Fig. 2.1, if there are congested lines detected, we isolate the related buses, remove the congested lines out of the subsystem and go back to the NR step.

2.3.8 Distributed ACOPF Computation for Each Subsystem

The ACOPF solution for the reduced system gives the sum of the control variables inside each subsystem. To decide the optimal dispatch inside the subsystem, it is still an ACOPF problem but with a smaller size and interchange power specified. Thus, ACOPF is computed to find the optimal settings for each subsystem. Finally, we obtain the detailed solution to the original ACOPF problem. With the nature of such a coarse-grained framework, we are able to distribute the computation of ACOPF for S subsystems to S processors to improve the speed.

We use the primal-dual interior point method to solve the ACOPF problem. It is worth mentioning that this framework works with different solvers and is in parallel with the performance of the optimization problem solvers.

2.4 Results

To test the proposed fast ACOPF computation algorithm for large-scale smart grids, we use two standard IEEE test power systems and two modified large test systems published in Matpower [54]. They are summarized in Table 2.1. In our tests, we use the total active power generation cost as the objective function of the ACOPF problem. We run all the tests on a laptop, which has an Intel Core2 Duo Processor of 2.26 GHz and 2 GB memory.

Table 2.1: Test Benchmarks

Benchmark	Bus No.	Branch No.	Generator No.
IEEE 30-bus	30	41	6
IEEE 300-bus	300	411	69
Case 3120sp	3,120	3,693	505
Case 21k	21,084	25,001	2,692

2.4.1 Experimental Results for the IEEE 30-Bus Test System

The IEEE 30-bus standard load-flow test system is used as a benchmark here. Figure 2.2 shows the network of the IEEE 30-bus system and it is partitioned into 6 subsystems. Two modified IEEE 30-bus systems with 5% more load demand and 10% more load demand and a modified IEEE 30-bus system with congestion are used to demonstrate the robustness of the proposed method including congestion forecast and congestion check. Real power costs for these 30-bus test systems were adapted from [16].

In this experiment, the congestion indicator threshold is set as 10% of the max value of all the congestion indicators C_is. For the non-congested test system, no additional bus is isolated. As shown in Fig. 2.2, the dash lines are the boundaries of the subsystems. To illustrate the capability of the congestion forecast module, we set the transmission capacity in line 2-5 to be 32 MW while the active power flow in this line was 63.01 MW in the standard case. Then bus 5 is isolated and the solid line shows the isolation.

Table 2.2 shows the experimental results for the IEEE 30-bus test system. Initially, we set the power generation to be {260.9, 40.0, 0.0, 0.0, 0.0, 0.0} (MW), which is a feasible setting for the test system, and the generation cost is 875.28 $/hr. By using the proposed method, the optimal setting is

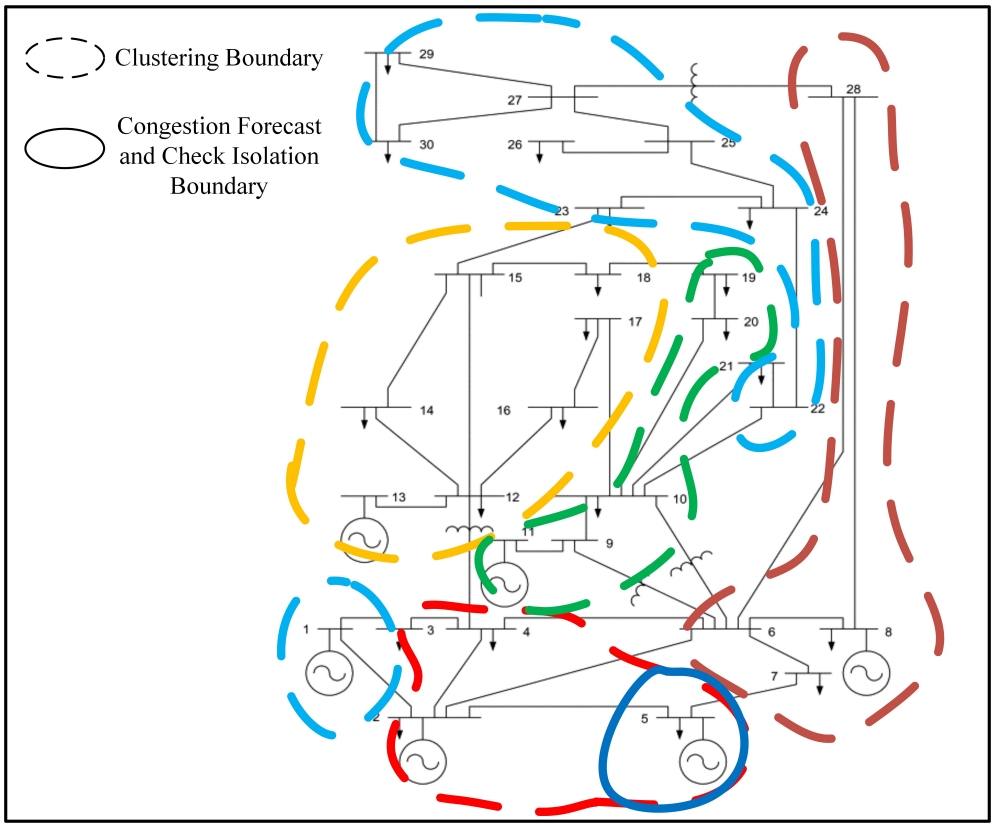


Figure 2.2: Clustering results for the IEEE 30-bus test system

{178.91, 48.50, 21.18, 21.14, 11.93, 11.40} (MW), with the generation cost of 802.35 $/hr. Note that bus 1 is the slack bus and we do not control the active power output. The total generation cost is reduced by 8.31%. The proposed method has 0.016% error on average compared to the most accurate full ACOPF. The proposed method reduces the error by 97.93% on average compared to DCOPF. The error of DCOPF is about 43 times larger than the proposed method in congestion-free test systems. In the congestion case DCOPF has a larger error which is about 2.1% and the error of DCOPF is about 106 times larger than the proposed method. The proposed method handles congested systems much better than DCOPF does.

2.4.2 Experimental Results for Larger Benchmarks

We also test this algorithm on larger benchmarks, including the IEEE 300-bus test system, case 3120sp, and case 21k from Matpower. In this approach, a 300-bus system is reduced to an 89-bus system with 112 lines; the 3120-bus

Table 2.2: Experimental Results of the IEEE 30-Bus Test System

	Standard 30-bus	With 5% DI	With 10% DI	With Congestion
Initial ($/h)	875.28	940.40	1008.05	875.28
ACOPF ($/h)	802.20	854.41	907.59	947.44
DCOPF ($/h)	806.97	859.70	913.44	967.67
PM ($/h)	802.35	854.54	907.69	947.63
DC error* ($/h)	4.77	5.26	5.75	20.23
DC error*	0.595%	0.612%	0.634%	2.135%
PM error* ($/h)	0.15	0.13	0.10	0.19
PM error*	0.018%	0.015%	0.011%	0.020%
Improvement**	96.86%	97.53%	98.26%	99.06%

PM: Proposed Method; DC error: DCOPF error; DI: Demand Increase.

*PM/DCOPF Error Compared to ACOPF

**PM Accuracy Improvement Compared to DCOPF

Table 2.3: Experimental Results of Accuracy Evaluation

	IEEE 30-bus	IEEE 300-bus	Case 3120sp	Case 21k
ACOPF ($/h)	802.20	719,725	2,142,704	2,732,880
DCOPF ($/h)	806.97	724,171	2,165,940	2,925,892
PM ($/h)	802.35	721,967	2,145,385	2,779,782
DC error ($/h)	4.77	4,446	23,236	193,012
DC error *	0.595%	0.618%	1.084%	7.06%
PM error ($/h)	0.15	2,242	2,681	46,902
PM error *	0.019%	0.311%	0.125%	1.72%
Improvement **	96.8%	49.6%	88.5%	75.70%

PM: Proposed Method; DC error: DCOPF error.

*PM/DCOPF Error Compared to ACOPF

**PM Accuracy Improvement Compared to DCOPF

Table 2.4: Experimental Results of Computation Time

	IEEE 30-bus	IEEE 300-bus	Case 3120sp	Case 21k
ACOPF (s)	0.6510	1.312	15.250	2552.8
DCOPF(s)	0.4720	0.5109	5.6130	400.7
PM (s)	0.4946	0.7966	7.2547	364.0
PM Speedup Compared to ACOPF	1.32×	1.63×	2.12×	7.01×
PM Speedup Compared to DCOPF	0.85×	0.64×	0.77×	1.10×

PM: Proposed Method.

system is reduced to a 449-bus system with 565 lines; and the 21k-bus system is reduced to a 4628-bus system with 5824 lines.

Table 2.3 shows the accuracy of proposed method. The proposed method has 0.54% error on average (1.72% error for the 21k-bus system) compared to the most accurate full ACOPF. The proposed method reduces the error by 77.6% on average (75.7% for the 21k-bus system) compared to DCOPF. As the size of the power system increases, the error of obtaining the optimal generation cost also increases. For the 21k-bus system, 7.06% error was observed in DCOPF. The power system will unnecessarily lose $193,012 per hour, which is $1.69 billion per year. This method can provide an accurate solution to ACOPF problems that can reduce the error by 75.7% compared to DCOPF. Thus we can save $146,110 per hour, which is $1.28 billion per year.

Table 2.4 shows the computation time of the proposed method. Compared to full ACOPF, the proposed method achieves 1.32×-2.12× speedup for small benchmarks (30-bus, 300-bus, 3120-bus) and 7.01× speedup for the largest benchmark (the 21,000-bus test system). The proposed method is slower than DCOPF for small benchmarks, but is faster than DCOPF for the largest benchmark.

The proposed method achieves better accuracy for all test systems compared to DCOPF. For large systems, the proposed method has the advantage over DCOPF in terms of both accuracy and speed.

2.5 Conclusion

ACOPF is very important in power system operation. In some applications, it cannot be approximated by DCOPF because of the DC power flow assumption. In addition, the poor accuracy of DCOPF results in great loss of social welfare. Therefore, a faster ACOPF algorithm needs to be developed for large-scale smart grids.

In this chapter, we propose a fast ACOPF analysis framework through power system network reduction to speed up the computation of ACOPF problems. This distributed framework works with different ACOPF solvers, such as primal-dual interior point method. We demonstrate that this approach can achieve 1.32× to 7.01× speedup over full ACOPF while introducing just 0.54% error on average. With congestion forecast and check, as long as ACOPF can converge to the optimal solution, our proposed method can find an optimal solution, which demonstrates its robustness. Compared to the widely used DCOPF, we reduce the error by 77.6% on average. It can potentially save millions of dollars in smart grid operation. Also, experimental results show that the computation time of the proposed algorithm grows almost linearly. The proposed method can be used to solve ACOPF for large-scale power systems in many applications, such as operational reliability analysis and power market management.

CHAPTER 3

CLUSRED: CLUSTERING AND NETWORK REDUCTION BASED PROBABILISTIC OPTIMAL POWER FLOW ANALYSIS FOR LARGE-SCALE SMART GRID

3.1 Introduction

The future smart grid in the U.S. has caused dramatic increases in the use of renewable energy sources, energy storage and demand response. Contrary to the traditional electric components, all the smart devices and economic agents present in the smart grid will need to make real-time decisions in order to achieve their individual objectives or to maximize their own profits [14]. Due to the inherent randomness of natural phenomena and the implicit and inaccurate assumptions related to modeling approaches, such as the assumption of constant load, balanced and steady-state conditions, many power system analysis and control problems are subject to uncertainties. The reliability of the power system and its ability to realize its objectives depend on the capabilities of making decisions. These decision-makers must be able to handle problems of growing complexity accounting for uncertain conditions, such as random disturbances, the effect of non-dispatchable resources, and customer participation. Because of the inefficiency and inaccuracy of previous approaches, new computational methods are needed to provide real-time and accurate solution to these complicated decision making problems.

In the last 50 years, OPF, which seeks to optimize an objective function by adjusting a set of control variables subject to certain physical, operational and policy constraints, has been widely used to support power system automation and planning. It has been generally addressed as a deterministic optimization problem. However, many uncertain conditions, such as the variation of loading conditions and the measuring and forecasting errors of the system parameters, exist in power system operation. Traditional determin-

istic OPF (DOPF) cannot analyze the probabilistic system behaviors. Thus, it is becoming increasingly important to incorporate probability into DOPF and transform DOPF into a probabilistic OPF (POPF) problem for smarter system operation.

POPF is concerned with the introduction of randomness or uncertainty into conventional DOPF problems. Since the early 1970s, probabilistic methods have been applied to power systems, such as the probabilistic power flow (PPF) problems [15]. Recently, a cumulant method (CM) was proposed to solve PPF using cumulants and Gram-Charlier expansion [55]. Later, point estimate method (PEM) was used to account for uncertainties in PPF problem [56]. *First-Order Second-Moment Method*, which uses a first-order Taylor series approximation to compute second-order statistical information, was proposed to solve the POPF problems [57]. However, the limitation of only analyzing the probabilistic features at a specific solution point largely limits its application. In [58], a two-point estimate method for POPF was proposed to gain the first three moments of the corresponding probability density functions (PDFs). However, 2PEM method does not perform well when the number of uncertain variables becomes large, which means that 2PEM is not scalable and not desirable in modern CPS. In [18, 19], CM was used to solve POPF. The key point in CM is to find the linear relationship between unknown variables and known variables to compute cumulants. Ref. [18] just simply uses the inverse of Hessian matrix as the linear mapping matrix. The effectiveness and accuracy rely on the accuracy of inversing Hessian matrix. However, Hessian matrix can sometimes be ill-conditioned and cannot be inversed accurately, which will result in large errors. Also, no alternative solutions are provided when Hessian is not invertible. In addition, inversing the Hessian matrix is very time consuming for a large-scale power system, such as the eastern interconnection in U.S.

Considering these limitations, we develop a new method to generate the linear mapping matrix efficiently and accurately. To solve the scalability issue, network reduction (NR) techniques are used to solve the problem. NR can efficiently reduce the computation burden by analyzing an equivalent reduced system. Traditional methods, such as Ward equivalent technique [13], are usually performed by computing the admittance and eliminating unnecessary elements that are not in the study area, which can reduce the network but cannot be used to solve OPF problems because it may not yield the same

power flow pattern as the original one. Sensitivity matrix based methods, such as power transfer distribution factor (PTDF) based method [16], are too time consuming and cannot be used for real-time operating purpose [17]. In our previous work [43], an analytical approximation based NR method was developed to solve DOPF efficiently and accurately. However, our previous approach cannot handle POPF. To solve this problem, we need to develop new probabilistic methods to handle clustering, aggregation, congestion check, equivalent POPF problem formulation, linear mapping matrix generation and so on.

In this chapter, we present a fast scalable algorithm, ClusRed, to solve POPF efficiently using clustering, NR techniques and a novel method to generate linear mapping matrix. Our main contributions are:

- This work provides a fast scalable probabilistic algorithm for POPF computation that is 4.57X faster than the previous approach.
- We developed a new linear mapping matrix generation method that is more robust when Hessian matrix is ill-conditioned.

The rest of this chapter is organized as follows. In Section 3.2, we give the necessary background, the DOPF and POPF formulation. Section 3.3 describes the framework and ClusRed algorithm. We present the numerical results in Section 3.4 and the conclusion in Section 3.5.

3.2 Preliminaries

Over the past 50 years, steady-state OPF problem was well formulated and many variations of OPF formulations were studied. In this section, we begin with the background of power system analysis. We will introduce power flow analysis, DOPF and POPF formulation.

3.2.1 Power Flow Analysis

We consider a power system with $N+1$ buses and L lines. We denote by $\mathcal{N} \triangleq \{0, 1, 2, \cdots, N\}$ the set of buses, with the bus 0 being the slack bus, and by $\mathcal{L} \triangleq \{\ell_1, \ell_2, \cdots \ell_L\}$ the set of transmission lines that connect the buses in

the set $\mathcal{N}$. We associate with each line $\ell \in \mathcal{L}$ the ordered pair of bus indices (i, j) and we write $\ell = (i, j)$. Each bus i is characterized by the voltage phasor $E_i = V_i e^{j\theta i}$ and the net injected complex power $S_{net_i} = P_{net_i} - jQ_{net_i}$. Here V_i is the nodal voltage magnitude and θ_i is the nodal voltage phase angle. The net power injection at each node i is $P_{net_i} = P_{g_i} - P_{L_i}$ and $Q_{net_i} = Q_{g_i} - Q_{L_i}$, where $P_{g_i}(Q_{g_i})$ is the real (reactive) power generated and $P_{L_i}(Q_{L_i})$ is the real (reactive) power consumed by the load at bus i. We denote by $\boldsymbol{Y}$ the $(N+1) \times (N+1)$ nodal admittance matrix, with Y_{ij} as the element in row $i+1$ and column $j+1$. We adopt the convention that $\boldsymbol{Y} = \boldsymbol{G} - j\boldsymbol{B}$, where $\boldsymbol{G}$ is the conductance matrix and $\boldsymbol{B}$ is the susceptance matrix. In power system, we have 3 types of buses: (1) *slack bus* 0 with V_0 and θ_0 specified; (2) *P,V-bus* with P_{net_i} and V_i specified; and (3) *P,Q-bus* with P_{net_i} and Q_{net_i} specified. A 3-bus power system example is shown in Fig. 3.1. In the 3-bus power system, bus 0 is a *slack bus* with V_0 and θ_0 specified; bus 1 is a *P,V-bus* with P_1 and V_1 specified; and bus 2 is a *P,Q-bus* with P_2 and Q_2 specified. The net complex power at bus i is given by

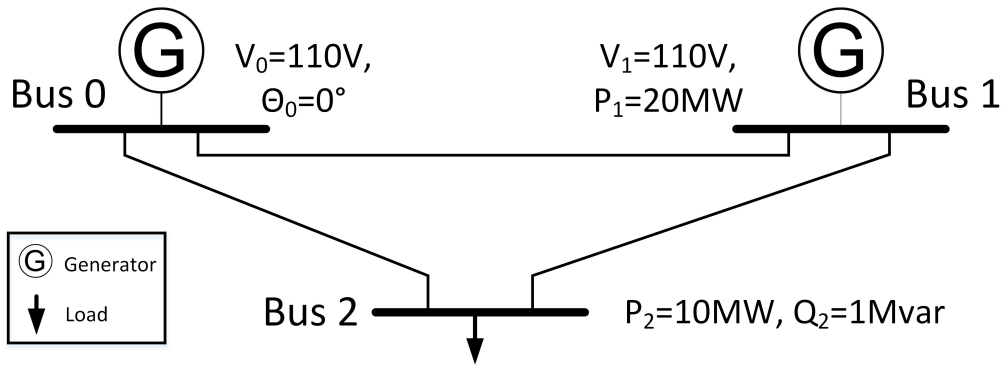


Figure 3.1: Single line diagram of 3-bus power system

$$S_{net_i} = P_{net_i} - jQ_{net_i} = E_i^* \, I_i = E_i^* \sum_{k=0}^{N} Y_{ik} \, E_k. \tag{3.1}$$

Therefore the power balance equations at each bus can be formulated as follows.

$$P_{net_i} = \sum_{k=0}^{N} V_i V_k [G_{ik} \cos\theta_{ik} - B_{ik} \sin\theta_{ik}], \tag{3.2}$$

$$Q_{net_i} = \sum_{k=0}^{N} V_i V_k [G_{ik} \sin\theta_{ik} + B_{ik} \cos\theta_{ik}], \tag{3.3}$$

where $i \in \mathcal{N}$, and $\theta_{ik} = \theta_i - \theta_k$ is the voltage angle difference between bus i and k. The complex power flow in the transmission line $\ell = (i, j)$ can be formulated as

$$S_{ij} = E_i^* I_{ij}. \tag{3.4}$$

The goal of power flow analysis is to solve the above nonlinear equations and obtain the states of the system.

3.2.2 Deterministic OPF

The OPF problem is formulated to optimize the steady state performance of a power system evaluated by an objective function, such as the total generation cost, or the total power transmission loss, under certain physical and operational constraints, including AC power flow, physical constraints and regulation policy constraints etc. With specified deterministic values of reference bus angle, power demands at each bus and network parameters, a general DOPF problem is formulated as follows:

$$\begin{aligned} \min_{u} \quad & f(\boldsymbol{x}, \boldsymbol{u}) \\ s.t. \quad & g(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0} \\ & h(\boldsymbol{x}, \boldsymbol{u}) \leq \mathbf{0} \end{aligned} \quad , \tag{3.5}$$

where $\boldsymbol{u}$ is the vector of independent (or control) variables and $\boldsymbol{x}$ is the vector of dependent (or state) variables. Here,

$$\boldsymbol{u} = [P_m, V_m, t_\ell\,], \quad for\ \forall\ P, V\text{-}bus\ m, \tag{3.6}$$

$$\boldsymbol{x} = [V_r, \theta_r, \theta_m\,], \quad for\ \forall\ P, V\text{-}bus\ m \text{ and } \forall\ P, Q\text{-}bus\ r, \tag{3.7}$$

where t_ℓ is vector of transformer tap settings. The equality constraints $g(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0}$ consist of nonlinear power balance equations in (3.2) and (3.3). The inequality constraints $h(\boldsymbol{x}, \boldsymbol{u}) \leq \mathbf{0}$ typically includes:

$$V_i^{min} \leq V_i \leq V_i^{max}, \tag{3.8a}$$

$$P_{g_i}^{min} \leq P_{g_i} \leq P_{g_i}^{max}, \tag{3.8b}$$

$$Q_{g_i}^{min} \leq Q_{g_i} \leq Q_{g_i}^{max}, \tag{3.8c}$$

$$S_{\ell_k} \leq S_{\ell_k}^{max}, \tag{3.8d}$$

$$t_{\ell_k}^{min} \leq t_{\ell_k} \leq t_{\ell_k}^{max}, \tag{3.8e}$$

for $\forall i \in \mathcal{N}$ and $\forall \ell_k \in \mathcal{L}$. Here, P_{g_i} and Q_{g_i} are the active power generation and reactive power generation of the generator at bus i. S_{ℓ_k} and t_{ℓ_k} are the power flow and the transformer tap setting on ℓ_k.

3.2.3 Probabilistic OPF

With some of the input variables to be uncertain, such as power generation and load, DOPF problem (3.5) becomes probabilistic. Thus the traditionally constant values will become random variables in POPF.

$$\begin{aligned} \min_{X} \quad & f(\boldsymbol{X}, \boldsymbol{S_g}, \boldsymbol{S_l}) \\ s.t. \quad & g(\boldsymbol{X}, \boldsymbol{S_g}, \boldsymbol{S_l}) = \boldsymbol{0} \\ & h(\boldsymbol{X}, \boldsymbol{S_g}) \leq \boldsymbol{0} \end{aligned} \tag{3.9}$$

where $X = [\boldsymbol{x}, \boldsymbol{u}]$ that contains all the control variables and state variables, S_g is the power generation vector (except the slack bus), and S_l is the load vector. Because of the randomness of S_l, the variables of interests are no longer deterministic. Instead, the desired outputs are PDFs. The goal of POPF is to calculate the PDFs of the variables of interests. For example, in power market analysis, locational marginal price (LMP) which evaluates the hypothetical production cost of one unit additional hypothetical power demand is very important. With random loading conditions, LMPs become probabilistic and the PDFs need to be calculated to make the most economic and efficient decisions on smart grids operation.

In reality, OPF can be used to minimize the total active power generation cost f:

$$f = \sum\nolimits_{i \in \mathcal{N}_\mathcal{G}} f^i(P_{g_i}), \tag{3.10a}$$

where $\mathcal{N}_\mathcal{G}$ =$\{$i $\mid$ bus i is connected to a generator$\}$, and $f^i(P_{g_i})$ is the active power generation cost at bus i. $f^i(P_{g_i})$ is usually modeled by a quadratic function,

$$f^i(P_{g_i}) = a_i P_{g_i}^2 + b_i P_{g_i} + c_i, \tag{3.10b}$$

where a_i, b_i, c_i are the cost coefficients. If this problem can be solved accurately in real time, optimal control operations will be updated timely to achieve the lowest generation cost with the largest probability and potentially a large amount of money can be saved.

3.2.4 Cumulants and Gram-Charlier Theory

Cumulants and Gram-Charlier (GC) theory are explained in [20]. For any random variable x, we denote by $\alpha_v(x)$ and $\gamma_v(x)$ the v^{th} moment and v^{th} cumulant of its distribution, m the mean, and σ the standard deviation. Suppose ξ and ξ' are random variables with $\xi = a\xi' + b$, we have

$$\gamma_1(\xi) = a\gamma_1(\xi') + b \quad \text{and} \quad \gamma_v(\xi) = a^v\gamma_v(\xi'). \tag{3.11}$$

Consider the sum of n independent random variables $\xi = \xi_1 + \xi_2 + ... + \xi_n$, the following property holds:

$$\gamma_v(\xi) = \gamma_v(\xi_1) + \gamma_v(\xi_2) + ... + \gamma_v(\xi_n). \tag{3.12}$$

We can compute unknown cumulants by using the known cumulants and the linear relationship between them.

For the standardized variable $\frac{x-m}{\sigma}$, its PDF is denoted as $f(x)$. According to GC Expansion, $f(x)$ can be written as

$$f(x) = \phi(x) + \frac{c_1}{1!}\phi(x)' + \frac{c_2}{2!}\phi(x)'' + \frac{c_3}{3!}\phi(x)''' + ... \ , \tag{3.13}$$

where $\phi(x)$ represents the PDF of normal distribution with $m = 0$ and $\sigma = 1$, and c_i are constant coefficients [20]:

$$c_i = \alpha_i - \sum_{m=1}^{i-1} \binom{i-1}{m-1} c_m \alpha_{i-m}. \tag{3.14}$$

3.3 Methods

In this section, we present the ClusRed algorithm that can solve POPF efficiently for large-scale smart grids. We take the objective function of minimizing total generation cost and known random loads with Gaussian distribution

as an example to illustrate this method. In this problem, the unknown variables that we are interested in are the LMPs. This method can be easily extended to solve any POPF problems with different objective functions, known random variables such as loads and wind generation, and unknown variables of interest such as LMP and power flow.

3.3.1 Approach Overview

The overall algorithm flow is shown in Fig. 3.2. We first generate similarity descriptors and congestion indicators for each bus. By clustering, we group buses into subsystems. NR is performed to generate an equivalent reduced system and an equivalent POPF problem is formulated. Then, we solve the equivalent POPF problem for the reduced system. After performing probabilistic congestion check, detailed solutions are obtained by solving POPF for each subsystem. At the solution point, we introduce a novel method to generate a linear mapping matrix that has better performance. Finally, by using cumulant method and Gram-Charlier expansion, we obtain the PDFs of the unknown variables of interest. Unlike NR-based method for DOPF, we consider PDFs of random loads during the clustering, NR and equivalent POPF formulation. In addition, we develop probabilistic congestion check method and compute POPF for each subsystem in a distributed way.

3.3.2 NR-based POPF Algorithm

3.3.2.1 Similarity Descriptor Generation

In order to identify the similarity between buses and group similar buses into one subsystem, we generate a similarity descriptor $\mathcal{D}_i$ that contains voltage, load/generator model, variance of random load and surrounding network topology information for bus i. By considering the variances of random loads, we can evaluate the randomness of the loads. Similarity descriptor provides a measure of how "close" two buses are. It allows us to aggregate similar buses in the reduced network by using clustering algorithms.

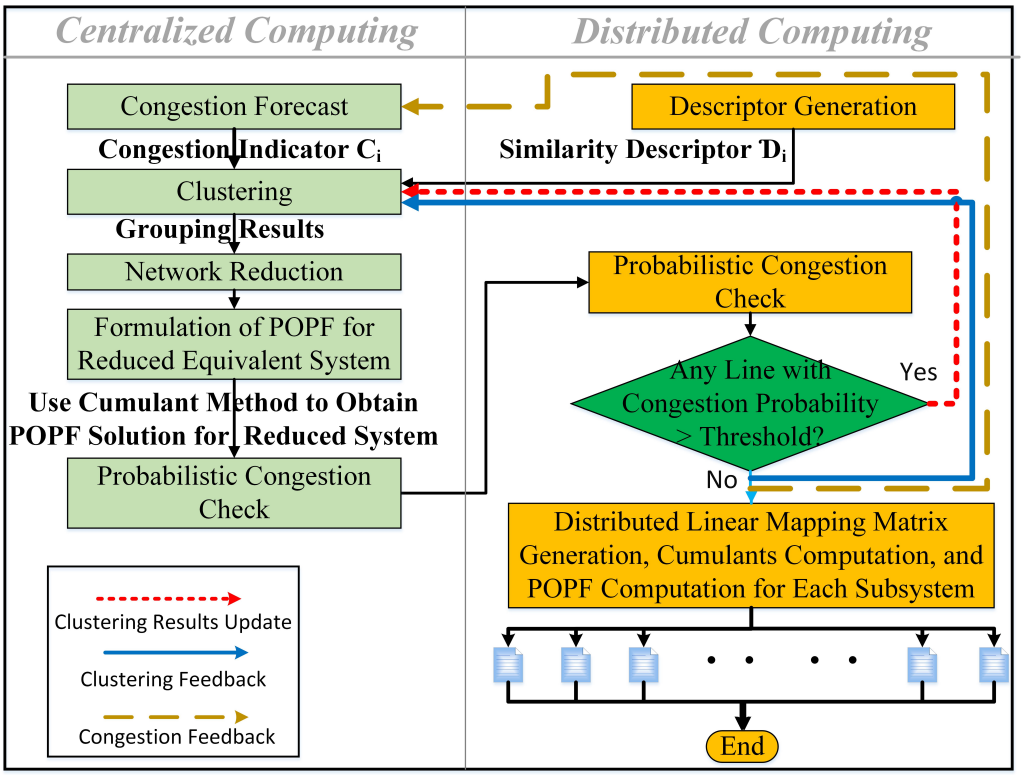


Figure 3.2: ClusRed algorithm

3.3.2.2 Clustering

We use the heuristic congestion forecast method in [43] to generate congestion indicator $C_i \in [0, 1]$. We tend to isolate bus i if C_i is close to 0 and group i into a subsystem if C_i is close to 1.

In the similarity identification process, each point is represented by Ψ_i, which is defined by

$$\Psi_i = \begin{cases} 0, & C_i < \delta \\ \mathcal{D}_i, & C_i \geq \delta \end{cases}, \tag{3.15}$$

where δ is a threshold for congestion indicator. System operator can select δ to meet their accuracy and performance requirements. Bus i will be isolated if $\Psi_i = 0$. Hot start K-means algorithm is used here. The most recent historical clustering result is used as the start point to improve the convergence speed of K-means algorithm. Due to the special physical features of slack bus and transformers, we isolate slack bus and make sure that lines with transformers are not grouped into subsystems. Then the system is divided into S subsystems, which means that set $\mathcal{N}$ is divided into S subsets, where

$\mathcal{N}_k \in \mathcal{N}$ and $\mathcal{N}_k \cap \mathcal{N}_m = \emptyset$ for $\forall k, m \leq S$. Subsystem k contains all the buses in $\mathcal{N}_k$. Let $c_k \in \mathcal{N}_k$ denote the centroid bus in subsystem k. Value at bus c_k is the average value of the cluster.

3.3.2.3 Network Reduction

The NR process follows the following strategies: (1) Buses inside each subsystem are aggregated into one bus; (2) Lines between two subsystems are aggregated into one line; (3) Lines inside each subsystem are ignored. The power network parameters are approximated to maintain the same power injection pattern and power flow pattern as the original system.

Power demand $S_{L_k}^{eq}$ and power generation $S_{g_k}^{eq}$ at bus k in the equivalent system are calculated by

$$S_{L_k}^{eq} = \sum\nolimits_{i \in \mathcal{N}_k} S_{L_i} \quad \text{and} \quad S_{g_k}^{eq} = \sum\nolimits_{i \in \mathcal{N}_k} S_{g_i} \,, \tag{3.16}$$

where S_{Li} and S_{g_i} are the power demand and power generation at bus i in the original system.

Suppose S_{L_i} has an independent normal distribution of $N(\mu_i, \sigma_i^2)$, where μ_i is the mean of the distribution and σ_i^2 is the variance of the distribution. Then $S_{L_k}^{eq}$ is also normally distributed, with mean $\sum_{i \in \mathcal{N}_k} \mu_i$ and variance $\sum_{i \in \mathcal{N}_k} \sigma_i^2$.

Since subsystem k is aggregated into bus k in the equivalent system, E_k^{eq} is approximated by the voltage of the centroid bus c_k in subsystem k.

$$E_k^{eq} = E_{c_k}. \tag{3.17}$$

Instead of using the time-consuming sensitivity matrix based method, we approximate the parameters by using the linearized power balance equations. We merge the power balance equations (3.2) and (3.3) of the buses in one subsystem and generate power balance equations for the single equivalent bus. Similarly, we merge the line flow equations from (3.4) where line flows in the equivalent system are

$$S_{mn} = \sum\nolimits_{r \in \mathcal{N}_m, v \in \mathcal{N}_n} S_{rv}. \tag{3.18}$$

In order to maintain the same power injection pattern and the same power flow pattern, the line admittance matrix is approximated by

$$Y_{ij}^{eq} = \sum_{s\in\mathcal{N}_i, t\in\mathcal{N}_j} \left(\frac{V_s V_t \cos\theta_{st}}{V_{c_i} V_{c_j} \cos\theta_{c_i c_j}} G_{st} + j\frac{V_s V_t \sin\theta_{st}}{V_{c_i} V_{c_j} \sin\theta_{c_i c_j}} B_{st}\right). \tag{3.19}$$

3.3.2.4 POPF for Reduced Equivalent System

After NR, new objective function, constraints and PDFs of aggregated loads are generated. Based on (3.16), the cost function of an equivalent generator is changed to a piecewise function by greedily minimizing the total cost without considering network constraints:

$$\begin{aligned} f^k(P_{g_k}^{eq}) = \min\nolimits_{i\in\mathcal{N}_k} \sum\nolimits_i f^i(P_{g_i}) \\ s.t.\ \ P_{g_k}^{eq} = \sum\nolimits_{i\in\mathcal{N}_k} P_{g_i},\ \ \forall k \in \mathcal{N}^{eq}. \end{aligned} \tag{3.20}$$

The equality constraints, which are the power balance equations, are changed to

$$S_i^{eq} = P_i^{eq} - jQ_i^{eq} = E_i^{eq*} \sum\nolimits_{k=0}^{N^{eq}} E_i^{eq} Y_{ik}^{eq}. \tag{3.21}$$

Inequality constraints are transformed based on (3.16) $\sim$ (3.19). Line limits in the equivalent system are relaxed to the sum of the corresponding line limits in the original system. After formulating the POPF problem for the reduced equivalent system, we obtain the optimal solution for the reduced system by performing POPF analysis. The line flows are then used to compute the power interchange between each subsystem and the external system.

3.3.2.5 Probabilistic Congestion Check

The random interchange power between each subsystem and the external system can be obtained by solving equivalent POPF problem. For each subsystem, we can obtain the PDFs of the line flows by performing probabilistic power flow analysis with random interchange power and random loads as random inputs while honoring all the constraints. We check the probability of violating line limits based on the PDFs of the line flows. If this probability is larger than a certain threshold, we will determine that this line is "prob-

ably" congested. As shown in Fig. 3.2, if there are "probably" congested lines detected, we isolate the related buses, remove the congested lines from the subsystem and update the clustering results. This step can be done in parallel.

3.3.2.6 Distributed POPF for Each Subsystem

The POPF solution for the reduced system gives the sum of the control variables inside each subsystem. To decide the optimal dispatch inside a subsystem, it is still an POPF problem but with a smaller size and interchange power specified. Thus, POPF is computed to find the optimal solution for each subsystem. Thus we obtain the detailed solution to the original POPF problem. Finally, by using cumulant method and Gram-Charlier expansion, we obtain the PDFs of the unknown variables. With the nature of such a multi-level framework, we are able to distribute the computation of ACOPF for S subsystems evenly to all the processors to improve the speed.

We summarize the overall algorithm in Algorithm 1.

Algorithm 1: POPF Algorithm

```
Data: Random input PDFs, Power system parameters
Result: PDFs of random variables
Calculate the cumulants of random inputs;
Calculate the similarity descriptors;
Calculate the congestion indicators;
while not Converged do
    Group buses into subsystems using k-means;
    Perform network reduction;
    Generate reduced equivalent system;
    Solve POPF for reduced equivalent system;
    Perform probabilistic congestion check;
    if congestion detected then
        Isolate congested lines and buses;
    else
        Solve POPF for subsystems; set Converged =1 ;
```

3.3.3 Linear Mapping Matrix

We introduce a new way to generate linear mapping matrix to calculate the PDFs of LMPs. Prime-dual Interior Point Method (PDIPM) is used to solve POPF in (3.9). We transform the POPF problem to the following:

$$\begin{aligned} \min_{x} \quad & [f(S_g) - \gamma \sum_{m=1}^{n_i} \ln(z_m)] \\ s.t. \quad & g(x, S_g, S_g) = 0 \\ & h(x, S_g) + z = 0 \\ & z \geq 0 \end{aligned} \quad . \tag{3.22}$$

For a certain γ, the Lagrangian for this equality constrained problem is

$$\begin{aligned} \mathcal{L}^{\gamma}(x, S_g, S_l, z, \lambda, \mu) =& f(S_g) + \lambda^T G(x, S_g, S_l) + \\ & \mu^T (H(x, S_g) + z) - \gamma \sum_{m=1}^{n_i} \ln(z_m). \end{aligned} \tag{3.23}$$

For any function $F(x, y)$, we denote by F_x the first order partial derivative of F with respect to (w.r.t.) x, F_{xx} the second order partial derivative w.r.t. x, and F_{xy} the second order mixed derivative w.r.t. x and y. The first order KKT condition gives:

$$(G_{xx}^T \lambda + H_{xx}^T \mu) * \Delta x + G_x^T * \Delta\lambda + H_x^T * \Delta\mu = -G_x^T \lambda - H_x^T \mu, \tag{3.24}$$

$$f_{S_g S_g}^T * \Delta S_g + G_{S_g}^T * \Delta\lambda + H_{S_g}^T * \Delta\mu = -f_{S_g}^T - G_{S_g}^T \lambda - H_{S_g}^T \mu, \tag{3.25}$$

$$G_{S_l}^T * \Delta S_l = -G_{S_l}^T \lambda, \tag{3.26}$$

$$[\mu] * \Delta Z + [Z] * \Delta\mu = -[\mu] Z + \gamma e, \tag{3.27}$$

$$G_x * \Delta x + G_{S_g} * \Delta S_g + G_{S_l} * \Delta S_l = -G(x, S_g, S_l), \tag{3.28}$$

$$H_x * \Delta x + H_{S_g} * \Delta S_g + I * \Delta Z = -H(x, S_g) - Z. \tag{3.29}$$

Solving these equations, we have

$$\Delta\lambda = T^{-1} G_{S_l} \Delta S_l + T^{-1}(G - G' * M^{-1} N), \tag{3.30}$$

where

$$G' = [G_x, G_{S_l}], \tag{3.31}$$

$$T = G' M^{-1} G'^T, \tag{3.32}$$

$$M = \begin{bmatrix} \mathcal{L}^{\gamma}_{xx} & 0 \\ 0 & f_{S_g S_g} \end{bmatrix} + \begin{bmatrix} H_x^T \\ H_{S_g}^T \end{bmatrix} [Z]^{-1} [\mu] [H_x, H_{S_g}], \tag{3.33}$$

$$N = \begin{bmatrix} \mathcal{L}^{\gamma}_{x} \\ \mathcal{L}^{\gamma}_{S_g} \end{bmatrix} + \begin{bmatrix} H_x^T \\ H_{S_g}^T \end{bmatrix} (\gamma e + [\mu][H_x, H_{S_g}]). \tag{3.34}$$

According to Taylor's expansion, in the neighbor of the solution point, this linear relationship holds for $\boldsymbol{\lambda}$ and $\boldsymbol{S_l}$.

$$\boldsymbol{\lambda} = T^{-1} G_{S_l} \boldsymbol{S_l} + T^{-1}(G - G' * M^{-1} N). \tag{3.35}$$

We use Cholesky decomposition to inverse the matrix to further speed up the computation. When M is not invertible, we use pseudo inverse instead and modify the matrix by adding $k * \boldsymbol{I}$ to ensure the robustness. After generating the cumulants of $\boldsymbol{S_l}$ using the PDFs, we use the linear mapping matrix to compute the cumulants of LMPs. With Gram-Charlier expansion, we can obtain the PDFs of the LMPs. By decomposing the large problem, we can operate on smaller problems with Hessian matrices that have smaller condition numbers. Instead of operating on a single high-dimensional Hessian matrix, we operate on smaller matrices and reduce the computation burden.

3.4 Results

To test ClusRed algorithm, we use five test benchmarks published in Matpower [54] which were also used in [18, 19]. They are summarized in Table 3.1. We run all the tests on a multi-core machine, which has Intel Core i5 processor of 3.10 GHz and 8 GB memory.

To demonstrate computation efficiency and accuracy of this method, we compare it with Monte Carlo simulation (MCS), which is considered "accurate". MCS repeats the process of DOPF using a particular set of values of the random variables that are sampled based on the corresponding PDFs. For the purpose of comparison we set the number of trials to be 10,000 to ensure the accuracy quality of the solution. However, because it takes months

Table 3.1: Test Benchmark

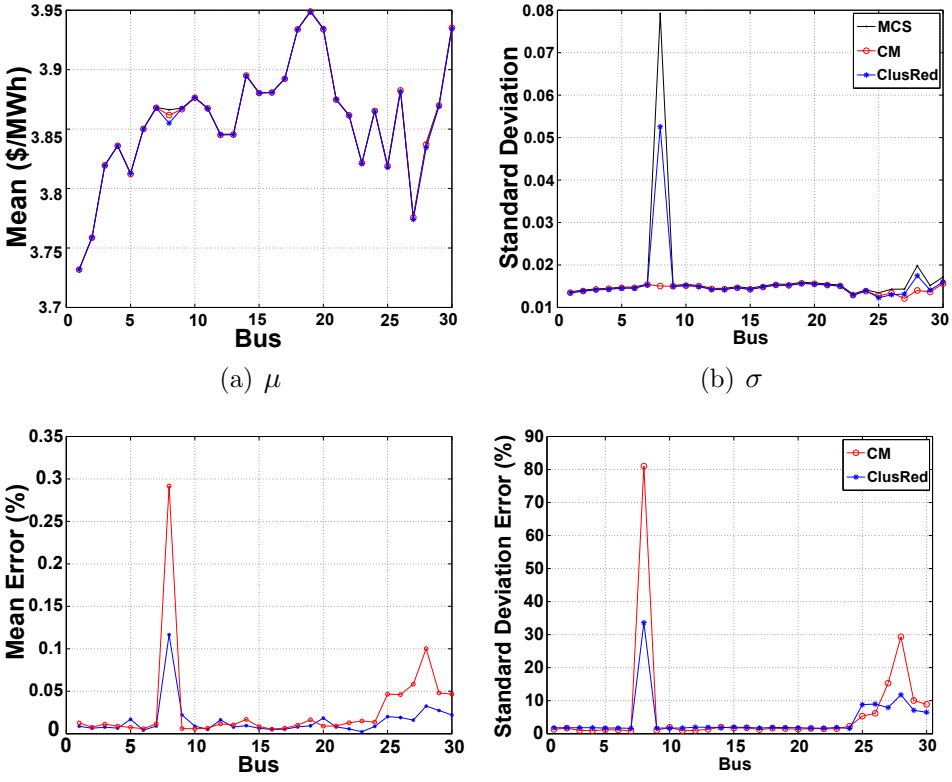
Benchmarks	Bus No.	Branch No.	Generator No.
Case 9	9	9	3
Case 30	30	41	6
Case 118	118	186	54
Case 3120	3,120	3,693	505
Case 21k	21,084	25,001	2,692

to run 10,000-sample MCS for the largest benchmark, case21k, we run 1,500-sample MCS. We also compare ClusRed with the previous method in [18], which we call CM.

3.4.1 Numerical Results on 30-bus System

Similar to the previous approach [18], we set the loads to vary with a small standard deviation that equals to 4% of the nominal bus loads. Figure 3.3 shows the mean and standard deviation of LMP at every bus obtained by MCS, ClusRed and CM.

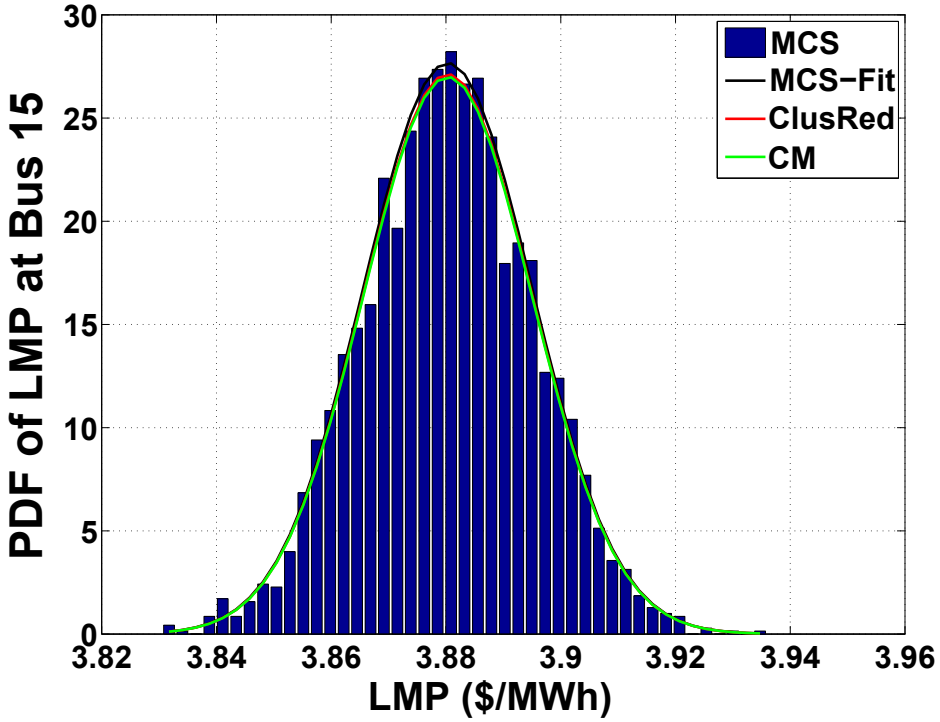
The PDF of LMP at bus 15 is shown in Fig. 3.4(a). Both ClusRed and CM can approximate the PDF accurately. Figure 3.4(b) shows the PDF of LMP at bus 8. Neither ClusRed nor CM can approximate the PDFs accurately. One reason is that some physical limits of the system's components are reached, such as generation limit at bus 6 and transmission capacity limit between bus 6 and 10. Cumulant-based estimation method cannot perform well since the linear mapping matrix cannot provide an accurate linear relationship between LMPs and random loads at the solution point. Another reason is that the system is working close to physical limits and the Hessian matrix is ill-conditioned. The condition number in this case is 1.77×10^6, which is very large. The poor numerical stability can introduce considerable inaccuracy in matrix operations, which will in turn affect the accuracy of generating linear mapping matrix and the solution of POPF. In ClusRed, we operate on smaller matrices and they have better numerical stability. Experimental results show that ClusRed has better accuracy than CM.



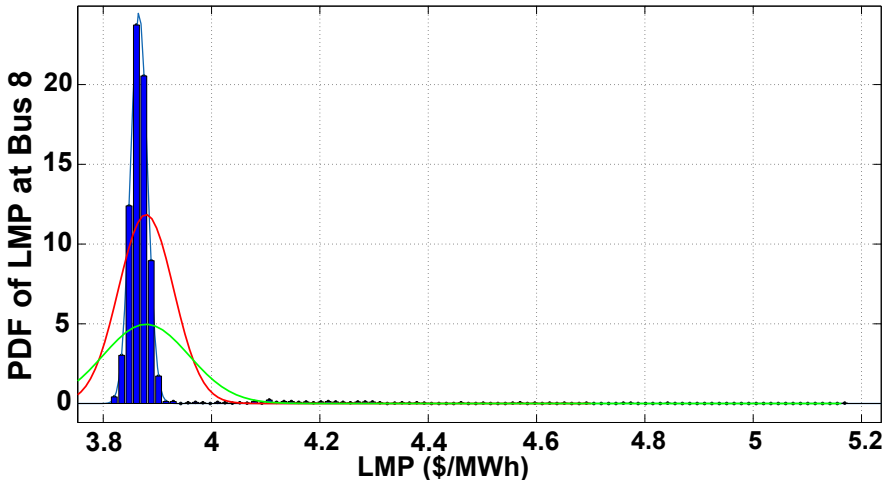
(a) μ (b) σ

(c) μ Error (d) σ Error

Figure 3.3: 30-bus test system



(a) PDF of LMP at bus 15



(b) PDF of LMP at Bus 8

Figure 3.4: LMPs at bus 15 and bus 8

3.4.2 Numerical Results for Other benchmarks

We test all the benchmarks to demonstrate the accuracy and efficiency of ClusRed. Mean absolute percentage error (MAPE), which computes the mean value of the absolute estimation errors of LMPs in percentage, is used to evaluate the estimation quality. Similarly, we compute worst absolute percentage error (WAPE) to evaluate the worst estimation. Experimental results show that both ClusRed and CM can accurately estimate the mean value of the distribution. As shown in Table 3.2, for small benchmarks, MAPE and WAPE are very small. For case-3120 and case-21k, MAPE and WAPE become larger. It shows that ClusRed is comparable with CM in terms of average accuracy, but performs 32.08% better for the worst estimation on average. The reason is that NR introduces a fixed amount of errors while ClusRed can generate the linear mapping matrix with better accuracy.

Table 3.2: Accuracy Evaluation of Mean

Benchmarks	CM		ClusRed	
	MAPE	WAPE	MAPE	WAPE
Case 9	0.08%	0.11%	0.08%	0.11%
Case 30	0.03%	0.29%	0.02%	0.13%
Case 118	0.02%	0.08%	0.02%	0.07%
Case 3120	1.25%	3.57%	1.28%	2.11%
Case 21k	3.73 %	10.84%	3.87%	5.22%

The standard deviation is also evaluated to demonstrate the accuracy of ClusRed. As shown in Table 3.3, compared to CM, ClusRed greatly improves the accuracy of the worst case scenarios when the Hessian matrix is ill-conditioned. In this experiment, both case-3120 and case-21k have very large condition numbers for their Hessian matrix. On average, the accuracy is improved by 36.76% in terms of WAPE. For case-21k, because of the extremely long run time, we only have 1,500 sample MCS results. Considering the size of this benchmark with much more random variables, at least 20,000 samples are desired. Thus the MCS results here are not accurate enough and biased the accuracy comparison.

We demonstrate the efficiency of ClusRed algorithm by comparing its runtime with those of other algorithms. As shown in Table 3.4, compared to the CM method, we achieved 4.57X speedup for the large-scale benchmark.

Table 3.3: Accuracy Evaluation of Standard Deviation

Benchmarks	CM		ClusRed	
	MAPE	WAPE	MAPE	WAPE
Case 9	3.17%	3.47%	3.15%	3.36%
Case 30	5.29%	80.99%	4.63%	33.9%
Case 118	2.84%	6.18%	2.43%	5.39%
Case 3120	6.78%	93.57%	5.28%	37.11%
Case 21k	10.24 %	148.84%	12.07%	67.22%

Table 3.4: Computation Time

Benchmarks	MCS	CM	ClusRed	SU1*	SU2**
Case 9	630s	0.24s	0.19s	3315	1.26
Case 30	890s	0.31s	0.22s	4045	1.41
Case 118	1160s	0.53s	0.27s	4296	1.96
Case 3120	13.29h	28.58s	12.56s	3809	2.28
Case 21k	156h	2998s	656s	856	4.57

*: Speedup compared to MCS; **: Speedup compared to CM.

3.5 Conclusion

With increasing uncertainties involved in smart grids, solving POPF is becoming more and more important in smart grid operation. Previous cumulant methods are neither fast nor accurate enough for large-scale smart grids. In this chapter, we propose a fast clustering and NR based cumulant method, ClusRed, that can solve POPF much faster and more accurately. We also developed a new linear mapping matrix based on NR that has better performance than previous approaches. For large-scale smart grids, compared to inversing high dimensional Hessian matrix, generating the new linear mapping matrix is much faster.

We demonstrate that ClusRed can achieve several thousandfold of speedup compared to MCS and up to 4.57X compared to previous CMs for large-scale smart grids. In addition, the experimental results show that we maintain a high level of accuracy compared to MCS. On average, we improve the worst estimation accuracy of mean value by 32.08% and standard deviation value by 36.76%. The proposed method can be used to solve POPF problems for large-scale smart grids in many applications, such as market management and reliability analysis.

CHAPTER 4

OPTIMAL BLOCKING DEVICE PLACEMENT FOR GEOMAGNETIC DISTURBANCE MITIGATION VIA BRANCH AND CUT ALGORITHM

4.1 Introduction

The quasi-dc geomagnetically induced currents (GICs) introduced by geomagnetic disturbances (GMDs), such as solar storms, have been concerning power grid operators for many years [21, 22]. Over the last several years, the power industry has seen more concentrated interest in this area. The US Federal Energy Regulation Commission (FERC) now requires US utilities to perform GMD vulnerability assessment, and to prepare corrective action plans and mitigation actions [23]. GMDs can cause rapid geomagnetic field variation, which in turn produces GICs flowing through transmission lines. GICs tend to flow through the neutral connection of transformers and can cause half-cycle saturation of transformers. As a result, power systems may suffer from transformer overheating and severe reactive power losses. GICs can damage the bulk power system assets, typically associated with transformers. This may eventually lead to system reliability issues, such as misoperations of proactive relays, and voltage instability [24]. The impact of GIC, GIC modeling and analysis are studied in [25–30].

Many methods were proposed to mitigate the effects of GIC [31, 32]. One widely recognized solution is to install blocking capacitors or switchable resistors to the neutral connection of Wye-connected transformers [33, 34]. In February 2015, the first blocking device was installed on a 345/138 kV transformer in Wisconsin to increase the power system's resilience to GMDs [59]. Considering that each blocking device is very expensive to install and maintain [34], it is crucial to place these BDs effectively and efficiently. We formulate this problem as an optimal BD placement (OBP) problem aiming

at selecting transformers in a power system to place BDs that minimize the damage of GICs.

In [35], a BD placement problem is proposed to minimize the purchase and installation costs of BDs while satisfying power system voltage and maximum generator reactive power limits. In [36], thermal limits of power equipment and power system operation constraints are considered in addition. However, the works mentioned above failed to solve the problem of minimizing GMDs' damages, which can usually be measured by the reactive power losses in a system. In [37], a BD placement problem that can minimize the reactive power losses is proposed and it is formulated as a mixed integer second-order cone programming (SOCP) problem via linear approximation. By taking advantage of the local blocking effect, an efficient blocking solution is developed by choosing substations, instead of transformers. However, due to the limitation of linear approximation, this relaxed SOCP method cannot always find the optimal solution. In addition, it can only provide blocking solutions on the substation level. Since a BD is physically placed at a transformer and a substation can contain several transformers, blocking a number of substations does not directly reflect the cost.

In this chapter, we formulate the OBP problem as a mixed integer nonlinear programming (MINLP) problem and use a branch and cut algorithm to search for the optimal solution [38]. Since there are several thousand high voltage transformers in the very large US electric grid, it is non-trivial to find the optimal solution. This OBP problem is a combinatorial optimization problem. Generic search algorithms are not guaranteed to find an optimal solution as the problem is a non-convex problem and generic search algorithms can easily get stuck at local optima. The key idea is that since OBP problem is constrained by the power system network equation, we can use bounding and pruning to reduce the search space and can potentially speed up the computation while guaranteeing to find the optimal solution. Our goal is to find the optimal solution which is indicated by the binary variables that represent whether to place blockers at transformers or not. In order to solve this problem efficiently, we introduce additional continuous variables to keep the low degree of unknown variables in the constraints in the OBP problem so that we can take advantage of the properties of the constraints when using branch and cut algorithm. Due to the relationship between the binary variables and the additional variables, we can branch

only on the binary variables instead of the continuous variables. In this way, we can solve the OBP problem efficiently. Our main contributions are as follows:

- We formulate the OBP problem as an MINLP problem that is suitable for branch and cut solvers.
- We apply a branch and cut based algorithm and develop proper branching strategy to find the globally optimal solution.

The rest of this chapter is organized as follows. In Section 4.2, we introduce the necessary background on power system modeling with GICs. Section 4.3 describes the formulation of the OBP problem. Solution method and detailed analysis regarding the theoretical structure of OBP problem and branch and cut method are presented in Section 4.4. We present the numerical results in Section 4.5 and the conclusion in Section 4.6.

4.2 Preliminaries

The modeling methodologies of GICs and their effect on power systems have been well studied in [25, 27, 28]. During geomagnetic disturbances, geomagnetic field variations introduce low-frequency GICs along transmission lines and flow through substation transformer to ground. Compared to the 50/60 Hz AC electric currents in the power network, the frequency of GICs is very low. As a result, GICs are modeled and analyzed based on DC analysis for all buses and substations.

4.2.1 Network Modeling

We start with DC network modeling for GICs. Consider a power system with N_B buses, N_L lines, N_S substations and N_K transformers. In GIC analysis, buses and substation neutrals form all the nodes in the GIC's DC network. We denote by $\mathcal{N} \triangleq \{1, 2, \cdots, N_B\}$ the set of buses and by $\mathcal{L} \triangleq \{\ell_1, \ell_2, \cdots \ell_{N_L}\}$ the set of transmission lines that connect the buses in the set $\mathcal{N}$. We associate with each line $\ell \in \mathcal{L}$ the ordered pair $\ell = (i, j)$. The series conductance of line ℓ is denoted by g_ℓ. We denote by $\mathcal{S} \triangleq \{0, 1, 2, \cdots, N_S\}$ the set of substations and $\mathcal{K} \triangleq \{0, 1, 2, \cdots, N_T\}$ the set of transformers.

Each bus i is characterized by the nodal voltage magnitude, V_i. We denote by $\boldsymbol{Y}$ the $(N + N_S) \times (N + N_S)$ nodal admittance matrix, with Y_{ij} as the element in row i and column j. In this case $\boldsymbol{Y} = \boldsymbol{G}$, where $\boldsymbol{G}$ is the conductance matrix. If a bus i is connected to another bus j, then Y_{ij} is the conductance of line (i, j). If a bus i is connected to a substation j then Y_{ij} is the conductance of transformer windings to the substation neutral. Similarly we have:

$$\boldsymbol{I} = \boldsymbol{G}\boldsymbol{V}, \tag{4.1}$$

where $\boldsymbol{I}$ is the vector of nodal current injection and $\boldsymbol{V}$ is the vector of nodal voltage.

4.2.2 Calculating Induced Voltage

To evaluate the GICs effects, let the geomagnetic field be

$$\vec{E} = [E^{north}, E^{east}],$$

where E^{north} is the geomagnetic field along the north direction and E^{east} is the geomagnetic field along the east direction. The induced voltage in the transmission between bus i and j is computed by integrating the geomagnetic electric field along the route of the line:

$$V_{ij} = \int_i^j \vec{E} d\vec{l}_{(i,j)}, \tag{4.2}$$

where $d\vec{l}_{(i,j)}$ is the incremental line segment length including direction. Integration path is the route of the line between bus i and bus j. As we assume that $\vec{E}$ is constant, then we have

$$V_{ij} = \vec{L} \cdot \vec{E} = L_{ij}^{North} E_{ij}^{North} + L_{ij}^{East} E_{ij}^{East}, \tag{4.3}$$

where L_{ij}^{North} is the northward distance and L_{ij}^{East} is the eastward distance between bus i and j. With the induced voltage, by doing Norton equivalent we can then calculate the GICs current injection at each bus.

$$I^{gic} = \sum_j g_{ij} V_{ij} = G_{Line} \cdot \vec{L} \cdot \vec{E} = H\vec{E}, \tag{4.4}$$

where matrix H is with the size of $N_b \times 2$. H can also be written as

$$H = [H_{North}, H_{East}]^T. \tag{4.5}$$

4.2.3 DC Network Analysis

The GICs, I^{gic}, at each bus is used to calculate the voltage at each bus using the DC model of the network constructed in the previous section.

$$V = G^{-1} I^{gic}. \tag{4.6}$$

4.2.4 GICs Effects on Transformers

As we have the DC voltage of all buses and all transformer neutrals, we can then calculate the GICs going through a transformer using the following equation:

$$I = g_{mn}(V_m - V_n). \tag{4.7}$$

The effect of GICs on each transformer is characterized as the effective current going through the transformer [28].

$$I_t^{effective} = I_H + I_L/\alpha_t, \tag{4.8}$$

where I_H is the transformer GIC going through the high-side winding for a conventional transformer or the series winding for an autotransformer, and I_L is the transformer GIC going through the low-side winding for a conventional transformer or the common winding for an autotransformer, and α_t is the transformer turn ratio. I_H and I_L can be calculated by (4.7). Expressing the above equations in matrix format, we have

$$I_t^{effective} = \phi V. \tag{4.9}$$

Then we can calculate the reactive power loss

$$Q_{loss}^t = \epsilon_t |I_t^{effective}|. \tag{4.10}$$

4.2.5 Effects of Placing Blockers

After placing blockers, the network conductance matrix is changed. Consider placing a blocker on transformer t connecting bus m and n, which sits at substation s. For a conventional transformer, it is equivalent to disconnecting both side windings from the substation. So G matrix is changed correspondingly by removing the contribution of the side windings to the G matrix.

$$\Delta G(s,s) = -g_{ms} - g_{ns}, \tag{4.11}$$

$$\Delta G(m,m) = -g_{ms}, \tag{4.12}$$

$$\Delta G(n,n) = -g_{ns}, \tag{4.13}$$

$$\Delta G(m,s) = \Delta G(s,m) = g_{ms}, \tag{4.14}$$

$$\Delta G(n,s) = \Delta G(s,n) = g_{ns}, \tag{4.15}$$

and $\Delta G(m,n) = \Delta G(n,m) = 0$, $\Delta G(x,y) = 0$ where $x, y \notin \{m,n,s\}$.

For blocking an auto transformer, it is equivalent to disconnecting common winding with the substation. Assume that bus n is the low side.

$$\Delta G(s,s) = -g_{ns}, \tag{4.16}$$

$$\Delta G(n,n) = -g_{ns}, \tag{4.17}$$

$$\Delta G(n,s) = \Delta G(s,n) = g_{ns}, \tag{4.18}$$

and $\Delta G(x,y) = 0$ where $x, y \notin \{n,s\}$.

4.3 OBP Problem Formulation

In this approach, in order to find the optimal blocker placement solution, we formulate this problem as a mixed integer nonlinear programming (MINLP) problem. For each transformer t_i, we assign a binary variable

$$x_i \in \{0,1\}$$

to indicate if there is a blocker or not. In our solution, $x_i = 0$ represents that no blocker is placed at transformer t_i while $x_j = 1$ represents that there is a

blocker at transformer t_j. Supposing that we have K transformers, the goal is to find $X = \{x_1, x_2, ..., x_K\}$ that can minimize the hazards of GICs.

Based on the analysis of GICs modeling above, for each blocker we placed, the admittance matrix (G) is changed correspondingly. Supposing that we have K transformers in the power system, then the admittance matrix is

$$G = G_0 + \sum_{i=1}^{K} x_i \Delta G_i, \tag{4.19}$$

where G_0 is the admittance matrix without any additional blockers in the system and ΔG_i is the change on admittance matrix by placing a blocker at t_i. After placing a number of blockers, the final admittance matrix is G.

As we discussed in last section, Φ will also change when placing blockers. Similarly, we have

$$\Phi = \Phi_0 + \sum_{i=1}^{K} x_i \Delta \Phi_i, \tag{4.20}$$

where Φ_0 is the correlation matrix with no additional blockers in the system and $\Delta\Phi_i$ is the change on the correlation matrix by placing a blocker at t_i. After placing a number of blockers, the final correlation matrix is Φ.

Consider the total loss of reactive power

$$Q_{loss}^{total} = \sum_{i=1}^{K} \epsilon_i |I_{t_i}|, \tag{4.21}$$

where I_{t_i} is the current going through transformer t_i and ϵ_i is the corresponding coefficient for transformer t_i which is related to the transformer itself only.

To compute the reactive power loss, firstly, we need to calculate the GICs introduced by geomagnetic field using (4.4). Secondly, we can compute the voltage at each bus and substation neutral using (4.6) and 4.4. Thus, we have the following equation:

$$V = G^{-1} I^{gic} = G^{-1} H \vec{E}. \tag{4.22}$$

Then, we can calculate the current going through each transformer

$$I_t = \Phi V = \Phi G^{-1} H \vec{E} = \Psi \vec{E}, \tag{4.23}$$

where V represents the voltage at each bus and Φ is the corresponding admittance matrix for calculating the currents flowing through the transformers. Ψ has the size of $N_t \times 2$. Last, we can use (4.21) to compute the reactive power loss. The above flow shows the GICs analysis process.

The overall hazards caused by GICs can be characterized by the total reactive power loss. As a result, we minimize the total reactive power loss in OBP. Given a power system together with the constraints on DC voltage at each bus and the total budgets on installing blockers, the goal of placing blockers in the power system is to minimize Q_{loss}. Modeling this mathematically gives us the following optimization problem.

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & Q_{loss}^{total} = \sum_{i=1}^{K} \epsilon_i |I_{t_i}| \\ s.t. \quad & I^{gic} = GV \\ & I_t = \Phi V \\ & I^{gic} = HE \\ & \sum_{i=1}^{K} x_i \leq N \end{aligned} \quad , \tag{4.24}$$

where N is the maximum number of blocking devices that can be installed in the system which is limited by the budget. The objective function simply is the total reactive power loss. The equality constraints are related to power system itself and are equivalent to (4.4), (4.22) and (4.23). The inequality constraints are constraints on max number of blockers and max DC voltage allowed.

Notice that matrix G and Φ actually contain unknown variable x_i. For each transformer, there are corresponding changes made on matrix G if a blocker is placed. So matrix G contains x_i on many rows and columns that correspond to transformer's substation and its related buses. If we explicitly express the inversed matrix G^{-1} using x_i by (4.22), the inversion of G will contain $\prod_{i \in B} x_i$ which has a very high degree. As a result, it is not desirable to explicitly express V in terms of x_i because it will be much more difficult to handle x_i in high degrees. Instead, we take in all the elements in vector V as unknown variables, which we denote as v_j.

Now, the goal is to find out X and V by solving the above OBP problem that can minimize the total loss of reactive power. Because X is a vector of binary values and V is a vector of continuous values, the problem is a mixed integer problem. In addition, both the constraints and the objective functions are nonlinear. Therefore, the OBP problem is a mixed integer nonlinear programming (MINLP) problem. We also notice that there is dependency between X and V, as we can always calculate V with a given value of X. So the problem is a combinatorial problem.

Because there are several hundred transformers in a typical power system, the number of potential solutions $\binom{N}{K}$ is extremely large. It is impossible to enumerate all the possible solutions, perform GICs analysis and pick the best solution simply because the running time is too long. On the other hand, we want to find the optimal solution as each blocker costs a large amount of money. Approximations and inequivalent relaxations are not desirable. As a result, we need to develop a fast method that can find the optimal solution.

4.4 Methods

In this section, we present our solution using branch and cut method that can find the optimal solution for the OBP problems in large-scale power systems. We take the objective function of minimizing total reactive power loss as an example to illustrate this method. In this problem, the unknown variable that we are interested in is the blocker placement vector X which indicates which transformers to be blocked. This method can be easily extended to solve OBP problems with any other nonlinear objective functions, such as minimizing the maximum reactive power loss, total sum of currents going through transformers and so on.

4.4.1 Approach Overview

The overall algorithm flow of our approach is shown in Fig. 4.1. Firstly, we need to construct the OBP problem. To construct the OBP problem, we gather the longitude and latitude information for all buses and substations. Then, we calculate the GICs in transmission lines by the geomagnetic electric field information and location information. We also construct the

corresponding graph for the given power system. Then we calculate the admittance matrix G before placing any blockers. By Kirchhoff's circuit laws, we can compute the effective GICs going through each transformer by using the voltage at each bus and each substation. Expressing the relationship between V and I_t in matrix format, we get Φ matrix. Then we compute ΔG and $\Delta\Phi$ for each blocker we placed using (4.19) and (4.20). After computing these necessary components, we can then put everything together and generate the OBP problem explicitly. Finally, we use branch and cut algorithm to solve this OBP problem.

4.4.2 OBP Problem Generation

As we want to address the problem of generating polyhedral relaxations of multivariate functions, we need to generate the OBP problem explicitly so that the convexity can be exploited. In order to improve the efficiency of solving this OBP problem using branch and cut method, we need to improve the convexity of the OBP problem so that we can take advantage of this condition to speed up the bounding process. In OBP problem formulation, we introduce additional variables V, which represents the voltages at each bus and each substation neutral. Given the value of X, V can be calculated by (4.6). As shown in (4.24), instead of explicitly expressing V in terms of X, we keep V and make it easier for convex relaxation of the constraints which defined the search space. It is worth mentioning that we also set upper and lower bounds for continuous variables in OBP problem using our knowledge of real power systems. For example, quasi-DC voltage cannot be too high in real systems as it can be dangerous to surrounding equipment especially under humid weather conditions. Based on the constraints on voltage, the search space is reduced significantly to a smaller rage. We may also achieve better convex relaxation and polyhedral outer approximations in a smaller range.

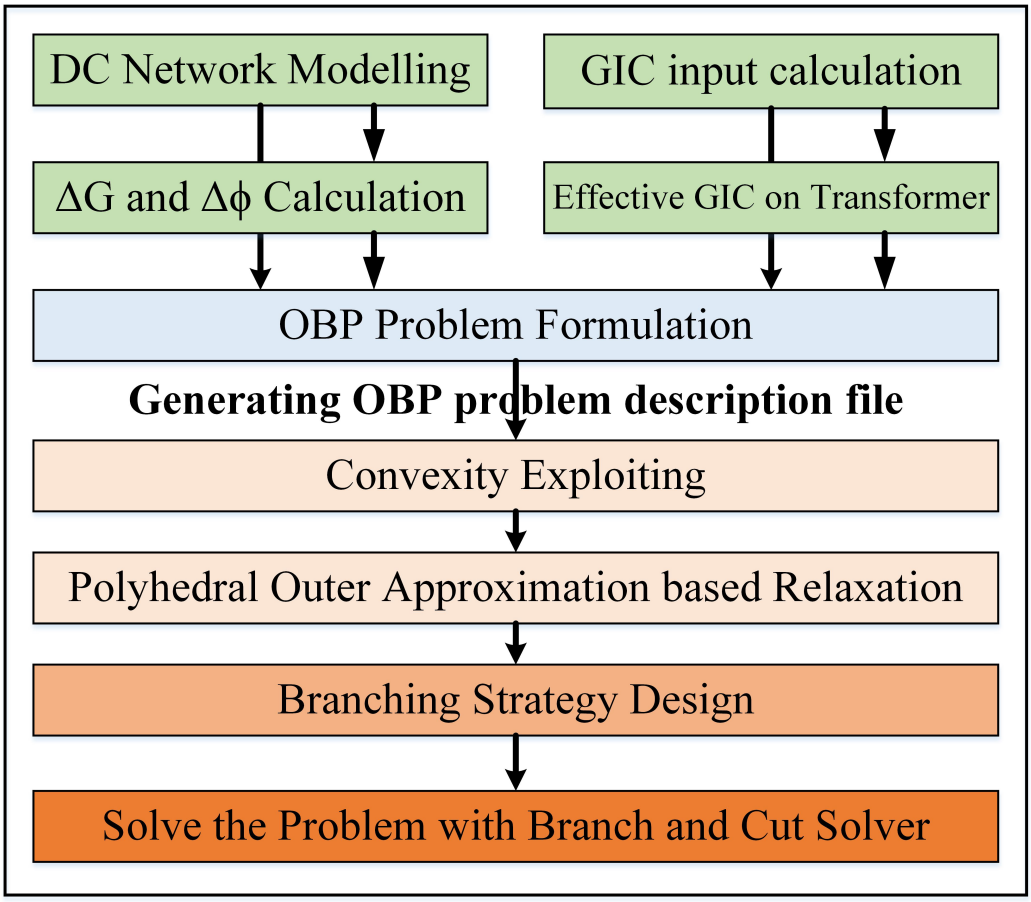


Figure 4.1: BC method

4.4.3 Branch and Bound Based Method

A branch and bound algorithm is an algorithm that explores branches of state space tree, which represent subsets of the solution set, by performing systematic enumeration of candidate solutions [38].

In the search tree, each node is a subset of the whole solution space. It recursively computes the upper and lower bounds of the minimum of the objective function with a given subset of solution space [38]. If the lower bound for a tree node A is greater than the upper bound for some other node B, then A can be safely discarded from the search. Assume the current minimum upper bound is $\mathcal{U}$, so any node whose lower bound is greater than $\mathcal{U}$ can be discarded. By enumerating the candidate solutions of branches in an recursive way, the search space is reduced. The recursion stops when the current candidate set $\mathcal{S}$ is reduced to a single element, or when the upper

bound for set $\mathcal{S}$ matches the lower bound. Either way, any element in $\mathcal{S}$ is a minimum of the objective function [38].

4.4.4 Branch and Cut Algorithm for OBP

The branch and cut algorithm is a branch and bound based algorithm. It firstly decomposes factorable function to derive relaxation of the OBP problem. Then it uses polyhedral outer approximation of the above derived relaxations to generate linear programming based relaxation. As the constraints of OBP problem contain DC power flow equation and the unknown variable x_i appears in G matrix, the constraints on the OBP problem are nonlinear. Therefore, the relaxations are also nonlinear which makes the outer approximation necessary as it is much easier to solve linear relaxations and perform domain reduction, branching and pruning. Lagrangian/linear programming duality is then used to develop domain reduction strategies. In particular, duality based domain reduction strategy is used in branch and bound algorithm [38]. A learning heuristic that improves initial branching decisions by relaying data across siblings in a branch and bound tree is also used [38].

More importantly, at each iteration, how the branch and cut algorithm chooses the next search node governs the structure of the search tree and thus determines the performance of the algorithm and the memory it uses. Priority based branching strategy is used in branch and cut algorithm and has been proved to be very efficient [38]. In this OBP problem, there is a very large search space on X and V, which makes the problem very difficult to solve. We notice that branching on continuous vector V in nonlinear programming may require infinitely many subdivisions. Fortunately, there are underlying relations between binary variables X and continuous variables V which is shown in (4.22). As X determines the value of continuous V, we can take advantage of this relation between them and solve this problem more efficiently. To be more specific, in our solution, we branch only on binary variables x_i which only creates two subproblems from each branching. We will not miss any solution subspace because the solution space of V reduces as the solution space of X reduces. In this way, with the help of the special structure of the OBP problem, the efficiency of the algorithm at solving the OBP problem is greatly improved.

4.5 Results

To test this algorithm, we use three test benchmark [24, 30, 60] which were also used in [37]. They are summarized in Table 4.1. We run all the tests on a multi-core machine, which has Intel Core i5 processor of 3.10 GHz and 16 GB memory. We use constant geomagnetic electric field $|E| = 1$ and let the direction be East to evaluate the performance of our solution.

Table 4.1: Test Benchmark

Benchmarks	Bus No.	Substation No.	Transformer No.
Case NERC 6-bus	6	3	3
Case EPRI 20-bus	19	8	15
Case large	1875	865	756

In order to demonstrate the optimality and computational efficiency of our method, we compare it with exhaustive search (ES) solution, which can find the "optimal" solution by searching the entire solution space. ES enumerates all the possible blocker placement combinations in a given power system, performs GIC analysis and then picks the optimal one that gives the best solution. We verify our results by comparing them with the results from PowerWorld [60]. However, for a power system in the real world, there are many transformers and thus many blocker placement solutions. It is impossible to use ES to enumerate all the solutions and find the optimal solution. For example, in the large benchmark used in this work, there are about 2^{756} possible ways to block the transformers, which is impossible for ES. We also compared our algorithm with linear approximation based method previously published in [37]. However, they assume to block transformers in a whole substation. The drawback of this assumption will be discussed later in this section.

4.5.1 Numerical Results on Small Systems

We test our algorithm on a small test system with 3 substations and 6 buses [24]. There are 3 transformers in the power system, and one of them is auto transformer. Experimental results show that our algorithm can find the optimal blocker placement solution given all number of blockers we can

place. As shown in the $Q_{Loss} - N_b$ curve in Fig. 4.2, there is 46.5 Mvar reactive power loss if no blockers are placed. After placing the first blocker at the conventional transformer, Q_{Loss} is reduced to 34.38 Mvar. If we place two, the reactive power loss is reduced to 0. Based on the analysis in Fig. 4.2, we do not need to place more than two blockers in the system as there is no benefit to place the third blocker.

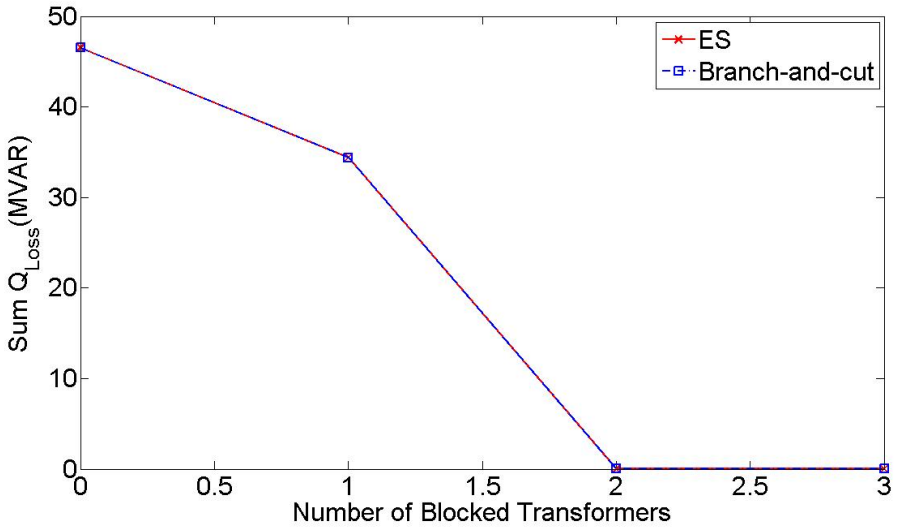


Figure 4.2: Blocker placement for NERC 6-bus system

We also test our algorithm on a system with 8 substations and 19 buses. Notice that by convention, this benchmark is called 20-bus test system [37]. Figure 4.3 shows the optimal solution for OBP problem. Given the maximum number of blockers that can be placed, our algorithm can find out the optimal blocker placement solution that minimizes the reactive power loss. As the number of blockers increases, reactive power loss decreases. We observe that the decrease in reactive power loss is not always the same. The analysis in this chapter can indicate the marginal gain of placing each blocker. For example, we notice that there is a sharp decrease of reactive loss when we place the third blocker. For all number of blockers we can place, our algorithm provides the optimal solutions and is verified by ES. Under the circumstance of placing a number of blockers in a real power system, we can make the decision by doing a cost-benefit analysis based on the $Q_{Loss} - N_b$ curve.

The previous method in [37] cannot always find the optimal solution for all numbers of blockers that can be placed. For example, as shown in the

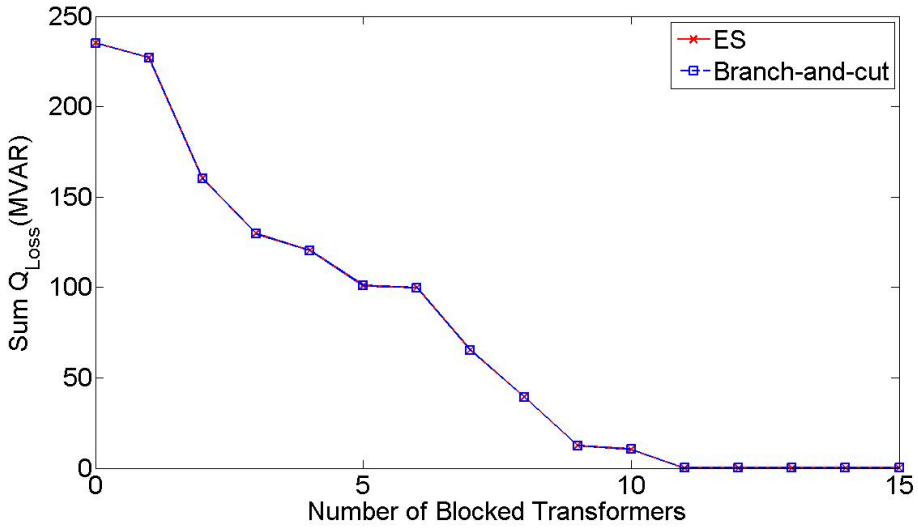


Figure 4.3: Blocker placement for 20-bus system

results of [37], optimal solution cannot be found if we are allowed to place blockers at 5 substations in the same 20-bus test system we are using here. In addition, the previous method in [37] assumes the basic blocking unit to be a substation which is not accurate enough. Since a blocker is physically placed at a transformer instead of a substation and a substation can contain several transformers, blocking a number of substations does not indicate how many transformers are blocked and how many blockers are needed. Thus the previous method in [37] cannot find the optimal blocker placement solution given the max number of blockers available.

In this benchmark, there are 15 transformers while there are only 8 substations. Compared with the formulation of OBP problem in [37], the OBP problem here is thorough and complete but more complex as well, and the searching space is much larger. Solving this OBP problem optimally represents a major milestone for the blocker placement problem.

4.5.2 Numerical Results on Large System

In order to demonstrate the effectiveness of our algorithm in real world application, we test our algorithm on a power system located in the upper Midwestern United States, an equivalent system model built from the Eastern Interconnection model. There are totally 756 transformers in the large

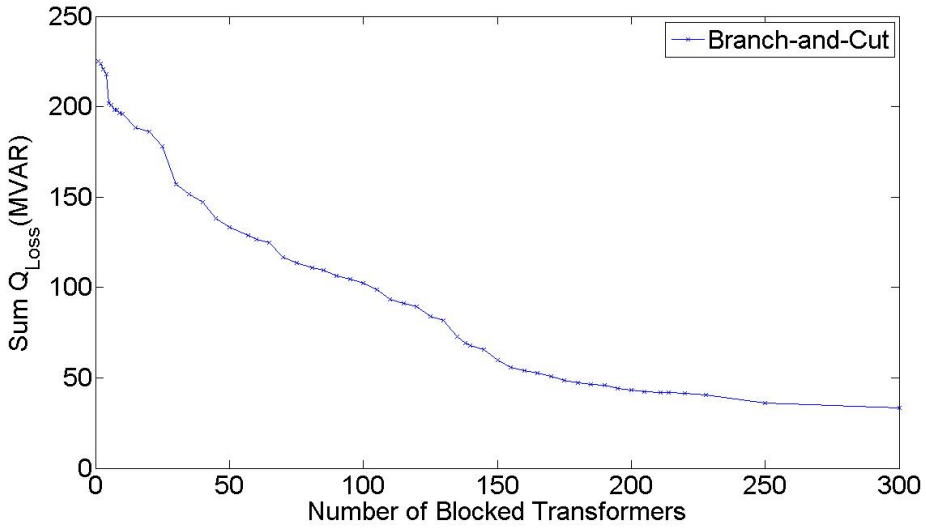


Figure 4.4: Blocker placement for a large-scale system

system, with 78 high voltage transformers which are connected to 200 KV networks. As discussed at the beginning of this section, ES cannot be used to solve this problem because there are too many possible solutions that make it impossible to search the entire solution space.

Our solution uses branch and cut algorithm and can efficiently reduce the solution space by branching and bounding. Given enough time, which is much less than it takes for ES to find the optimal solution, branch and cut based algorithm can find the optimal solution as it only removes the solution subspace that does not contain the optimal solution. When the N in the OBP problem is large, the convergence rate sometimes becomes small. Although we do not place that many blockers in reality, we collect the results for large Ns to show the performance of our method and to make the results complete. Due to the limitation on the time we have for collecting results, we set the maximum running time to be 10^4 seconds when N is large and pick the best solution found in the limited time.

Figure 4.4 shows the relationship between the number of blockers and the minimum reactive loss. This analysis provides a lot of useful information. For example, it is not very helpful to place just one blocker in the system. We may want to place at least around 5 blockers to obtain a significant decrease in reactive power loss. Also, it is not quite beneficial to place more than 200 blockers because the reactive power loss decreases very slowly as the number

of blockers increases. In addition, experimental results show that we cannot place too many blockers as it may violate the voltage constraints and make the OBP problem infeasible. For example, if 700 blockers are placed, then the voltage of some buses will become too high.

4.6 Conclusion

This chapter presents a method that can find the optimal solution for OBP problem which mitigates the effects of GICs using branch and cut algorithm. We formulate the OBP problem to MINLP and carefully design the branching strategy to further reduce the searching efforts. The previous method suffers from the fact that it cannot always find the optimal solution. However, it is ideal to maximize the benefit, especially under the condition that the cost of installing a blocker in a real power system is very high. As a result, a method that can provide optimal solution is highly desirable. We demonstrate that our method can provide optimal solution by testing it on two small test systems and one large-scale power system. This solution can be used to guide the installation of blockers in real power systems. Also, the marginal benefit of placing a blocker can be obtained using this solution. Our method can not only provide a solution to OBP problem, but also offer a framework for solving OBP related problems. Our algorithm can be applied to solve a large range of OBP related problems with different objective functions and different constraints.

CHAPTER 5

OPTIMAL BLOCKING DEVICE PLACEMENT FOR GEOMAGNETIC DISTURBANCE MITIGATION VIA SIMULATED ANNEALING

5.1 Introduction

In the Chapter 4, we formulate the OBP problem as a mixed integer nonlinear programming (MINLP) problem and develop a branch and cut (BC) based algorithm to solve it. However, the BC method does not scale up well to large-scale power systems. Furthermore, all previous methods suffer from some significant limitations.

First, the impact of BD placement on neighboring interconnected systems has never been studied. Previous studies show that blocking GICs through one transformer will likely result in the redistribution of GICs to nearby transformers [35, 37]. Since power systems are highly interconnected, when blocking transformers in one system, it is necessary to evaluate and limit the impact on neighboring systems. As a result, solving the block by system (BBS) scenario for the OBP problem is very important as we do not want to redistribute the damages to other interconnected systems. In the BBS scenario, limiting the BD placement impact to other systems introduces non-convex constraints to the OBP problem, which significantly increase its complexity. Second, all previous methods do not use a realistic time-varying geoelectric field (TVGF). The geoelectric field introduced by GMDs has been assumed to be constant to simplify GIC and OBP analysis in previous studies [25, 26, 28, 35, 61]. However, the amplitude and direction of GMDs usually keep changing violently throughout time [62]. As a result, it is more realistic and accurate to evaluate GMD damage and solve OBP problems using a TVGF. In addition, as stated in the TPL-007-2 Geomagnetic Disturbance Reliability Standard [23] [63], utilities are currently required to conduct GMD

vulnerability assessment using a TVGF in the March 13-14 1989 GMD event that caused the Hydro Quebec blackout. This standard further motivates us to consider the TVGF scenario for the OBP problem.

The community is in need of a new algorithm that addresses the issues mentioned above and can produce a high-quality solution in the given affordable computational time. In respond, we developed a simulated annealing (SA) [39, 40] based algorithm that can accelerate the solution process, produce a near-optimal solution in a relatively short time, and solve the BBS and TVGF scenarios for OBP problems. When computational time is not restricted, we can achieve the exact optimum since the SA algorithm statistically guarantees global optimal [64]. The performance of the SA algorithm depends highly on its optimization strategy. In this chapter, several strategies are carefully designed and evaluated. We select one optimization strategy that can find the high-quality solutions most efficiently. Instead of concentrating on the physical modeling aspect of OBP problems, we focus on the algorithmic and mathematical aspects of solving OBP problems. In reality, different system operators have various practical considerations and limitations. With this in mind, we solve several different and complex scenarios of OBP problems and demonstrate that our solution framework can be easily extended to solve OBP problems with various configurations. To sum up, our main contributions are as follows:

- We develop a fast SA method that can produce high-quality near-optimal solutions for OBP problems.
- For the first time, we study the BD placement impact on interconnected systems and solve for the BBS scenario.
- For the first time, we study the time-varying nature of geoelectric field and solve for the TVGF scenario.
- We demonstrate that our SA method can be easily extended to solve OBP problems with various configurations.

The rest of this chapter is organized as follows. In Section 5.2, we introduce the formulation of the basic OBP problem, followed by the BBS and TVGF scenarios. Section 5.3 describes our SA algorithm and the detailed SA design. We present the numerical results in Section 5.4 and conclude in Section 5.5.

5.2 OBP Problem Formulation

The modeling of GIC and their effects on power systems has been presented in Section 4.2 in Chapter 4. The OBP problem is formulated as a mixed integer nonlinear programming (MINLP) problem. Compared to the formulations of the BD placement problems in previous studies [35–37], our formulations are more thorough, complete and realistic. At the same time, it is also more mathematically complex and more challenging to solve.

Similar to the OBP formulation in Section 4.3, for each transformer t, we assign a binary variable $x_t \in \{0, 1\}$ to indicate whether there is a BD. The expression $x_t = 0$ represents that no BD is placed at transformer t while $x_t = 1$ represents that a BD is placed. The overall effects of GICs on a power system can be quantified by the total reactive power losses over all transformers, or by the maximum reactive power losses, also over all transformers, as defined by

$$Q_{total}^{loss} = \sum\nolimits_{t=1}^{N_T} Q_t^{loss} = \sum\nolimits_{t=1}^{N_T} \epsilon_t |I_t^{eff}|, \tag{5.1}$$

$$Q_{max}^{loss} = \max_{t=1}^{N_T} Q_t^{loss} = \max_{t=1}^{N_T} \epsilon_t |I_t^{eff}|. \tag{5.2}$$

Besides the constraints in the OBP problem in (5.3), another important yet difficult constraint is the limit on per-phase effective GIC. The thermal assessment in [65] has shown that 75 A is a conservative threshold for per-phase effective GIC and NERC has set the transformer thermal impact screening criterion to 75 Amps in TPL-007-002 [63]. As a result, solving OBP problems with this constraint is necessary. Throughout the chapter, we will use either Q_{total}^{loss} in (5.1) or Q_{max}^{loss} in (5.2) as the objective function to illustrate our formulation and solution. OBP problems using other metrics can also be solved with the same solution framework. We will introduce the formulation of the basic OBP problem, the block by system (BBS) scenario and time-varying geoelectric field (TVGF) scenario in the rest of this section.

5.2.1 Basic OBP

Based on the formulation in (4.24), we add the per-transformer GIC constraint and condense the expressions to get the new formulation for the basic

OBP problem. It is described as follows: Given a power system, the budget on installing BDs and the maximum effective GIC constraint on individual transformers, find the optimal way to place these BDs so that Q_{total}^{loss} is minimized.

$$\begin{aligned} \min_{X} \quad & Q_{total}^{loss} = \sum\nolimits_{t=1}^{N_T} \epsilon_t |\boldsymbol{I}_t^{\mathit{eff}}(\boldsymbol{X})| \\ s.t. \quad & \boldsymbol{H}\vec{\boldsymbol{E}} = \boldsymbol{G}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}) \\ & \boldsymbol{I}_t^{\mathit{eff}}(\boldsymbol{X}) = \boldsymbol{\Phi}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}) \quad , \\ & \sum\nolimits_{t=1}^{N_T} x_t \leq N \\ & \boldsymbol{I}_t^{\mathit{eff}} \leq \mathbf{I}_{max}, \forall t \end{aligned} \tag{5.3}$$

where N is the maximum number of BDs, which is limited by the budget, and $\mathbf{I}_{max}$ is the limit on per-phase effective GIC. Solving this MINLP problem is not trivial. The total number of potential solutions is $\binom{N_T}{N}$. Since there are several hundred transformers in a typical power system, the solution space is extremely large. For example, given 200 transformers and 10 BDs, the total number of possible solutions is 2.24e^{16}. It is impossible to enumerate all possibilities, perform GIC analysis and find the best solution. On the other hand, we want to find, or at least approach, the optimal solution because each BD is very expensive.

5.2.2 Block By System (BBS) Scenario

Consider a power system A that is connected to another power system B, and we are looking for the OBP solution for system A. Each system contains a set of transformers, denoted by $T^A = \{T_1^A, T_2^A, ...\}$ and $T^B = \{T_1^B, T_2^B, ...\}$. For a transformer t in system B, let $Q_t^{loss\_base}$ be the reactive power losses before any BD is placed in system A and Q_t^{loss} be the reactive power losses after placing BDs in system A. The impact on system B of placing BDs in system A can be quantified by the percentage difference of the total reactive power losses in system B before and after BDs are placed in system A.

$$\Delta Q_B^{loss} = \frac{\sum_{t \in T^B} Q_t^{loss}}{\sum_{t \in T^B} Q_t^{loss\_base}} - 1. \tag{5.4}$$

In reality, system operators may consider limiting the impact on interconnected systems to be non-negative or restrict the negative impact to a certain

degree. The OBP problem's BBS scenario is formulated as

$$\begin{aligned} \min_{X} \quad & Q_A^{loss} = \sum\nolimits_{t\in T^A} \epsilon_t |\boldsymbol{I}_t^{\mathit{eff}}(\boldsymbol{X})| \\ s.t. \quad & \boldsymbol{H}\vec{\boldsymbol{E}} = \boldsymbol{G}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}) \\ & \boldsymbol{I}_t^{\mathit{eff}}(\boldsymbol{X}) = \boldsymbol{\Phi}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}) \\ & \sum\nolimits_{t=1}^{N_T} x_t \le N, \quad \Delta Q_B^{loss} \le \alpha \end{aligned} , \tag{5.5}$$

where α is a number between 0 and 1 and represents how much negative impact is allowed on system B.

5.2.3 Time-varying Geoelectric Field (TVGF) Scenario

In previous studies, the geoelectric field was assumed to be constant to simplify GIC and OBP analysis. However, in reality, the amplitude and direction of the geoelectric field usually keep changing [62]. Because geoelectric fields with different amplitudes and directions can introduce significantly different GICs into the same system, the optimal BD placement can be very different. The assumption of a constant geoelectric field significantly limits the practicality of previous studies. In the TVGF scenario, we incorporate a TVGF to provide better and more realistic solutions.

We use time series to facilitate the time-domain analysis of GMDs' impacts in OBP. Given a TVGF $\vec{\boldsymbol{E}}_\Psi = \{\vec{\boldsymbol{E}}_\psi, \psi \in \Psi\}$, we quantify the overall impact of GMDs by the average GMDs' impact during period Ψ. In this chapter we use Ψ and ψ to represent time so that we can reserve T and t for transformers. Let the time series start at ψ_0 and end at $\psi_{|\Psi|}$, so there are $|\Psi|$ time intervals in the time series. Denote $Q_t^{loss}(\psi)$ the reactive power loss at transformer t at a particular time ψ. Because the damages on transformers are highly correlated to the lasting time of GMDs, we use the time-weighted average of $Q_t^{loss}(\psi)$ to measure the GMDs' impact on a single transformer t for the whole period Ψ, denoted by

$$\begin{aligned} Q_t^{loss}(\Psi) = \langle Q_t^{loss}(\psi)\rangle_{\psi\in\Psi} &= \frac{\sum_{k=0}^{|\Psi|-1} Q_t^{loss}(\psi_k) * (\frac{\psi_{k+1}-\psi_k}{\psi_{|\Psi|}-\psi_0})}{|\Psi|} \\ &= \frac{\sum_{k=0}^{|\Psi|-1} \epsilon_t |I_t^{\mathit{eff}}(\vec{\boldsymbol{E}}_{\psi_k})| * (\frac{\psi_{k+1}-\psi_k}{\psi_{|\Psi|}-\psi_0})}{|\Psi|}, \end{aligned} \tag{5.6}$$

where $\frac{\psi_{k+1}-\psi_k}{\psi_{|\Psi|}-\psi_0}$ is the weight at time ψ_k, which represents the percentage-wise duration of geoelectric field being $\vec{\boldsymbol{E}}_\psi$. Similarly we can calculate the time-weighted average per-phase effective GIC.

$$I_t^{\mathit{eff}}(\Psi) = \langle I_t^{\mathit{eff}}(\psi)\rangle_{\psi\in\Psi} = \frac{\sum_{k=0}^{|\Psi|-1} I_t^{\mathit{eff}}(\psi_k) * (\frac{\psi_{k+1}-\psi_k}{\psi_{|\Psi|}-\psi_0})}{|\Psi|}. \tag{5.7}$$

Similar to (5.1) and (5.2), the overall impact of GMDs on the whole power system can be quantified by the sum of $Q_t^{loss}(\Psi)$ over all transformers or by the maximum of $Q_t^{loss}(\Psi)$ over all transformers, as defined by

$$Q_{total}^{loss}(\Psi) = \sum\nolimits_{t=1}^{N_T} Q_t^{loss}(\Psi) = \sum\nolimits_{t=1}^{N_T} \langle Q_t^{loss}(\psi)\rangle_{\psi\in\Psi}, \tag{5.8}$$

$$Q_{max}^{loss}(\Psi) = \max_t Q_t^{loss}(\Psi) = \max_t \langle Q_t^{loss}(\psi)\rangle_{\psi\in\Psi}, \tag{5.9}$$

where $I_t^{\mathit{eff}}(\vec{\boldsymbol{E}}_\psi)$ is the effective current injection at transformer t with constant geoelectric field $\vec{\boldsymbol{E}}_\psi$ and $\langle Q_t^{loss}(\psi)\rangle_{\psi\in\Psi}$ is the time-weighted average of $Q_t^{loss}(\psi)$ in (5.6).

The $Q_{total}^{loss}(\Psi)$ metric focuses on the average impact of a GMD event. In contrast, $Q_{max}^{loss}(\Psi)$ focuses on evaluating the worst damage a transformer takes during Ψ. We use (5.8) to illustrate the formulation of OBP problem in TVGF scenario:

$$\begin{aligned}
\min_X \quad & Q_{total}^{loss}(\Psi) = \sum\nolimits_{t=1}^{N_T} \langle \epsilon_t | I_t^{\mathit{eff}}(\vec{\boldsymbol{E}}_\psi)|\rangle_{\psi\in\Psi} \\
s.t. \quad & \boldsymbol{H}\vec{\boldsymbol{E}} = \boldsymbol{G}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}), \\
& \boldsymbol{I}_t^{\mathit{eff}}(\boldsymbol{X}) = \boldsymbol{\Phi}(\boldsymbol{X})\boldsymbol{V}(\boldsymbol{X}), \\
& \sum\nolimits_{t=1}^{N_T} x_t \le N, \\
& I_t^{\mathit{eff}}(\Psi) \le \mathbf{I}_{max}, \forall t.
\end{aligned} \tag{5.10}$$

5.3 Solution Methods

As discussed in Section 5.1, all previous studies come with some limitations. They are either not accurate or not scalable. More importantly, they cannot solve OBP problems with realistic constraints, such as limiting the BD

placement impact on other interconnected systems or using a TVGF. In this chapter, we develop an SA method to tackle these issues. We will first outline our SA method in detail and use it to solve the basic OBP problem using a constant geoelectric field. Then, we extend our SA method to cover the BBS and TVGF scenarios.

5.3.1 Simulated Annealing Algorithm

SA is a generic stochastic algorithmic approach for finding the global optimum of an optimization problem that may have many local optima [39, 40]. It can handle cost functions with arbitrary degrees of nonlinearities, discontinuities, and stochasticity and arbitrary boundary conditions and constraints. What is more, it statistically guarantees global optimal given enough running time [64]. SA emulates the physical cooling process of a solid in a heat bath. If this solid is cooled down slowly enough, its structure will eventually be "frozen" at a minimum energy configuration. SA uses this strategy to search for the optimal solution. The difficulty with traditional optimization algorithms for solving the OBP problem is that while many methods can easily find a local optimum, they fall short of finding the global optimum. SA uses two techniques in its searching strategy. First, it uses the "metropolis algorithm". In the searching process, SA probabilistically accepts worse solutions which allows it to jump out of the basin that contains local optimum and explore a larger solution space. As a result, SA can have a more extensive search and potentially find the global optimum. Second, SA interprets the slow cooling process as a slow decrease in the probability of accepting worse solutions. Our SA algorithm is described in Algorithm 2. In the next subsection, we will discuss the detailed design of our SA method.

5.3.2 SA Design for OBP Problems

It is worth mentioning that the configuration of SA is crucial and can solely determine the effectiveness and efficiency of an SA algorithm. In this chapter, we evaluated several different SA designs, compared their solution quality and running time. We chose the best one to solve OBP problems. When there is a trade-off between solution quality and running time, we select the

Algorithm 2: Simulated Annealing Algorithm for OBP

```
Set initial temperature T = T_0;
Generate a random initial BD placement solution s = s_0;
Calculate the objective function E(s) in the OBP problem;
for k = 0 to MaxIterations do
    Pick a neighbor, s_new = neighbor(s);
    if s_new violates OBP constraints then
        Set E(s_new) = ∞
    else
        Calculate E(s_new) using (4.6)-(4.22);
    end
    Update temperature T = temperature(T_0, k);
    if E(s_new) < E(s) then
        Accept new solution s = s_new;
    else
        if P(s, s_new, T) ≥ random(0, 1) then
            Accept new solution s = s_new;
        end
    end
end
if E(s) = ∞ then
    There is no feasible solution to the OBP problem;
end
```

saturation point where the improvement of solution quality approaches zero with longer running time. We will also list some alternative SA configurations we considered.

5.3.2.1 Solution Representation and State Space

In Algorithm 2, a *state* s corresponds to a BD placement solution, which is represented by a binary vector $\boldsymbol{X}$. Given N_T transformers that can be blocked and N BDs to place, the solution space contains $\binom{N_T}{N}$ possible states.

5.3.2.2 Objective Function

$E(s)$ in Algorithm 2 represents the objective function in Equation (5.3). It is analogous to the internal energy of the physical system in the respective state.

5.3.2.3 Neighborhood Structure and Neighbor Generator

The search space of the OBP problem can be modeled as a search graph with verticals as all states and edges as candidate moves defined by *neighbor()*. The design of *neighbor()* is extremely critical to the performance of the SA algorithm. As shown in Algorithm 2, SA first generates a random blocking solution s_0. At each iteration, a neighbor state s_{new} is generated using *neighbor()* based on current state s. The SA algorithm probabilistically decides whether to move the system to a neighbor state s_{new} or to stay at state s. This step is repeated until the system reaches a state that is good enough or until a given computation budget has been exhausted. One essential requirement here is that the *neighbor()* function must provide a sufficiently short path on the search graph from any arbitrary state to any state that may represent the global optimum. Intuitively, it means that the SA algorithm can move to the global optimal state, from any state, quickly.

In our design, we also use the heuristic in the Metropolis-Hastings algorithm [66], which has been proved to be effective. This heuristic tends to generate s_{new} that has an objective function value similar to that of current state s. More specifically, we order the transformers' indices in the binary vector $\boldsymbol{X}$ in such a way that the transformers in the same substation have consecutive indices and transformers that are geographically close to each other also have close indices. Then *neighbor()* swaps 1 and 0 that are next to each other in the binary vector $\boldsymbol{X}$. Because of the local blocking effect [37], swapping two adjacent transformers' blocking decisions is expected to have a modest effect on the objective function value, which is measured by reactive power losses. This design allows us to move from any state to any other state in $N_T(N_T - 1)$ steps. In an alternative approach, we considered generating neighbors by swapping two arbitrary transformers' blocking decisions, which could provide a somewhat shorter path ($N_T - 1$ swaps) from one state to another. However, this design is likely to dramatically change the objective function value because it is equivalent to moving a BD randomly in the system. Experimental results also show that the former design performs better.

5.3.2.4 Initial Temperature and Stopping Criteria

The higher the initial temperature, the more likely the SA algorithm can find a global optimal solution yet the longer it takes to converge. We evaluate solution quality and running time with different initial temperatures. Finally, we choose an initial temperature $T_0 = 1000$. This initial temperature provides very good solution quality and the solution quality improvement saturates at this point.

5.3.2.5 The Annealing Schedule

The annealing schedule defines how temperature T systematically decreases as the searching proceeds. Temperature T is updated as

$$T = \text{temperature}(T_0, k) = T_0 * \beta^k, \quad \beta = 0.95 \tag{5.11}$$

where k is the iteration number. It ends with $T = 0$ at the end of the algorithm. We experimented with different scheduling functions and scaling factors β. With a larger β, the SA algorithm converges slower and will more likely find a better solution. The scheduling function in (5.11) performs the best. We chose 0.95 for β because the solution quality improvement saturates at this point.

5.3.2.6 Acceptance Probability Function

The probability of moving from the current state s to a candidate new state s_{new} is specified by an *Acceptance Probability Function* (APF) $P(s, s_{new}, T)$:

$$P(s, s_{new}, T) = \begin{cases} e^{\frac{E(s)-E(s_{new})}{T}} & , \text{if } E(s_{new}) > E(s) \\ 1 & , \text{if } E(s_{new}) \leq E(s) \end{cases}. \tag{5.12}$$

The APF function $P(s, s_{new}, T)$ is positive when $E(s_{new}) > E(s)$ so that it can prevent the algorithm from getting stuck at a local minimum. When $E(s_{new}) > E(s)$, $P(s, s_{new}, T)$ decreases as T decreases, which means that the SA algorithm will be less likely to accept moves to worse solutions as it cools down.

5.3.3 SA for the BBS Scenario

To solve the BBS scenario using the SA method, we add an additional step in Algorithm 2 to check if the BBS constraint is violated. If any BBS constraint is violated, we set the objective function value to be infinite and continue. Eventually, we will either get the optimal solution that does not violate the BBS constraint or an infinite large objective function value, which indicates that there is no solution to the OBP problem.

5.3.4 SA for the TVGF Scenario

The computation time increases significantly when solving the TVGF scenario using the SA method as it takes a much longer time to evaluate the objective function values. We parallelize the computation in our SA method and the computation time for the TVGF scenario in (5.10) is reduced down to almost the same as that for the basic OBP problem. With this parallelization technique, we can easily extend our current SA method and solve the OBP problem for a probabilistic time-varying geoelectric field with Monte Carlo simulation.

5.4 Results

To validate our method, we use four GIC test systems of various sizes [24,30]. The test systems are summarized in Table 5.1. The 6-bus, 20-bus and large test system were also used in [37] and [61]. The other one is a newly created test case based on the UIUC 150-bus system. The PowerWorld simulator [60] was used to process the data.

We ran the tests on a multi-core machine, featuring a 3.10 GHz Intel Core i5 processor and 16 GB memory. First, we solve OBP problems under the constant geoelectric field (CGF) scenario. Specifically, we took the common approach of using an eastward constant $\vec{\boldsymbol{E}}$, where $|\vec{\boldsymbol{E}}| = 1$. We then extend our studies and solve OBP problems for CGFs of several different directions. We evaluate the solution quality and the running time of our method and compare them with previous methods. Secondly, we solve the OBP prob-

lem under the BBS and TVGF scenarios and study the behaviors of BD placement.

Table 5.1: Test Benchmark

Benchmarks	Bus No.	Substation No.	Transformer No.
Case NERC 6-bus	6	3	3
Case 20-bus	19	8	15
Case 150-bus	150	20	46
Case large	1875	865	756

5.4.1 Numerical Results for the CGF Scenario

In order to demonstrate the optimality and computational efficiency of our SA method, we compare it with two greedy search approaches, the relaxed SOCP method, the BC method and the exhaustive search (ES) method. We use the problem formulation in (5.3). Since all previous methods cannot solve OBP problems with the maximum effective GIC constraint in (5.3), we will first solve the basic OBP problem without GIC constraint as that in [37, 61]. In Section 5.4.1.3, we will show the results for solving the OBP problem with a maximum GIC constraint using our SA method. Although the solutions from the ES method are guaranteed to be optimal, enumerating all the solutions is not practical for a realistically large power system.

5.4.1.1 Basic CGF Scenario

We compare all methods on the 6-bus, 20-bus and 150-bus systems. As mentioned earlier, greedy search based methods can also solve the CGF scenario. We will evaluate two variations of greedy search based methods: Static Greedy Search (SGS) and Iterative Greedy Search (IGS). Assume that we will install N BDs. In SGS, we calculate GMD damages for all transformers before placing any BD and select the top N transformers with the highest GMD damages to block. In IGS, we iteratively select transformers to block until we find all N transformers to block. At each iteration, we calculate the GMD damages using the current BD placement and select the transformer that has the highest GMD damage as the next one to block. The difference between SGS and IGS is that SGS calculates GMD damages for only

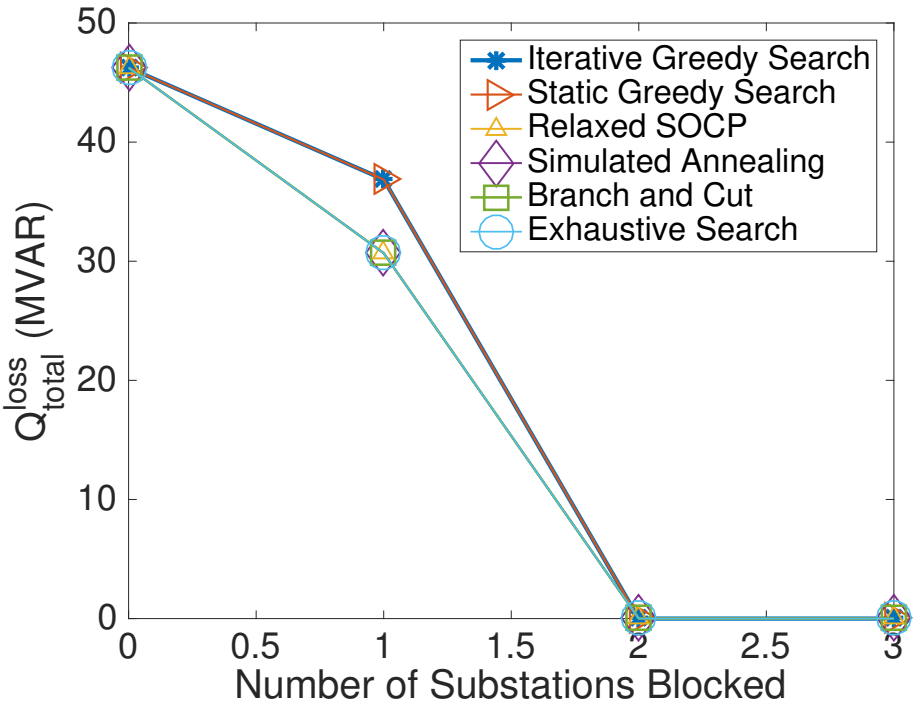


Figure 5.1: Comparison of all methods using the 6-bus system (E=1V/km, East)

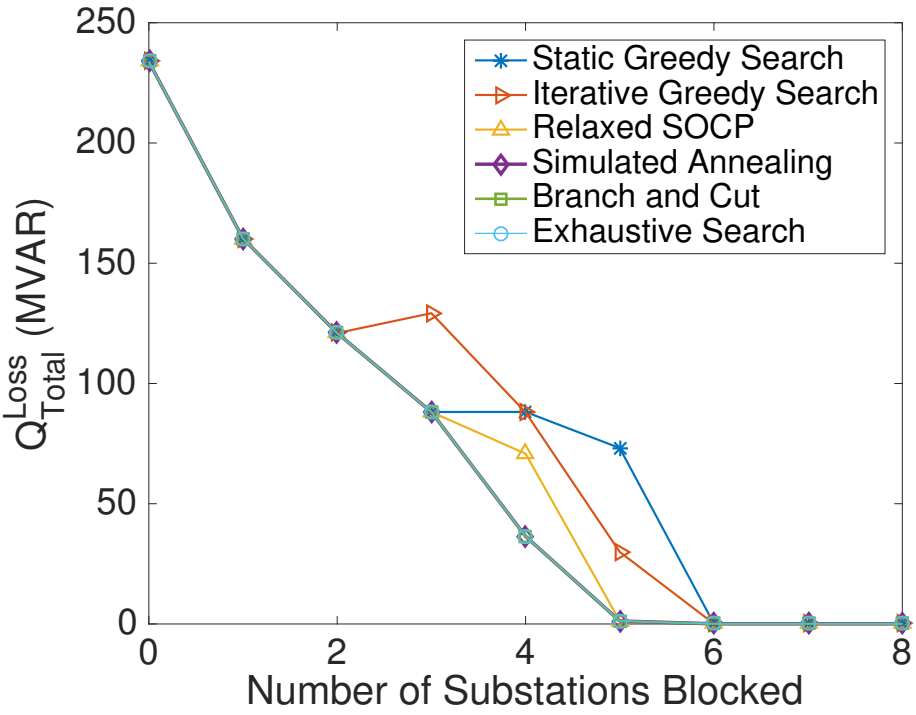


Figure 5.2: Comparison of all methods using the 20-bus system (E=1V/km, East)

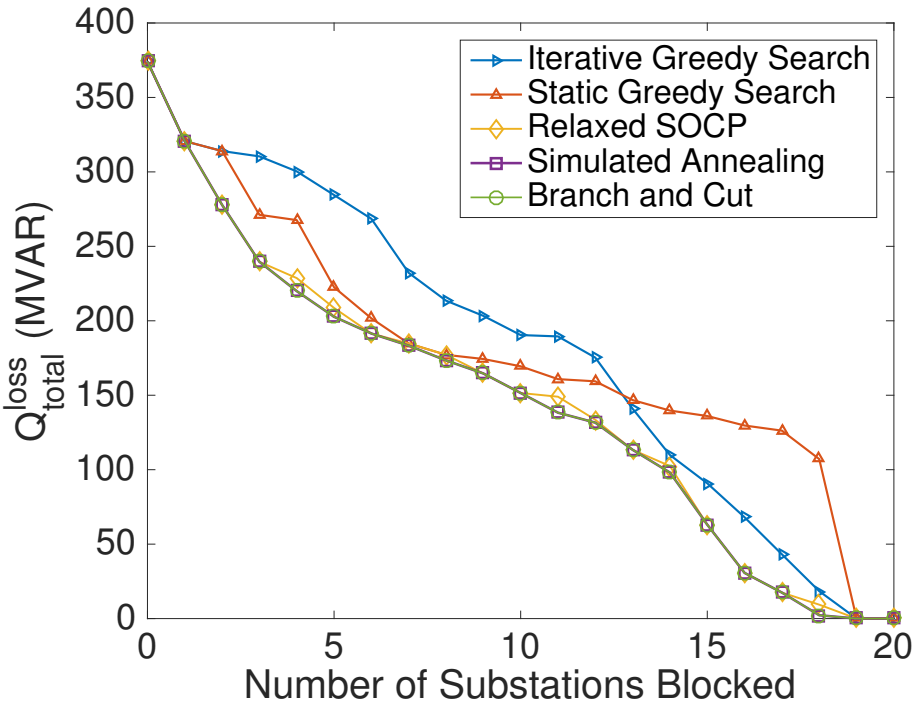


Figure 5.3: Comparison of all methods using the 150-bus system (E=1V/km, East)

one time while IGS calculates GMD damages N times. Because the relaxed SOCP method can only provide blocking solutions at substation level, to compare with this method, we show the results of blocking by substation in this section.

The 6-bus system [24] is a small system with 3 transformers (T_1-T_3) and 3 substations. Experimental results show that relaxed SOCP, SA, BC and ES can find the optimal BD placement solutions for any number of BDs while SGS and IGS cannot. As shown in Fig. 5.1, Q^{loss}_{total} is 46.24 MVar before placing any BD. The optimal solution for $N = 1$ is to block T_3 and Q^{loss}_{total} is reduced to 30.68 MVar. Both SGS and IGS select T_1 to block because it has the largest Q^{loss} before installing any BDs. The optimal solution for $N = 2$ is to block T_1 and T_3, and accordingly Q^{loss}_{total} is reduced to 0.

The 20-bus system [30] is medium-sized, with 15 transformers and 8 substations. As shown in Fig. 5.2, for any given number of substations to block, the BC and SA methods can always find the optimal solutions, while the SGS, IGS and relaxed SOCP method may not. For example, the relaxed SOCP method cannot find the optimal solution when blocking 4 substations and it can only reduce the Q^{loss}_{total} down to 70.7 MVar while the optimal solution is

36.45 MVar. Similarly, the SGS method cannot find optimal solutions for blocking 4 and 5 substations and the IGS method cannot find optimal solutions for blocking 3, 4 and 5 substations. The IGS method fails to find the optimal solution for blocking 4 substations because there is a local minimum trap. The optimal solution for blocking 2 substations is $\{S4, S6\}$, leading to the next substation of highest Q^{loss} to $S8$. However, due to local blocking effect, blocking $S8$ will increase Q^{loss} at $S3$ and $S5$. Actually, the the optimal blocking solution for 3 substations is $\{S1, S4, S6\}$. The SGS method also fails to find optimal solutions for similar reasons. These two methods cannot assess global optimality of the whole system when solving the OBP problems. The SA method performs significantly better than other methods here. In addition, we observe that the marginal gain of blocking each additional substation could vary, confirming the combinatorial nature of this problem.

The 150-bus system is medium-sized, with 46 transformers and 20 substations. The benchmark solution by ES is not available here due to the size of the system. As shown in Fig. 5.3, for any given number of substations to block, the BC and SA methods can always achieve the best solutions compared to all other methods, corroborating their effectiveness in attaining the optimal solutions.

5.4.1.2 CGFs in Multiple Directions

The 20-bus system [30] has been tested using CGFs in multiple directions for minimizing both Q^{loss}_{total} in (5.1) and Q^{loss}_{max} in (5.2). We will present results for using two representative CGF directions: eastward and northward. Here, we use transformer as the basic blocking unit. Since the relaxed SOCP method cannot provide blocking solutions at the transformer level, it is not included for comparison. Figures 5.4 and 5.5 show the results for eastward and northward CGFs using Q^{loss}_{total}. Figures 5.6 and 5.7 show the results for eastward and northward CGFs using Q^{loss}_{max}. In all four cases, the magnitude of GF is 1 V/km

As shown in Fig. 5.4 - 5.7, both BC and SA methods can find optimal solutions while SGS and IGS methods suffer from low solution quality. Experimental results also show that BD placement by transformer is sensitive to the directions of GFs. Specifically, the minimal Q^{loss}_{total} and Q^{loss}_{max} are differ-

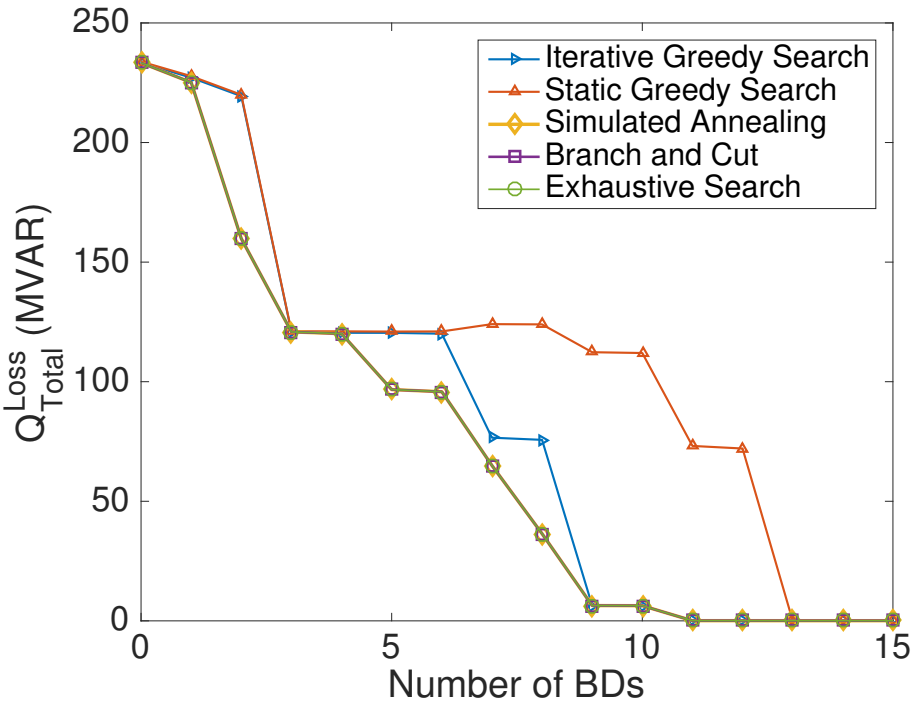


Figure 5.4: Comparison of all methods using the 20-bus system and Q^{loss}_{total} for 1V/km, eastward CGF (block by transformer)

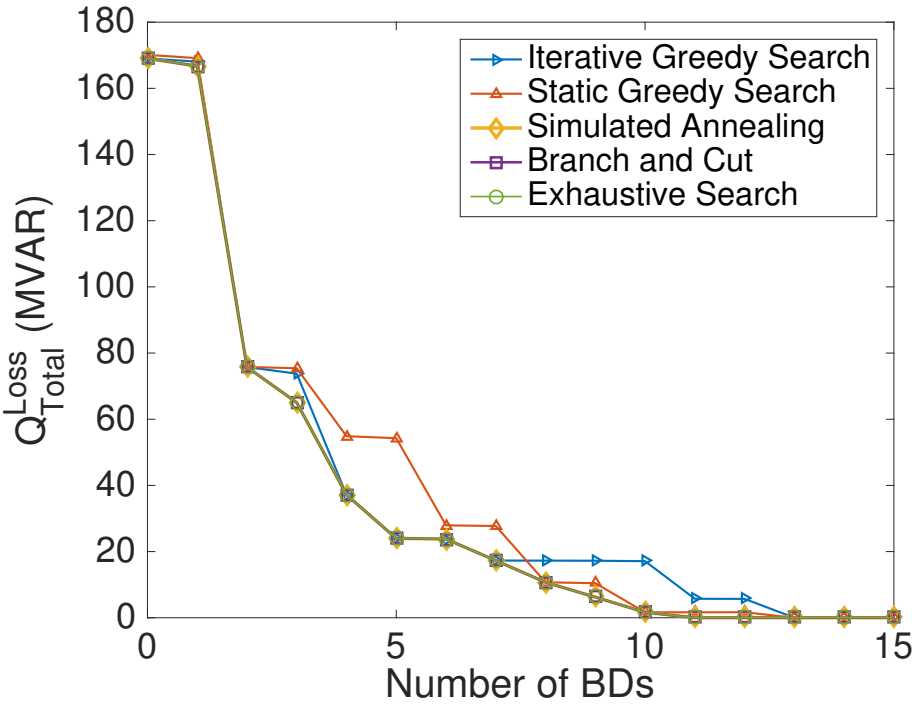


Figure 5.5: Comparison of all methods using the 20-bus system and Q^{loss}_{total} for 1V/km, northward CGF (block by transformer)

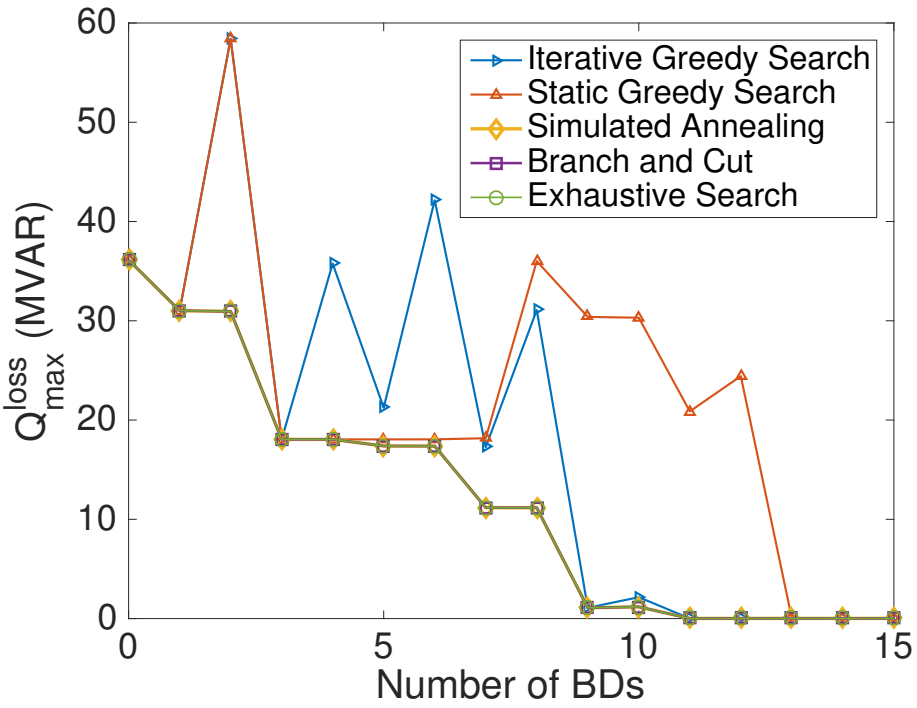


Figure 5.6: Comparison of all methods using the 20-bus system and Q^{loss}_{max} for 1V/km, eastward CGF (block by transformer)

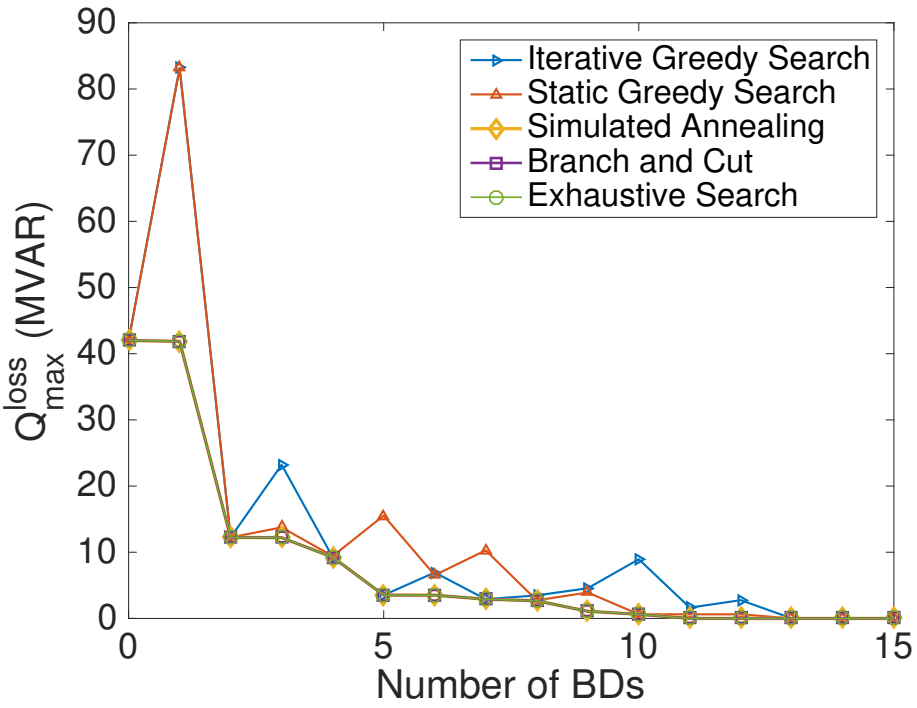


Figure 5.7: Comparison of all methods using the 20-bus system and Q^{loss}_{max} for 1V/km, northward CGF (block by transformer)

ent between using eastward GF and northward GF for some number of BDs. By comparing Fig. 5.4 with Fig. 5.5, we observe that, on average, the 20-bus system is more sensitive to eastward GF at higher reactive power losses. However, between Fig. 5.6 and Fig. 5.7, the Q_{max}^{loss} levels are similar under both GF directions. More importantly, the optimal blocking solutions are very different between using these two GFs. As a result, it is not sufficient to measure the GMDs impact or make BD placement decisions using a single CGF direction.

In addition, the optimal BD placement solutions are not always the same between minimizing Q_{total}^{loss} and Q_{max}^{loss} even under fixed GF directions. This implies the objective function is critical in determining the BD placement solutions. More in-depth study is needed to understand the best objective function to use. Again, greedy search based methods have proven ineffective in solving the OBP problem. Neither SGS or IGS can provide satisfactory solutions due to the local blocking effect. This is extremely important when using Q_{max}^{loss} as objective function. For example, in Fig. 5.7, when blocking 1 transformer, the SGS method would redistribute the GIC to nearby transformers in the same substation and accordingly increase Q_{max}^{loss}.

These observations speak for the importance of the ensuing TVGF scenario studies, which can account for the average impact of GMDs in practical systems.

5.4.1.3 CGF Scenario with Max GIC Constraints

The new feature of our proposed SA method is to solve OBP problems with different max GIC constraints and compare the results. In this experiment, we use the 20-bus system and a 2 V/km eastward CGF. Figure 5.8 shows the minimal Q_{total}^{loss} under different GIC constraints. The incomplete lines indicate that there are no feasible solutions at lower max GIC limits and smaller number of BDs. For example, before installing any BDs, the maximum per-phase effective GIC is 137 A. When the GIC limit is 75 A, there is no feasible solution with less than 3 BDs. In addition, the GIC constraint affects the minimal Q_{total}^{loss} that can be obtained. The larger the max GIC limit, the smaller the objective function in the OBP can be. Further investigation shows that the BD placement solutions are also different depending on the GIC limits. For example, when placing 2 BDs, the blocking solution

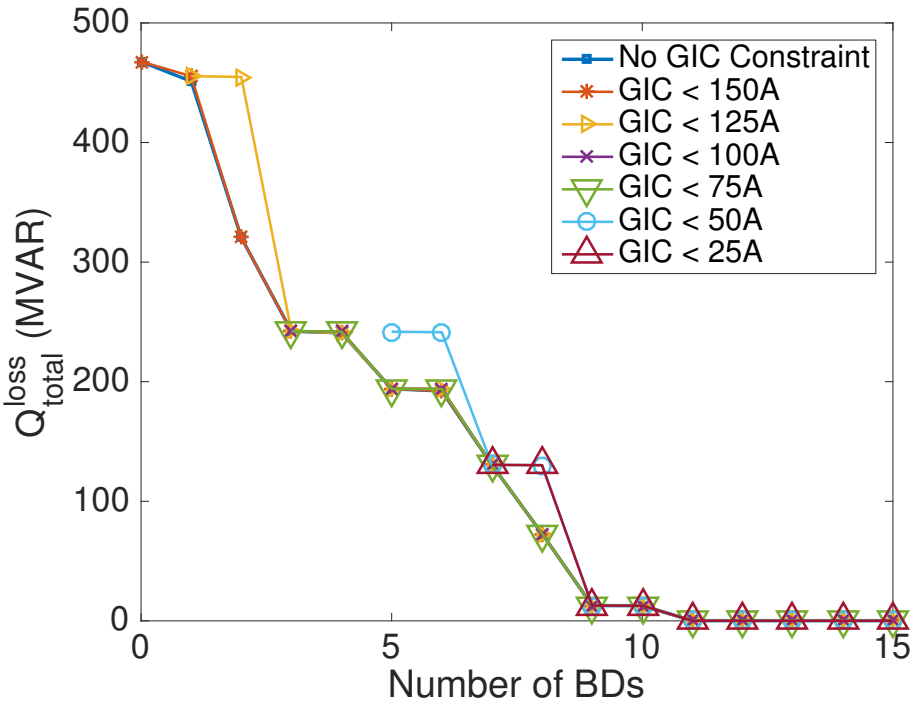


Figure 5.8: BD placement under different GIC constraints (E=2V/km, East)

is $\{T_1, T_7\}$ when the GIC limit is 125 A, and is $\{T_8, T_9\}$ when the GIC limit is 150 A or higher. Therefore, the ability to include the Max GIC limit makes our proposed SA method more suitable for practical system studies.

5.4.1.4 CGF Scenario for a Large System

In order to demonstrate the effectiveness of our method in real-world applications, we test our method on a system located in the upper Midwestern United States, with neighboring systems built from equivalents of the Eastern Interconnection. In total, there are 756 active transformers in this system, with 78 high voltage transformers. The ES method cannot be used to solve the OBP problem here because the size of the power system is too large. For example, there are $1.58e^{22}$ possible solutions when placing only 10 BDs. For the BC method, we generously restrict the maximum running time to be 48 hours and pick the best solution found within this time limit. The running time of the BC method depends heavily on the search graph of the OBP problem and the branching heuristics. For example, if the BC method can quickly prune several "nodes" in the early phases, it can converge faster [38].

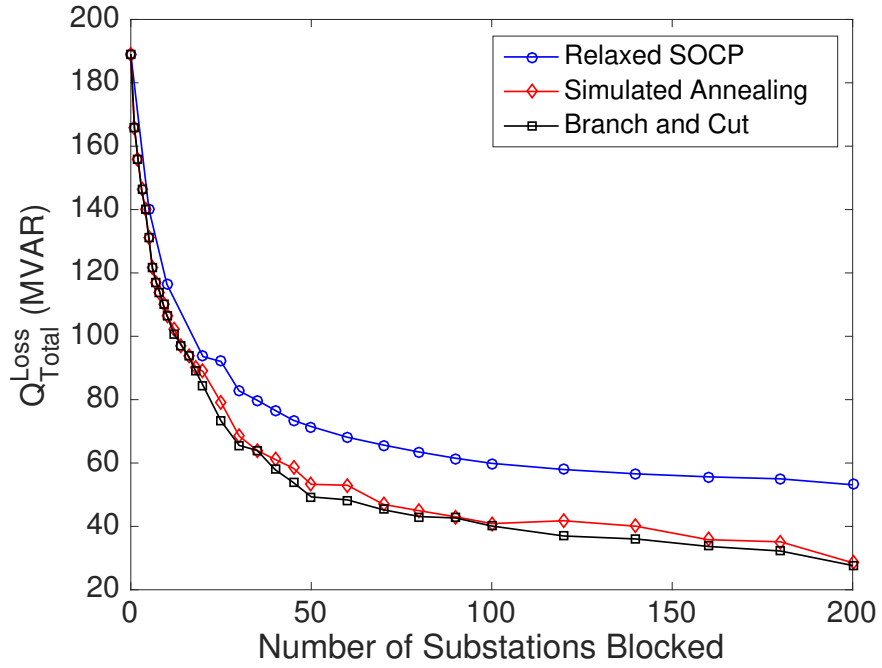


Figure 5.9: BD placement for the large-scale system (block by substation)

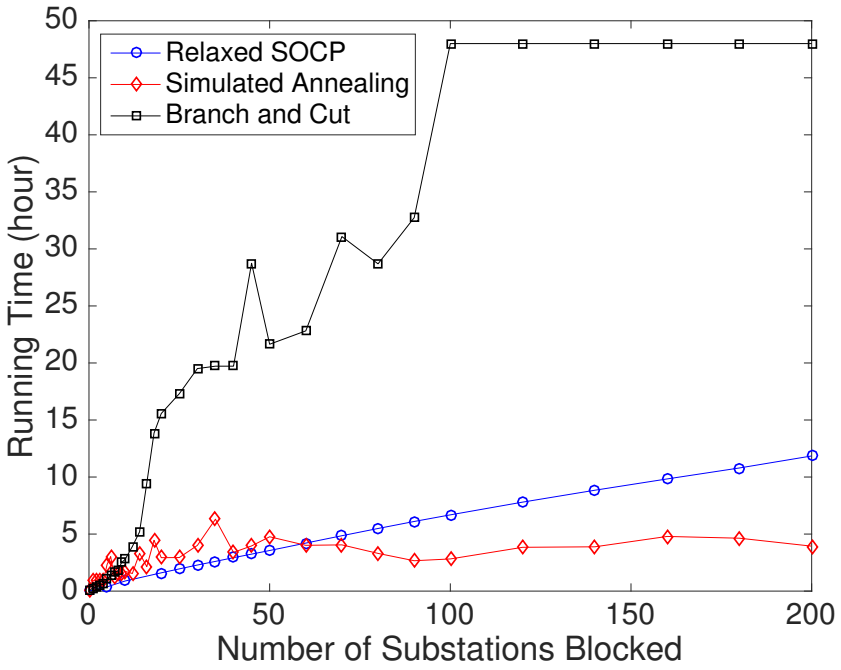


Figure 5.10: Running time of OBP problem for the large-scale system (block by substation)

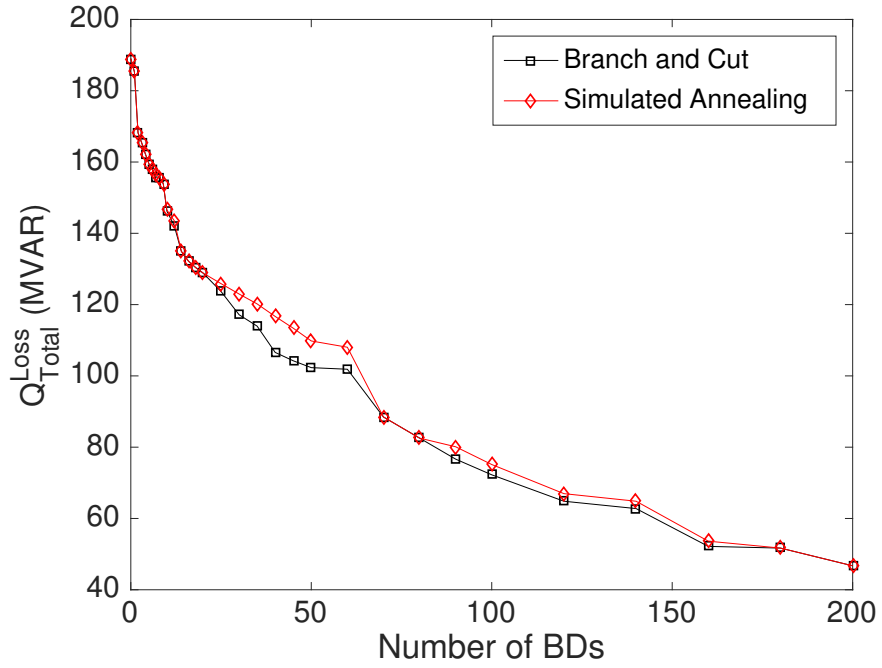


Figure 5.11: BD placement for the large-scale system (block by transformer)

In practice, there are some converging and memory consumption issues for the BC method. We notice the BC method can handle cases with up to 40 BDs without any converging issues. Convergence issues start to appear when there are more than 45 BDs. Multiple runs with different BC options, including branching strategies, relaxation options and many others, are necessary to achieve convergence at higher number of BDs [38,67]. When blocking 100 or more transformers, we observe that BC methods cannot converge within 48 hours. We solve the OBP problems for 30 different N values between 0 and 200. The solution quality, measured by the total reactive power losses, and running time for all methods (block by substation) are shown in Fig. 5.9 and Fig. 5.10.

Overall, the SA method provides much better mitigation results than the relaxed SOCP method and is still comparable to the BC method when N is large. However, the BC method is not scalable as N increases. Clearly, the SA method is much faster than the BC method and the relaxed SOCP method when N is large. In terms of the achieved Q^{loss}_{total} values, the solutions provided by the relaxed SOCP method are 32.31% worse than those generated by the SA method. On the other hand, the solutions provided by the

BC method are only 4.75% better than those provided by the SA method. On average, the SA method is 7.4 times faster than the BC method and is 1.3 times faster than the relaxed SOCP method. When N is large (e.g. $N = 200$), the SA method can be 12.2 times faster than the BC method. The running time of the SA method fluctuates a lot because of its random nature. Also, it has to take a certain number of iterations to "cool down" no matter how small N is. Hence, it is slower than the relaxed SOCP method when N is relatively small. Figure 5.11 shows the solutions of blocking by transformer using the BC and SA methods. Data from the relaxed SOCP method is absent as it cannot provide BD placement solutions by transformer. The results show a similar pattern as in Fig. 5.9. The SA method is very effective and it provides high-quality solutions. Solution quality from SA method only falls short by 5.0% when compared to the BC method.

5.4.2 Numerical Results for the BBS Scenario

To test the BBS scenario, we use the 20-bus system and divide it into two sub-systems based on their geographical locations [30]. System A contains substations $1, 2, 4$, and system B contains the rest of the substations. Figure 5.12 shows the total reactive power losses in system A using the formulation in (5.5) with different BBS constraints, quantified by α. The α value controls how much negative impact is allowed on interconnected systems. Experimental results show that placing BDs in one system can indeed introduce negative impact on neighboring interconnected systems and the BBS constraints do change the optimal solutions significantly. The stricter the BBS constraint, which corresponds to a smaller α value, the less room there is to achieve a better solution for the target system.

As shown in Fig. 5.12, if we do not allow any negative impact on system B ($\alpha = 0$), there is no valid solution for placing any number of BDs (represented by infinite Q_A^{loss}). If we allow 1% negative impact on system B ($\alpha = 1\%$), when placing 1 BD in system A, the optimal solution can only reduce 0.23 MVar reactive power losses, which is far less than the optimal solution that reduces 23.75 MVar losses at no BBS constraint. If we allow 10% negative impacts on system B ($\alpha = 10\%$), the BBS constraints do not change the optimal solutions for placing 1 and 2 BDs, while they significantly affect

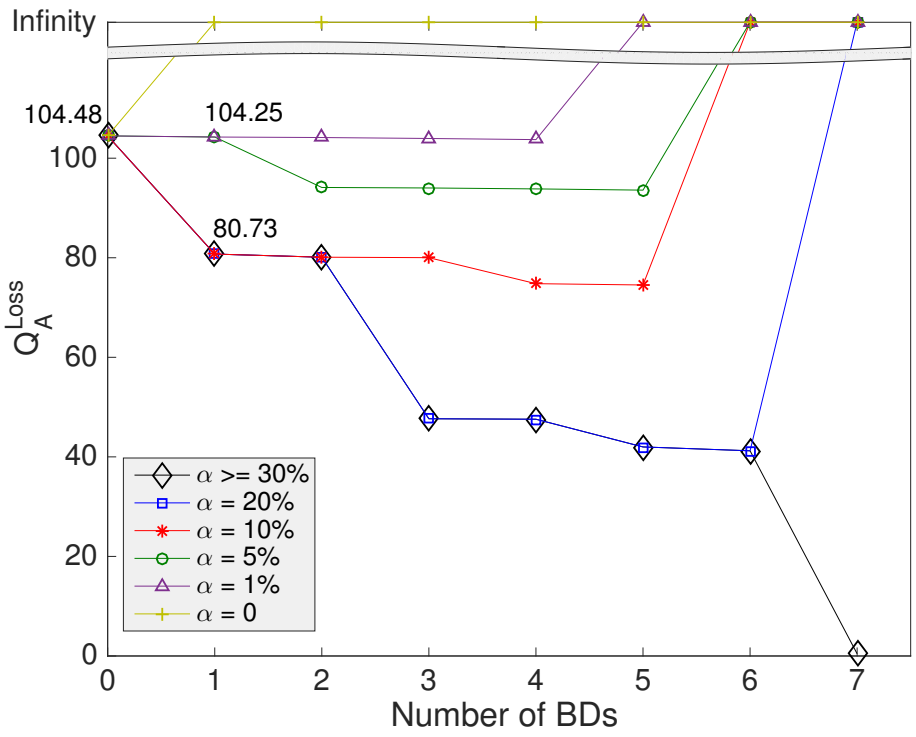


Figure 5.12: BD placement for with different BBS constraints using the 20-bus system

the optimal solutions for placing 3 and 4 BDs. The BBS constraints do not change the optimal solutions if we allow 30% or more negative impact on system B ($\alpha \geq 30\%$), which means that the largest impact on system B from placing BDs in system A is less than 30% of $\sum_{t \in T^B} Q_t^{loss\_base}$. The GIC's local blocking effect is the reason behind this behavior. When a transformer is blocked, the GICs will flow to its nearby transformers. To further understand this behavior, let us take a closer look at the results. Before placing any BD, Q_A^{loss} is 104.48 MVar. Now consider the case of placing 1 BD in system A. When $\alpha = 1\%$, the optimal solution is blocking T_{14} and $Q_A^{loss} = 104.25$ MVar. When $\alpha = 10\%$, the optimal solution is blocking T_1 and $Q_A^{loss} = 80.73$ MVar, which is much better than the solution for $\alpha = 1\%$. Note that transformer T_1 is the only transformer in its substation, while there are three transformers in T_{14}'s substation. Because of GIC's local blocking effect, blocking T_1 redistributes the GICs through T_1 to transformers nearby, which include transformers in system B. Transformer T_{14} is also close to system

B. However, blocking T_{14} redistributes most of the GIC flowing through T_{14} to transformer T_{13} in the same substation, which is still in system A. As a result, compared to blocking T_{14}, blocking T_1 is more beneficial to system A but has larger negative impact on system B. This is why the SA method chooses to block T_{14} when α is small and T_1 when α is large. Also, we can achieve lower reactive power losses with a larger α value.

5.4.3 Numerical Results for the TVGF Scenario

We use the 20-bus system to test SA method in the TVGF scenario. We use the TVGF from the "benchmark" GMD event [62] in NERC TPL-007-2 standard [23] [63]. We also run the same set of experiments using a similar TVGF from the "supplement" GMD event [68]. As the results are similar for both events, we will only present the benchmark GMD event, with the waveform shown in Fig. 5.13. Moreover, Fig. 5.14 presents its magnitude while Fig. 5.15 shows the histogram of the angle. Note that Northward direction cor-

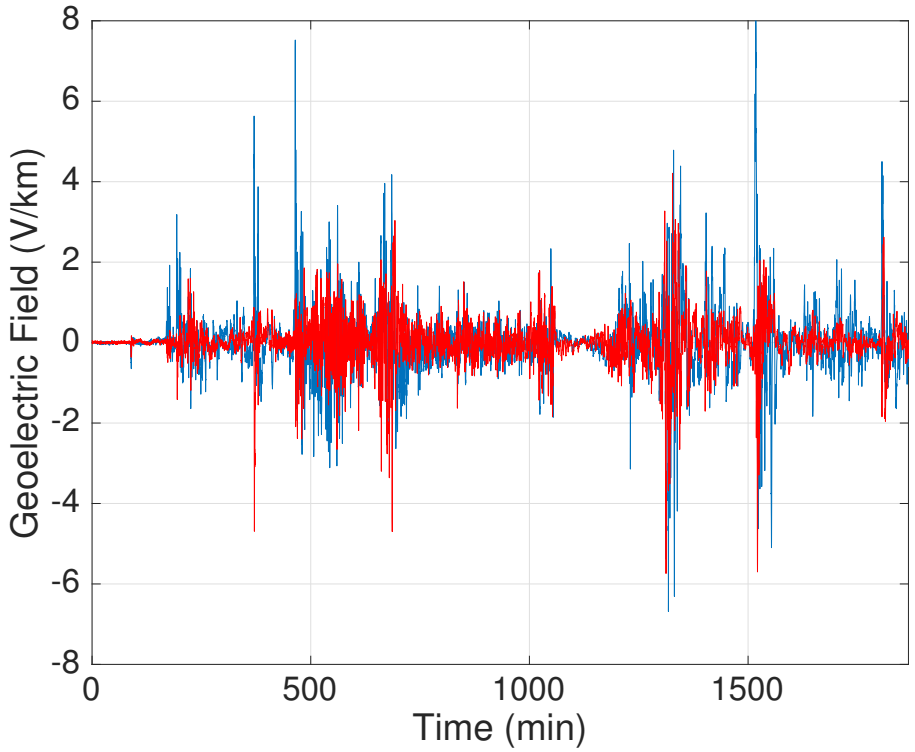


Figure 5.13: Time-varying geoelectric field waveform (red: northward, blue: eastward)

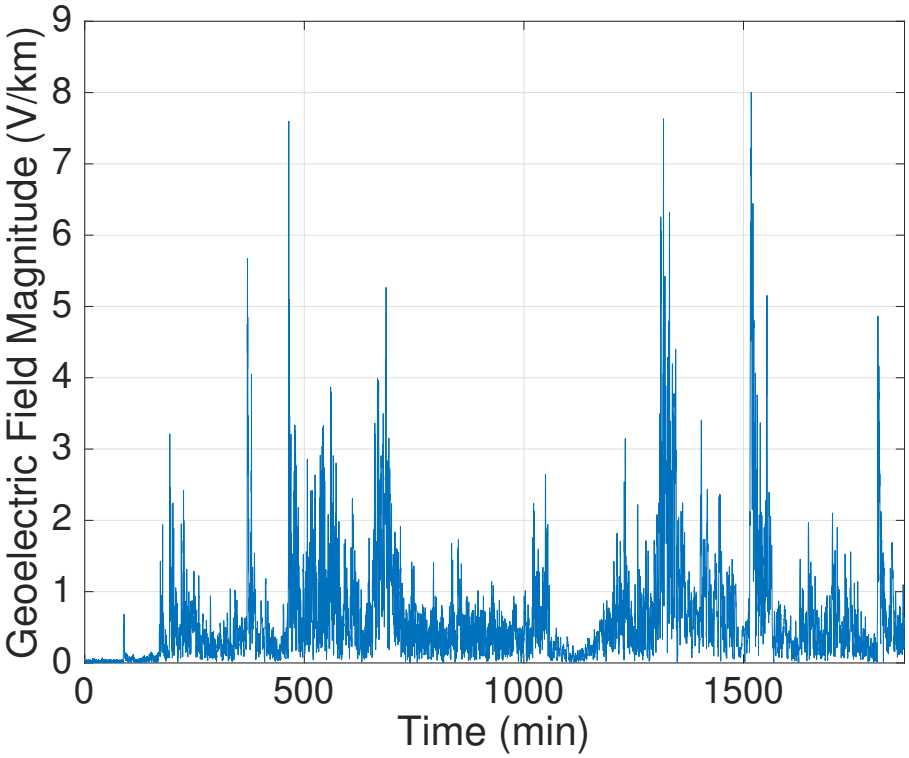


Figure 5.14: Magnitudes of the benchmark TVGF

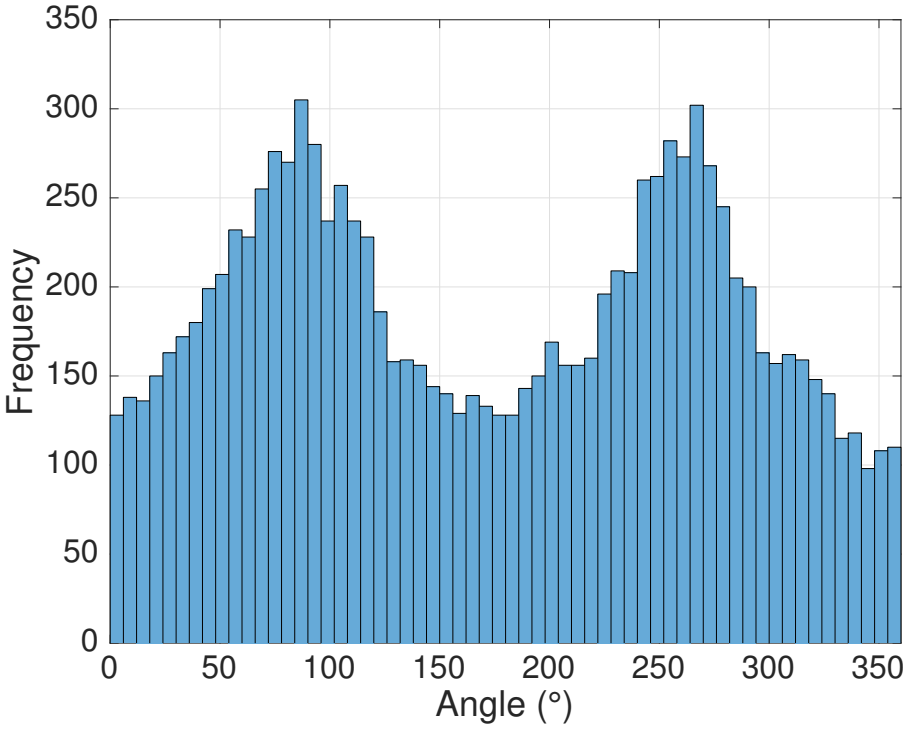


Figure 5.15: Histogram of the angles in the benchmark TVGF

responds to 0° in angle while Eastward to 90°. Hence, the benchmark event is shown to align more with the East-West axis than the North-South one. We use the problem formulation in (5.10) and first limit the time-weighted average per-phase effective GIC at 75 A in Sections 5.4.3.1 and 5.4.3.2. We then study the impact of varying maximum GIC constraints in the TVGF scenarios in Section 5.4.3.3.

5.4.3.1 Comparison of Multiple TVGFs

We consider both $Q_{total}^{loss}(\Psi)$ in (5.8) and $Q_{max}^{loss}(\Psi)$ in (5.9) in this study. To better assess the impact of TVGFs, we also rotate the benchmark event by a range of angles in $[30°, 150°]$ for studying the BD placement. This is because the problem is exactly symmetric with a 180° difference in GF angle.

Figures 5.16 and 5.17 show the BD placement solutions for the benchmark TVGF (0°) and the rotated TVGFs using $Q_{total}^{loss}(\Psi)$ and $Q_{max}^{loss}(\Psi)$. Clearly, transformers and BD placement solutions are sensitive to the directions of the GFs. In Fig. 5.16, for all TVGFs, the minimal $Q_{total}^{loss}(\Psi)$ follows a similar decreasing trend as the number of BDs increases. These results are par-

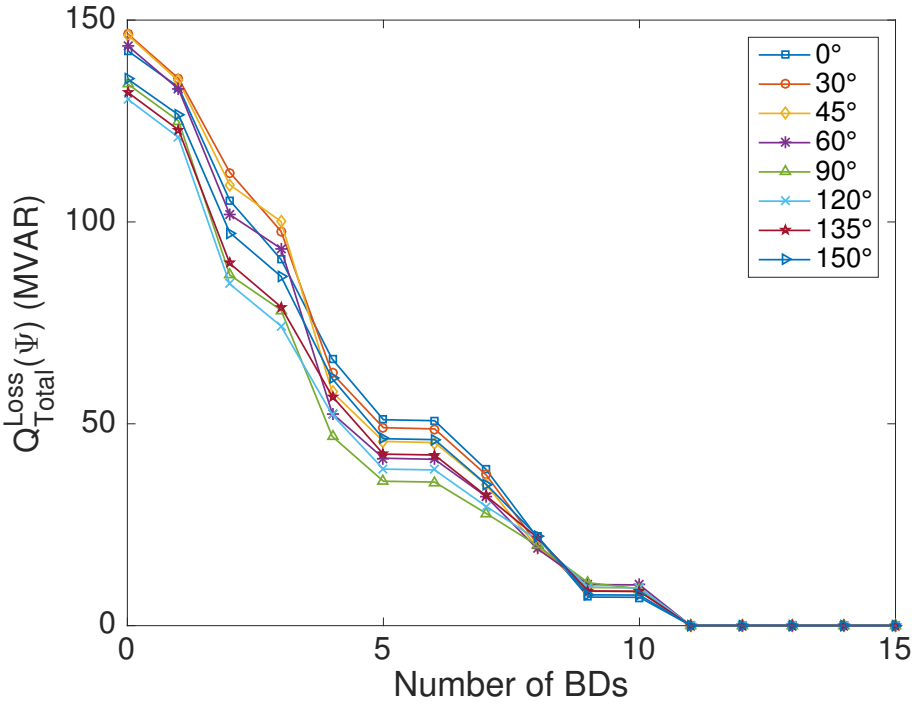


Figure 5.16: BD placement for the TVGF scenario ($Q_{total}^{loss}(\Psi)$)

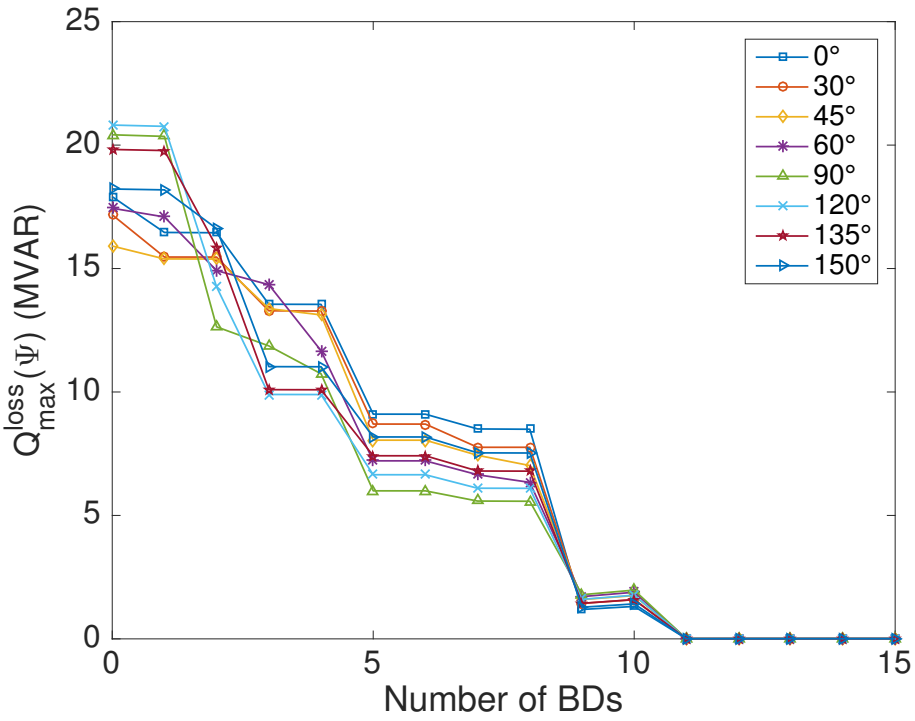


Figure 5.17: BD placement for the TVGF scenario ($Q^{loss}_{max}(\Psi)$)

ticularly helpful for planning stage studies. We also notice that the best BD placement solutions are different depending on rotating angles. For example, when placing 3 BD devices, the optimal solutions for TVGFs with $0°, 15°, 30°, 165°$ rotations are $\{T_1, T_8, T_9\}$, while those are $\{T_1, T_6, T_7\}$ for all other rotations.

As shown in Fig. 5.17, when measuring $Q^{loss}_{max}(\Psi)$, not all TVGFs follow the same decreasing trend as the number of BDs increases. The benefits of placing the first 4 BDs for TVGFs with $0°, 30°, 45°, 90°$ rotations are slightly less significant than those for other TVGFs. Similarly, the most effective BD placement solutions are different. For example, when placing 3 BD devices, the optimal solutions for TVGFs with $15°, 30°, 45°$ rotations are $\{T_1, T_{12}, T_{13}\}$, while those are $\{T_1, T_6, T_7\}$ for all other rotations. Because rotating TVGFs actually shifts the distribution of GFs' angles by a certain degree, these results indicate that the distribution of the angles of GFs plays a key role in assessing the impact of GMD and the blocking behaviors of the system.

5.4.3.2 Comparison between TVGF and CGF

We also compared the solutions for the OBP problem using TVGF and CGF. To make a fair comparison, we selected several constant GFs based on the benchmark TVGF. First, we use the average magnitude (0.6526 V/km) and average angle (4.48°) of the benchmark TVGF to compose one of the constant GFs. We also choose 4 constant GFs in 4 directions (North, East, Northeast and Northwest) using the same magnitude (0.6526 V/km). In fact, we have run experiments using CGFs in additional directions. Since the observations are very similar, we only pick the 4 representative directions. In the CGF scenario, the reactive power losses is the same for all ψ in the period Ψ, which means that $Q^{loss}_{total}(\Psi) = Q^{loss}_{total}(\psi)$ and $Q^{loss}_{max}(\Psi) = Q^{loss}_{max}(\psi)$. Therefore, with a CGF, it is equivalent to solving the basic OBP problem in (5.3) using the total and maximum reactive power losses in (5.1) and (5.2).

Figures 5.18 and 5.19 show the comparisons of the TVGF scenario and various CGF scenarios. Experimental results show that the behavior of the power system under TVGF cannot be approximated by using any single CGFs. Experimental results again confirm that GMDs impact on transformers and BD placement solutions are sensitive to the directions of GFs.

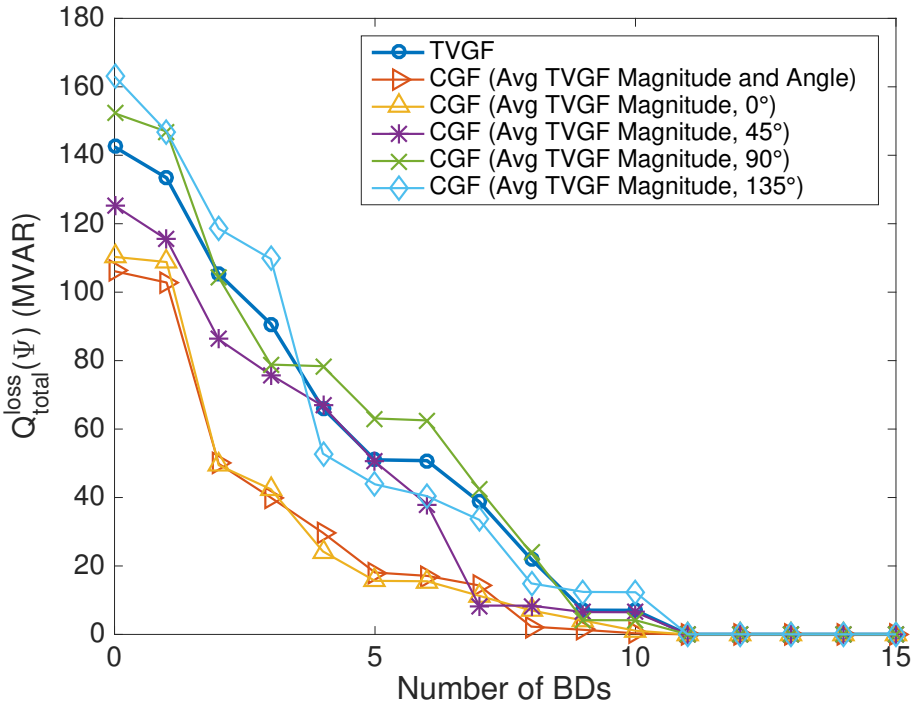


Figure 5.18: BD placement for the TVGF and CGF scenarios ($Q^{loss}_{total}(\Psi)$)

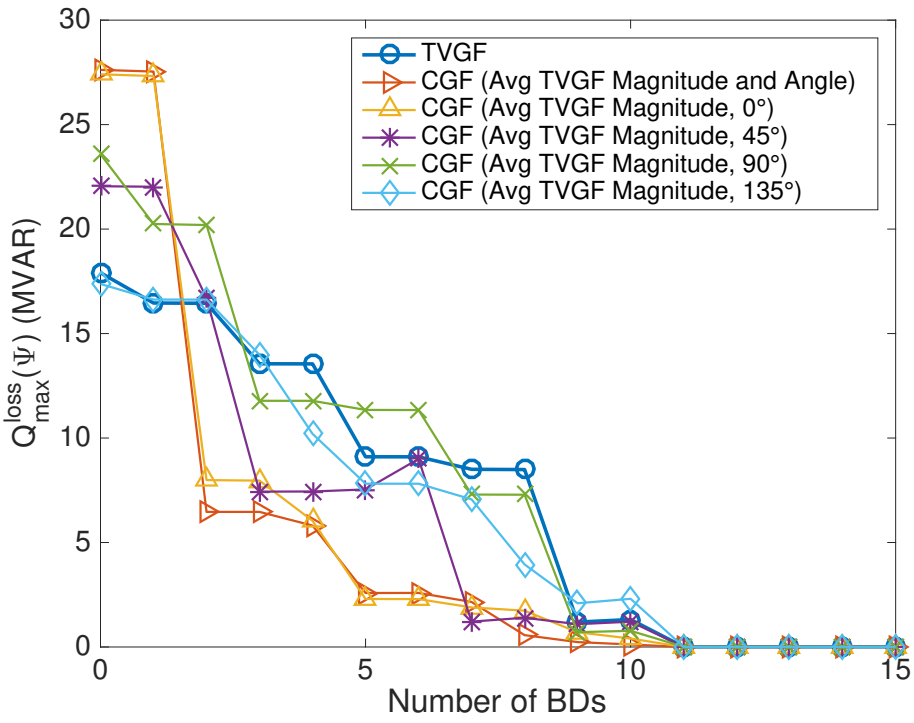


Figure 5.19: BD placement for the TVGF and CGF scenarios ($Q^{loss}_{max}(\Psi)$)

For example, before placing any BD, $Q^{loss}_{total}(\Psi)$ and $Q^{loss}_{max}(\Psi)$ are all different between using TVGF and CGF. A similar pattern exists when installing 1 to 10 BDs. More importantly, the BD placement solutions ($\boldsymbol{X}$) are different between using TVGF and CGF for both objective functions. For example, when installing 3 BDs to minimize $Q^{loss}_{total}(\Psi)$ (Fig. 5.18), the optimal solutions are to block transformer $\{T_1, T_8, T_9\}$ for TVGF and CGF (90°), $\{T_1, T_{12}, T_{13}\}$ for CGF (135°), and $\{T_1, T_6, T_7\}$ for the rest. When installing 3 BDs to minimize $Q^{loss}_{max}(\Psi)$ (Fig. 5.19), the optimal solutions are to block transformer $\{T_1, T_6, T_7\}$ for TVGF and CGF (45°), $\{T_3, T_6, T_7\}$ for CGF (Average of TVGF), $\{T_6, T_7, T_{11}\}$ for CGF (0°), $\{T_1, T_8, T_9\}$ for CGF (90°), and $\{T_1, T_{12}, T_{13}\}$ for CGF (135°). These observations confirm again that the objective function and GMD event scenarios would affect the solutions of the OBP problem.

Hence, our experiments show that the input GMD event along with the distribution of TVGF's angles and magnitudes would play a key role in determining the BD placement solutions. Since future GMD events are hard to predict, generating a synthetic TVGF using the distribution of GF's angles and magnitudes in previous significant GMD events may be worth consider-

ing for assessing GMDs' impact and solving OBP problems. This is exactly why we have used the benchmark GMD event for the TVGF studies. Regardless of the input GMD event, our proposed SA method is guaranteed to solve the OBP problems effectively and efficiently.

5.4.3.3 TVGF Scenario with Max GIC Constraints

To demonstrate the capability and flexibility of our SA method, we further consider the maximum GIC limit constraints. Figures 5.20 and 5.21 show the minimal $Q_{total}^{loss}(\Psi)$ and $Q_{max}^{loss}(\Psi)$ under different GIC limits. Again, the incomplete lines are due to the feasibility issue. For example, when GIC limit is 25 A, at least 3 BDs are required to keep all $I_t^{eff}(\Psi)$ under limit. The observation here is similar to that in Fig. 5.8. The larger the max GIC limit, the smaller the objective function in the OBP can be.

As shown in Fig. 5.20 and Fig. 5.21, when the GIC limit is 75 A, there is no additional impact on the blocking solutions. This behavior is of course dependent on how we limit effective GIC over time, the choice of objective function, and the duration of the TVGF. Note that in Fig. 5.8, an 150 A limit

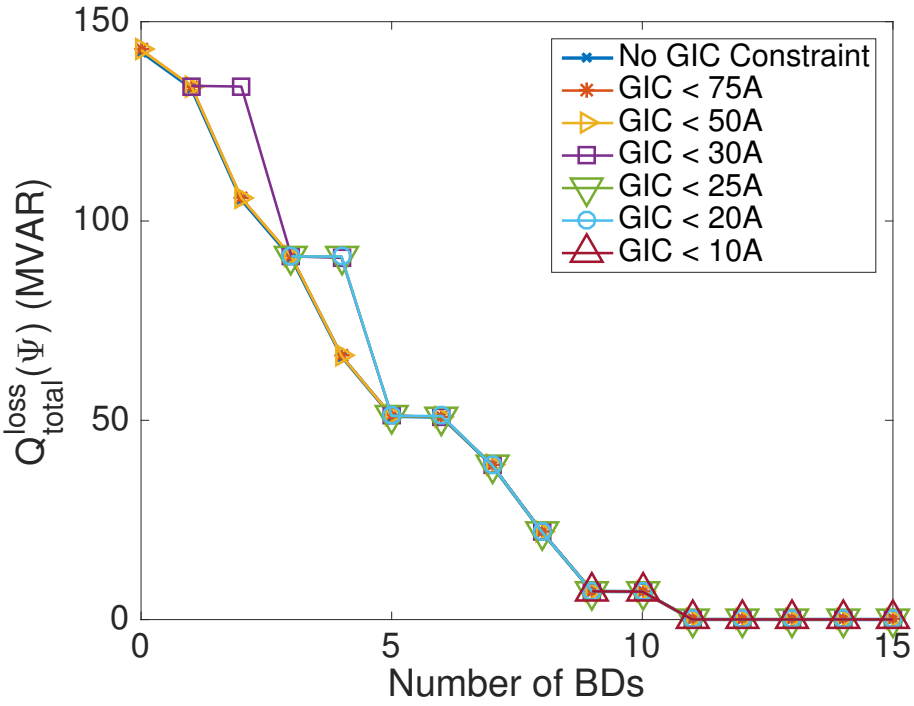


Figure 5.20: BD placement under different GIC constraints using TVGF and $Q_{total}^{loss}(\Psi)$

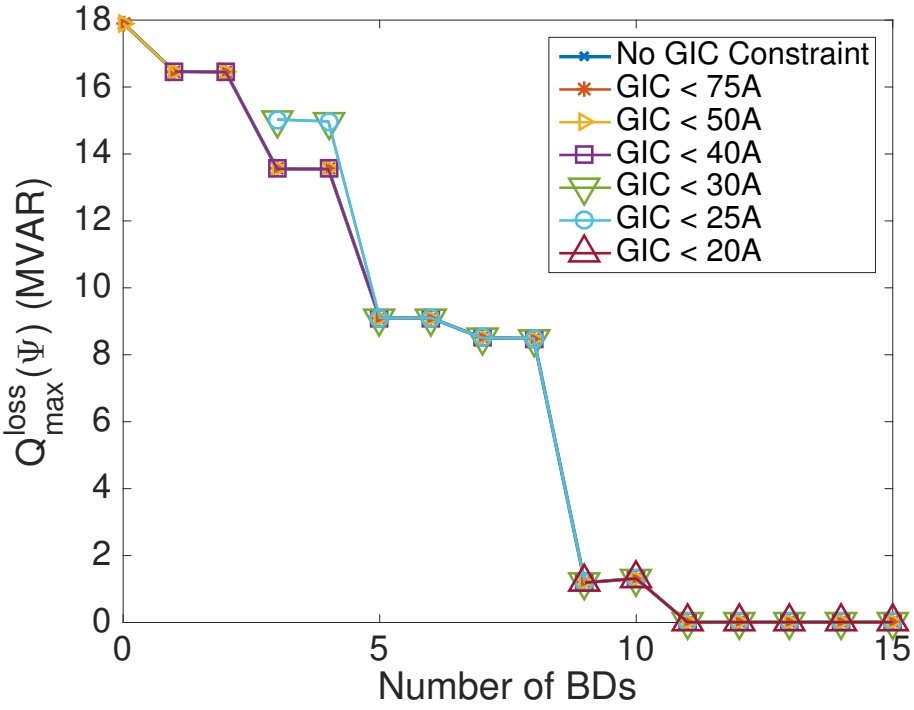


Figure 5.21: BD placement under different GIC constraints using TVGF and $Q^{loss}_{max}(\Psi)$

is needed for the CGF scenario. This comparison implies that the max GIC limit may be less of a concern for TVGF scenario. This is perhaps because the varying angle and magnitude of GFs during the TVGF event could reduce the max GIC levels when the temporal averaging effects are considered. Since the peaks of the benchmark GMD event are random and isolated (shown in Fig. 5.13), the thermal impact of TVGF on transformers is less severe than that of a fixed CGF. Therefore, with the capability of including max GIC limits, our proposed SA method is very suitable for practical system planning studies under realistic TVGF scenarios.

5.5 Conclusion

This chapter presents a simulated annealing (SA) based method that can solve various scenarios of the optimal blocking device placement (OBP) problem, by which effects of geomagnetically induced currents (GICs) are mitigated. We study the basic OBP problem, the OBP problem with maximum effective GIC constraints and two new scenarios, the block by system (BBS)

and time-varying geoelectric field (TVGF) scenarios, which have never been studied before. We solved the OBP problem with the maximum effective GIC constraint for individual transformers, which is necessary in practice. In the BBS scenario, we solve the OBP problem for more realistic use cases where we consider impact on other systems. In the TVGF scenario, we use a more realistic geoelectric field, which varies over time, in GMD modeling to better estimate and mitigate the real GMDs impact. In the OBP problem, considering these three new scenarios greatly improves the GIC modeling. At the same time, the mathematical complexity of the problem is increased significantly and it becomes much more difficult to solve. All previous methods fall short on solving these new scenarios. In this chapter, a novel SA method is developed to tackle the challenge. We compared the solution quality and the running time between our SA method and methods from previous works. For the basic OBP problem, we demonstrated that, with the SA design in this chapter, the solution quality of the SA method is close to that of the BC method and is much better than that of the relaxed SOCP method. Note that the SA algorithm statistically guarantees global optimal when the run time is not restricted. In cases where higher quality solution is required, we can always tune the SA design to accommodate that. In terms of running time, the SA method is much faster than the BC method and is even faster than the relaxed SOCP method when the number of BDs is large. More importantly, we demonstrate that the SA method can handle the three new scenarios mentioned above. Covering these new scenarios represents a major milestone for mitigating the effects of GICs as they are crucial in realistic and practical applications. Furthermore, as shown in the chapter, the SA method we proposed offers a framework that solves not only the specific OBP problems formulated in this chapter, but also OBP-related problems with various objective functions and constraints.

CHAPTER 6

POISONING ATTACK ON SHORT-TERM LOAD FORECASTING

Short-term load forecasting systems for power grids have demonstrated high accuracy and have been widely employed for commercial use. However, classic load forecasting systems, which are based on statistical methods, are subject to vulnerability from training data poisoning. In this chapter, we demonstrate a data poisoning strategy that effectively corrupts the forecasting model even in the presence of outlier detection. To the best of our knowledge, poisoning attack on short-term load forecasting with outlier detection has not been studied in previous works. Our method applies to several forecasting models, including the most widely-adapted and best-performing ones, such as multiple linear regression (MLR) and neural network (NN) models. Starting with the MLR model, we develop a novel closed-form solution to quickly estimate the new MLR model after a round of data poisoning without retraining. We then employ line search and simulated annealing to find the poisoning attack solution. Furthermore, we use the MLR attacking solution to generate a numerical solution for other models, such as NN. The effectiveness of our algorithm has been tested on the Global Energy Forecasting Competition (GEFCom2012) data set with the presence of outlier detection.

6.1 Introduction

In the era of internet of things, cybersecurity is of growing concern to power industries [41]. As power systems benefit from stronger connectivity and advanced probabilistic modeling, they also become more vulnerable to attacks that target these aspects. Malicious cyberbehaviors and technologies that used to challenge security in areas unrelated to power systems, such as information integrity or privacy, have suddenly started to endanger the safety

of large-scale smart grids [69]. Short-term load forecasting (STLF) is one of many aspects that are subject to these attacks. Today, accurately conducting STLF is essential to power systems [42]. The power industry relies heavily on accurate predictions to increase efficiency, reduce waste and maintain stability. If the forecast is corrupted, not only could there be financial losses, in extreme but realistic cases, the bulk power system assets could be damaged, resulting in safety hazards. Among many cybersecurity issues, data integrity attacks, where malicious attackers access and modify sensitive data, pose a great threat to STLF.

Over the past years, many methods were proposed to conduct STLF [42]. A regression model was first introduced in [70] and many regression-based models were later proposed to further improve the forecasting accuracy [71, 72]. In GEFCom 2012 [73], a regression-based model won the top place [74]. Researchers also approach the problem using other methods such as artificial neural network [75–77], support vector machine [78], and fuzzy logic regression [79]. For all of the statistical methods mentioned above, integrity of training data is essential. People face a dilemma: on one hand, these probabilistic models have demonstrated very impressive performance; on the other hand, it is hard to abstract any information from these model that can be easily understood by humans. These methods are purely data driven; all meaningful rules or guidelines resulting from these methods, whether inspiring or totally wrong and damaging, are extracted from, and only from, the training data. This is the reason why the community should pay close attention to the data integrity of the load forecasting system.

The authors of [80] pioneered the research on data integrity attacks on load forecasting systems. They conducted studies to understand how poor training data could affect forecasting accuracy. This work simulates the random attack by selecting $k\%$ of the training data and multiplying the original load by $1+p\%$, where p follows a normal or uniform distribution. The attack is only noticeable when the amount of data to attack and the magnitude of the change are large. For example, without any attack, the mean absolute percentage error (MAPE) using the original vanilla model in [80] is 5.22%. For a normally-distributed ($p \sim \mathcal{N}(\mu, \sigma^2)$) attack, to reach the attack goal of increasing MAPE from 5.22% to 10%, the random attack needs to modify 30% of the data with $\mu = 0.4$ when $\sigma = 0.4$. This implies that on average, 30% of the original load values need to be 1.4X their original value. There-

fore, the random attack is not realistic as it can easily be identified by visual inspection of the historical load curve, let alone if there is a dedicated outlier detector.

To the best of our knowledge, no previous work has considered developing an attacking strategy that not only modifies the training data effectively, but also fools an outlier detector. In this chapter, we will present a data poisoning algorithm that reduces forecasting accuracy without setting off an outlier detector. We will start with MLR model and extend our study to other models, including neural networks. Before we illustrate the details of our attaching strategy in Section 6.3, we first cover the background in Section 6.2. We then present our experimental findings in Section 6.4. We conclude our work in Section 6.5. Future research directions have also been proposed in the same section.

6.2 Preliminaries

6.2.1 Short-term Load Forecasting

In STLF problem, load in the short-term future is forecasted using information such as historical load, date and time, and temperature. Many statistical models [42, 70–79, 81], have been proposed to solve this problem. In this section we will discuss two of the most representative and best performing models that are widely used for load forecasting. Nonetheless, our numerical solution can be applied to a wide range of other models.

6.2.1.1 Load Forecasting Models

We will first discuss the multiple linear regression (MLR) model [82]. It attempts to learn the relationship between multiple explanatory variables and a scalar response variable by fitting a linear mapping between them. It has been proven effective for load forecasting [72]. Note that the model constructed by MLR can be non-linear when the explanatory variables contain non-linear terms. Since the goal is to predict load in the future, time and temperature are treated as explanatory variables, and the load is treated as

a response variable. The MLR model is usually fitted using the least squares method [82].

In [70], linear regression is proposed to solve the load forecasting problem. Later, [71] proposed to use 24 MLR models to forecast load in 24 hours. Several other regression models were proposed in [72], [81]. Although there are many variations of this method, their core concept is similar. In this chapter, we will use the vanilla benchmark model in GEFCom2012, which is also used in [74, 80], to establish a fair comparison with previous studies. It is worth mentioning that our method works for various linear regression models and does not depend on any specific model. This vanilla model [80] is defined as

$$\begin{aligned} & Y = X\beta \\ =&\beta_0 + \beta_1\chi_T + \beta_2\chi_M + \beta_3\chi_H\chi_W + \beta_4\chi_t\chi_H + \beta_5\chi_t^2\chi_H \\ &+ \beta_6\chi_t^3\chi_H + \beta_7\chi_t\chi_M + \beta_8\chi_t^2\chi_M + \beta_9\chi_t^3\chi_M, \end{aligned} \tag{6.1}$$

where χ_T is a nature number that represents a linear trend, χ_M is a 12-dimensional class variable representing 12 months of a year, χ_W is a 7-dimensional class variable representing 7 days of a week, χ_H is a 24-dimensional class variable representing 24 hours of a day, χ_t represents the temperature and β is the parameter for the MLR model. In total, β in the vanilla model has 289 dimensions. In this chapter, we generate the 289-dimensional training data $\{X, Y\} = \{(x_i, y_i), i = 1, 2, ..., n\}$ using the original temperature, date and time, and load information in the data set. The parameters for MLR model are obtained by solving a least squares problem

$$\min_{\beta} \sum_{x \in X} (y - \hat{y})^2 = (Y - X\beta)^T (Y - X\beta). \tag{6.2}$$

Neural networks (NNs) have also been widely used for load forecasting [75–77]. Similar to MLR, NN also attempts to learn the relationship between explanatory variables (temperature and time) and load. Thus, NN usually takes temperature and time data as input and outputs the load. To have a fair comparison with previous studies, we adopt a feed-forward neutral network [72] that was also used in [80]. This NN has 45 input neurons, including month of a year, day of a week, hour of a day, temperature and

trend, and has one output neuron, which represents load. It has one hidden layer with 22 neurons and is fully connected.

6.2.1.2 Outlier Detection

Outlier detection is a well established area of study [83, 84]. We use the outlier detection method developed in [85] for reference. This approach has been proven to be effective in the GEFCom competition. The method first trains a benchmark model using all of the historical data and calculates the absolute percentage error (APE) for each hourly load. The data points with APE values larger than a certain threshold ϕ ($\phi = 50\%$) are treated as outliers. For these outliers, the machine learner will replace the original data with predicted values in the training process. This predicted value is also called the regulated value. We will show that our attacking method can also fool purely statistics-based outlier detection method in Section 6.3.

6.2.2 Poisoning Attack

Machine learning algorithms rely on training data to generate predictive models. Poisoning attack contaminates the learner's training data to mislead the model from being accurate and it has recently caught attention in the machine learning research community [86, 87].

6.2.2.1 Attacking Model

In load forecasting, the training data include date and time, temperature and load. In this chapter, we assume that attackers only have the ability to modify temperature and load data. To ensure that the data attack is realistic, we do not allow attackers to modify the date and time as these modifications can easily be detected.

6.2.2.2 Attacking Goal

The goal of a poisoning attack is to reduce the accuracy of load forecasting while not being detected by any outlier detector. Specifically, we maximize a loss function over a selected data set $\{X_a, Y_a\}$. We can achieve certain

attacking goals by choosing the desired data set. For example, we can choose the original testing data set $\{X_t, Y_t\}$ to attack the overall load forecasting accuracy or we can choose to use peak hour data to attack peak hour load forecasting accuracy. Denote β the model obtained from training data $\{X, Y\}$ and $\hat{y_i}$ the predicted load for input x_i using model β. The poisoning attack problem can then be formulated as

$$\begin{aligned} \max_{X,Y} \quad & \boldsymbol{L}(\{X, Y\}, \{X_a, Y_a\}) = \sum_{\{x_a, y_a\} \in \{X_a, Y_a\}} (y_a - \hat{y_a})^2 \\ s.t. \quad & \mathit{IsOutlier}(\{x, y\}) = \mathit{False}, \ \ \forall \ \ \{x, y\} \in \{X, Y\}, \end{aligned} \tag{6.3}$$

where $\boldsymbol{L}$ is the loss function, where we use l^2-norm to measure the loss of the model over the selected data set $\{X_a, Y_a\}$ and the constraint here is none of the poisoned training data shall be flagged as an outlier.

6.3 Methods

To attack load forecasting models, we start with the MLR model as it has many good mathematical properties. Then we will describe the numeric approach for attacking load forecasting using the NN Model.

6.3.1 Attack Multiple Linear Regression Model

The goal here is to poison the training data and reduce the forecasting accuracy while not being detected by an outlier detector. It is obviously impossible to enumerate all possible poisoning attacks to training data, retrain the model and identify the optimal poisoning attack solution. The bottleneck here is the solution space of poisoning attack and the running time of model retraining. The ability to quickly estimate the behavior of the new model without retraining is the key to solve this problem. In this chapter, we develop a closed-form solution for estimating the new model after poisoning without the need of retraining.

Denote the training data $\{X, Y\}$ where X includes date and time, temperature, and Y represents load. We use the well-known and accurate vanilla model [80] in (6.1). We will first discuss the scenario of attacking only load

data in Y (Section 6.3.2) and then extend it to attacking both temperature in X and load data in Y (Section 6.3.3). The reason we separate these two scenarios is that temperature is treated as an explanatory variable, thus the mathematical properties are dramatically different for these two scenarios.

6.3.2 Attack Load Only

In this scenario, the data attribute subject to attack is only the historical load. As shown in [82], given training data $\{X, Y\}$, solving (6.2) yields the MLR model, denoted by $\hat{\beta}$:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \tag{6.4}$$

We choose to maximize the l^2-norm loss over the original testing data set $\{X_t, Y_t\}$. Therefore, based on (6.3), the mathematical formulation of the attacking goal becomes

$$\boldsymbol{L}(\{X, Y\}, \{X_t, Y_t\}) = \sum_{\{x_t, y_t\} \in \{X_t, Y_t\}} (y_t - \hat{y}_t)^2. \tag{6.5}$$

Note that the predicted load $\hat{y}_t$ can be calculated by

$$\hat{y}_t = x_t \hat{\beta}. \tag{6.6}$$

Therefore, the attack goal, or the objective function to maximize, in the poisoning attack problem becomes

$$\boldsymbol{L} = \sum_{\{x_t, y_t\} \in \{X_t, Y_t\}} (y_t - \hat{y}_t)^2 = (Y_t - X_t \hat{\beta})^T (Y_t - X_t \hat{\beta}). \tag{6.7}$$

The partial derivative of $\boldsymbol{L}$ over $\hat{\beta}$ can be calculated by

$$\begin{aligned} \frac{\partial \boldsymbol{L}}{\partial \hat{\beta}} =& 2(X_t X_t^T \hat{\beta} - X_t^T Y_t) \\ =& 2(X_t X_t^T (X^T X)^{-1} X^T Y - X_t^T Y_t). \end{aligned} \tag{6.8}$$

Algorithm 3: Line Search based Poisoning Attack Algorithm

Set $Y_0 = Y$;
Set converge counters and thresholds for Y and $\mathbf{L}$;
for $k = 0 \rightarrow k_{max}$ *(max training iterations)* **do**
    Set $Y_k^0 = Y_k$;
    for $j = 0 \rightarrow j_{max}$ *(max line search iterations)* **do**
        Compute the descent direction $\mathbf{p}_j$ using (6.10);
        Choose α_j to "loosely" minimize $h(\alpha_j) = \mathbf{L}(Y_k^j + \alpha \mathbf{p}_j)$ over $\alpha \in \mathbb{R}_+$;
        Update $Y_k^{j+1} = Y_k^j + \alpha \mathbf{p}_j$;
    end
    if $Y_k^{j_{max}+1}$ *is outlier* **then**
        Set $Y_{k+1} = Y_k$;
    else
        Set $Y_{k+1} = Y_k^{j_{max}+1}$;
        Update surrounding points to smooth the load curve;
    end
    Stop if S or $\boldsymbol{L}$ converges;
end

With (6.4), the partial derivative of $\hat{\beta}$ over Y becomes

$$\frac{\partial \hat{\beta}}{\partial Y} = (X^T X)^{-1} X^T. \tag{6.9}$$

As a result, the derivative of $\boldsymbol{L}$ over Y can be calculated as

$$\begin{aligned} \frac{\partial \boldsymbol{L}}{\partial Y} =& 2(X^T X)^{-1} X^T (X_t^T X_t \hat{\beta} - X_t^T Y_t) \\ =& 2(X^T X)^{-1} X^T (X_t^T X_t (X^T X)^{-1} X^T Y - X_t^T Y_t). \end{aligned} \tag{6.10}$$

The satisfying property here is that we need not re-train the model to estimate the loss function value after a small change. This closed-form model estimation enables the search for the poisoning solution. As described in Algorithm 3, we use a line search algorithm to search along the gradient for the poisoning solution. For each point that we attack, we also modify the surrounding points accordingly to smooth the curve using polynomial interpolation. If the current poisoning solution is an outlier, we set it to the corresponding regulated value to ensure that the attack is not detectable.

6.3.3 Attack Both Temperature and Load

In this scenario, we will discuss how to conduct poisoning attack on both temperature and load. We will first discuss model estimation and then present the poisoning attack algorithm.

6.3.3.1 Model Estimation

Similar to Algorithm 3, we first estimate the model β after modifying training data without retraining. In this case we use a similar loss function to measure the attacking goal.

$$\textit{Attack Goal} = \boldsymbol{L}(\{X, Y\}, \{X_t, Y_t\}). \tag{6.11}$$

Let $\hat{\beta}^{(-i)}$ be the trained model without training point x_i, $\hat{\beta}^{(+\delta)}$ be the trained model with additional point x_δ, and $\hat{y}_i$ be the predicted value for point $\{x_i, y_i\}$ using model $\hat{\beta}$. Denote $(X^{(-i)}, Y^{(-i)})$ as the training input after removing one point $\{x_i, y_i\}$. With (6.4), the new model can be calculated by

$$\hat{\beta}^{(-i)} = ({X^{(-i)}}^T X^{(-i)})^{-1} {X^{(-i)}}^T Y^{(-i)}. \tag{6.12}$$

Removing x_i and y_i is equivalent to removing one row from X and Y. As a result

$${X^{(-i)}}^T X^{(-i)} = X^T X - x_i^T x_i, \tag{6.13}$$

$${X^{(-i)}}^T Y^{(-i)} = X^T Y - x_i^T y_i. \tag{6.14}$$

With the Sherman-Morrison formula,

$$\begin{aligned} &(X^T X - x_i^T x_i)^{-1} \\ &= (X^T X)^{-1} + \frac{(X^T X)^{-1} x_i^T x_i (X^T X)^{-1}}{1 - x_i (X^T X)^{-1} x_i^T}. \end{aligned} \tag{6.15}$$

Denote hat matrix

$$H = X(X^T X)^{-1} X^T. \tag{6.16}$$

It is obvious that $H_{ii} = x_i (X^T X)^{-1} x_i^T$. Note that

$$\hat{\beta} = (X^T X)^{-1} X^T Y, \tag{6.17}$$

$$\hat{y}_i = x_i \hat{\beta}. \tag{6.18}$$

Then we have

$$\begin{aligned}
\hat{\beta}^{(-i)} =& (X^{(-i)^T} X^{(-i)})^{-1} X^{(-i)^T} Y^{(-i)} \\
=& (X^T X - x_i^T x_i)^{-1} (X^T Y - x_i^T y_i) \\
=& ((X^T X)^{-1} + \frac{(X^T X)^{-1} x_i^T x_i (X^T X)^{-1}}{1 - H_{ii}})(X^T Y - x_i^T y_i) \\
=& \hat{\beta} - (X^T X)^{-1} x_i^T y_i + \frac{(X^T X)^{-1} x_i^T \hat{y}_i}{1 - H_{ii}} \\
& - \frac{(X^T X)^{-1} x_i^T H_{ii} y_i}{1 - H_{ii}} \\
=& \hat{\beta} - \frac{(X^T X)^{-1} x_i^T}{1 - H_{ii}} (y_i - \hat{y}_i).
\end{aligned} \tag{6.19}$$

This equation indicates that we can calculate the new model $\hat{\beta}^{(-i)}$ with just X, H, x_i, y_i and $\hat{y}_i$. As a result, there is no need for retraining on $(X^{(-i)}, Y^{(-i)})$. Similarly if we added a new point $\{x_k, y_k\}$ to the input training data (X, Y), we have

$$\hat{\beta}^{(+k)} = \hat{\beta} + \frac{(X^T X)^{-1} x_k^T}{1 - H_{ii}} (y_k - \hat{y}_k). \tag{6.20}$$

If we replace $\{x_i, y_i\}$ with $\{x_i + \delta, y_i + \delta\}$, denoted by $\{x_\delta, y_\delta\}$, we can approximate the model by

$$\begin{aligned}
\hat{\beta}^{(-i,+\delta)} =& \hat{\beta^{(-i)}}^{(+\delta)} \\
=& \beta^{(-i)} + \frac{(X^{(-i)^T} X^{(-i)})^{-1} x_\delta^T}{1 - x_\delta (X^{(-i)^T} X^{(-i)})^{-1} x_\delta^T} (y_\delta - \hat{y_\delta}^{(-i)}).
\end{aligned} \tag{6.21}$$

Denote $H_{\delta\delta} \doteq x_\delta (X^{(-i)^T} X^{(-i)})^{-1} x_\delta^T$, as $x_\delta \to x_i, y_\delta \to y_i$

$$\begin{aligned}
H_{\delta\delta} &= x_\delta [(X^T X)^{-1} + \frac{(X^T X)^{-1} x_i x_i^T (X^T X)^{-1}}{1 - x_i^T (X^T X)^{-1} x_i}] x_\delta^T \\
&\approx H_{ii} + \frac{H_{ii}^2}{1 - Hii} = \frac{H_{ii}}{1 - Hii}.
\end{aligned} \tag{6.22}$$

With (6.19), as $x_\delta \to x_i, y_\delta \to y_i$, we have

$$y_\delta - \hat{y_\delta}^{(-i)} \quad = y_\delta - x_\delta \hat{\beta}^{(-i)} = \frac{1}{1 - H_{ii}} (y_i - \hat{y}_i). \tag{6.23}$$

Therefore, (6.21) can be expressed as

$$\begin{aligned}&\hat{\beta}^{(-i,+\delta)}|_{\{x_\delta \to x_i, y_\delta \to y_i\}} = \\ =&\hat{\beta}^{(-i)} + \frac{(X^{(-i)^T} X^{(-i)})^{-1} x_i^T}{1 - 2H_{ii}}(y_i - \hat{y_i}).\end{aligned} \tag{6.24}$$

Different from the scenario in 6.3.2, we cannot directly calculate gradient. As shown in (6.4), there is no simple closed-form solution for calculating the derivative of β over X. With (6.24), we can estimate the model $\hat{\beta}^{(-i,+\delta)}$ without retraining the model after making a small change to one training point. We then use a simulated annealing based algorithm to find the poisoning attack solution for this scenario.

6.3.3.2 Simulated Annealing (SA) based Poisoning Attack Algorithm

The SA-based poisoning strategy is described in Algorithm 4. Unlike the previous case, we use SA in the inner loop to search for the poisoning solution that maximizes $\boldsymbol{L}$. Similar to Algorithm 3, we iteratively verify and regulate the poisoning solution in the outer loop to ensure that the poisoning attack is not detectable. SA is a generic stochastic algorithmic approach for finding the global optimum of an optimization problem that may have many local optima [39, 40]. It can handle cost functions with arbitrary degrees of nonlinearities, discontinuities, stochasticity, arbitrary boundary conditions and constraints. It first selects an initial solution and iteratively searches for the optimal solution by emulating the physical cooling process of a solid in a heat bath. In the searching process, SA probabilistically accepts worse solutions which allows it to jump out of the basin that contains local optimum and explore a larger solution space.

At each outer iteration, the point $\{x_i, y_i\}$ with the highest influence value [82] will be modified. In the inner SA loop, similar to (6.10), we calculate the partial derivative $\frac{\partial \boldsymbol{L}}{\partial y_i}$ and modify the selected point based on it. We also ensure that the modification is less than 1% for both temperature and load. The modified point is the neighbor of the current point and will be used in the next SA iteration. The annealing schedule, which defines how SA temperature T systematically decreases as the searching proceeds, is defined

Algorithm 4: SA based Poisoning Attack Algorithm

Set $\{X_0, Y_0\} = \{X, Y\}$;
Set converge counters and thresholds for S and $\mathbf{L}$;
for $k = 0 \rightarrow k_{max}$ *(max converge iterations)* **do**
    Set initial temperature $T = T_0$;
    Set initial solution $S_k^0 = \{X_k, Y_k\}$;
    for $i = 0 \rightarrow i_{max}$ *(max SA iterations)* **do**
        Pick a neighbor, $S_k^{new} = \text{neighbor}(S_k^i)$;
        Estimate objective function values, $\mathbf{L}(S_k^i)$ and $\mathbf{L}(S_k^{new})$, using (6.24);
        Update SA temperature $T = \text{temperature}(T_0, i)$;
        if $P(S_k^i, S_k^{new}, T) \geq random(0, 1)$ **then**
            Accept new solution $S_k^{i+1} = S_k^{new}$;
        else
    if $S_k^{i_{max}+1}$ *is outlier* **then**
        $\{X_{k+1}, Y_{k+1}\} = \{X_k, Y_k\}$;
    else
        $\{X_{k+1}, Y_{k+1}\} = S_k^{i_{max}+1} = \{X_k^{i_{max}+1}, Y_k^{i_{max}+1}\}$;
        Update surrounding points to smooth the temperature and load curve;
    Stop if S or $\boldsymbol{L}$ converges.

as

$$T = \text{temperature}(T_0, k) = T_0 * \eta^i, \quad \eta = 0.95 \tag{6.25}$$

where i is the iteration number and T_0 is the initial SA temperature. We set a high initial SA temperature ($T_0 = 1000$) so that the SA algorithm may find a good solution in the full solution space. Last, the probability of moving from the current state S to a candidate new state S_{new} is specified by an acceptance probability function (APF) $P(S, S_{new}, T)$:

$$P(S, S_{new}, T) = \begin{cases} e^{\frac{L(S_{new}) - L(S)}{T}} & , \text{if } L(S_{new}) < L(S) \\ 1 & , \text{if } L(S_{new}) \geq L(S) \end{cases}. \tag{6.26}$$

The APF function $P(S, S_{new}, T)$ is positive when $L(S_{new}) < L(S)$ so that it can prevent the algorithm from getting stuck at a local minimum. When $L(S_{new}) < L(S)$, $P(S, S_{new}, T)$ decreases as T decreases, which means that the SA algorithm will be less likely to accept moves resulting in a worse solution as it cools down.

6.3.4 Constructing Final Poisoning Solution

As shown in (6.1), the input features to MLR are not the organic date and time, temperature, and load. When attacking both temperature and load, we will need an additional step to compute the attacking solution for temperature. We first use the poisoning strategy in Algorithm 4 to find the poisoning solutions for the nonlinear terms $(\chi_t\chi_H, \chi_t^2\chi_H, \chi_t^3\chi_H, \chi_t\chi_M, \chi_t^2\chi_M, \chi_t^3\chi_M)$ and then solve a group of nonlinear equations to construct the final poisoning solutions for temperature. In this chapter, the standard implementation of Trust-Region Dogleg Method in Matlab [88] is employed. We notice that the MLR model can estimate the load curve well and has forecasting accuracy comparable to other models, including NN [74]. In other words, other load forecasting models can be fitted by MLR. As a result, we use the numerical poisoning solutions developed for MLR to attack other models.

6.4 Results

6.4.1 Load Forecasting Data

The data set we use in this chapter comes from The 2012 Global Energy Forecasting Competition (GEFCom2012) [74]. This data set is widely adapted in the load forecasting research community [73, 80]. The data set contains 4.5 years of load data for 21 zones (20 zones and the whole system that contains all zones) and temperature information for 11 stations from a US utility. The resolution of the data goes down to the scale of one hour. It ranges from the 1^{st} hour of 2004/1/1 to the 6^{th} hour of 2008/6/30. In this chapter, the same setup as that of [80] is shared across experiments to ensure fair comparisons. More specifically, we use the load data in 2005 and 2006 as training data and the load data in 2007 as test data. In this experiment, we only allow up to 10% of data to be attacked. There are 17520 data points in the training data set and the number of points that can be attacked is up to 1752. We also follow the methodology in [80, 89] to handle temperature data.

6.4.2 Numerical Results

We implemented a baseline MLR and achieved the same mean APE (5.22%) as that in [80]. We compared the poisoning attacking performance among three different scenarios, including the *strategic attack on load only* in Section 6.3.2, the *strategic attack on temperature and load* in Section 6.3.3, and the *random attack on temperature and load using a normal distribution* ($p \sim \mathcal{N}(0.5, 0.5^2)$ with the methodology from [80]. For random attack, we also ran experiments to check the attacking effectiveness using a uniform distribution. Since there is no fundamental difference between normal and uniform distributions, we only keep results of normally-distributed attack for simplicity. In this experiment, we conduct random attack 20 times and take the average of them as the final attacking performance. We measure the attack performance by the mean APE and maximum APE. It is worth mentioning that randomly modify training data cannot fool outlier detectors. As a matter of fact, these attacks can easily be spotted by visual inspection.

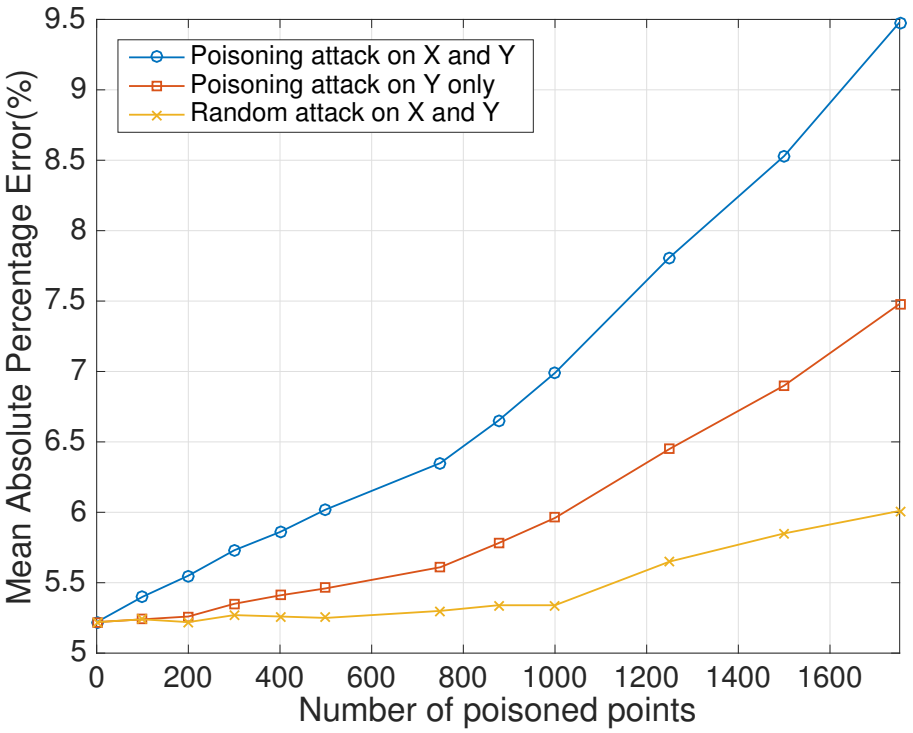


Figure 6.1: Mean APE for poisoning attack on load forecasting

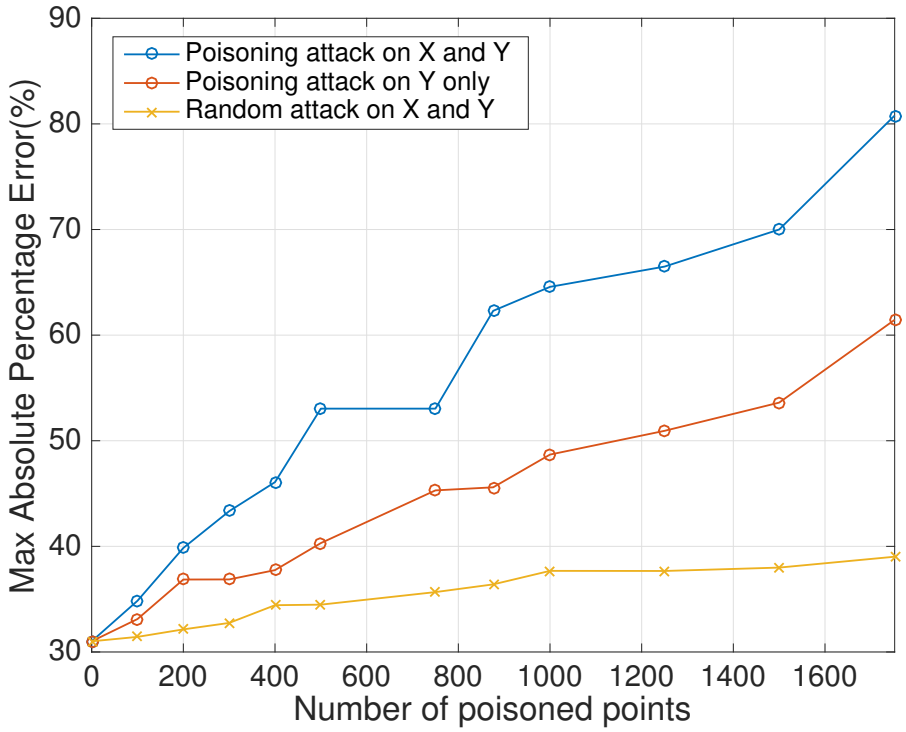


Figure 6.2: Max APE for poisoning attack on load forecasting

Figures 6.1 and 6.2 show the results of mean APE and max APE for all three attacking scenarios in the aggregated zone (Zone 21). Poisoning both temperature and load has better attacking performance (higher mean and max APE) than just poisoning the load. This finding makes sense because the forecasts depend on both temperature and load, and are also very sensitive to the nonlinear terms with x_t in (6.1). As we can see, strategic poisoning attacks perform much better than random poisoning attack. On average, when comparing the increase of mean APE, our non-detectable attacking method that poisons both temperature and load is 5X better than the detectable random attack.

Let us take a closer look at how the distribution of APEs changes after poisoning attack. Figures 6.3 - 6.5 show the histograms of APEs when attacking 10% of the training data. We can see that strategic poisoning attack on both temperature and load has more points with larger APE than the other two. Both strategic poisoning attacks move the distribution towards the right side more significantly than the random attack. We also notice that the random attack sometimes makes significant changes to some points and

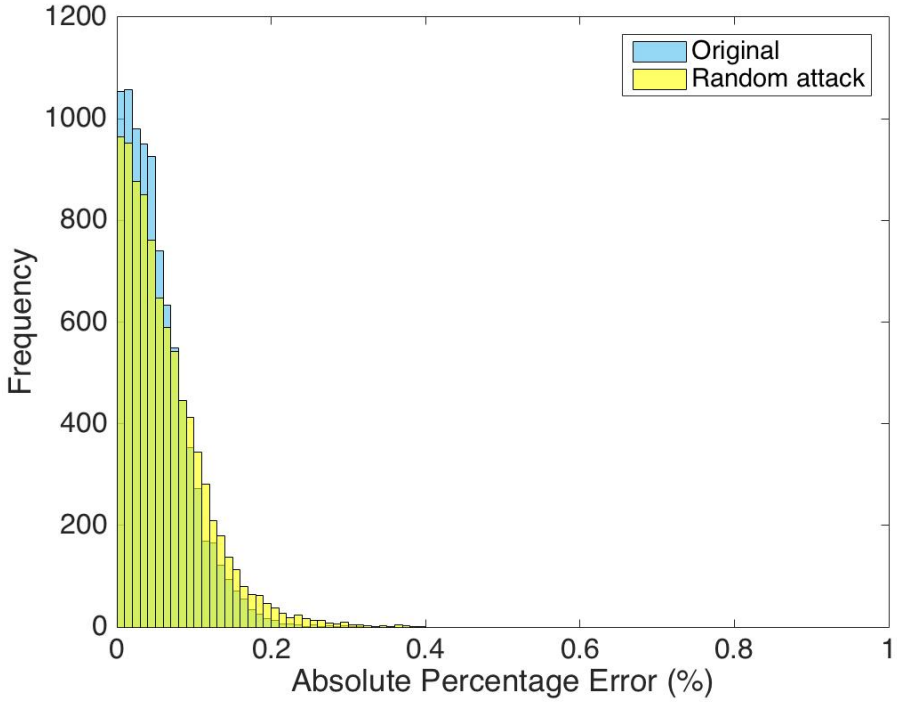


Figure 6.3: Histogram of APE for random attack

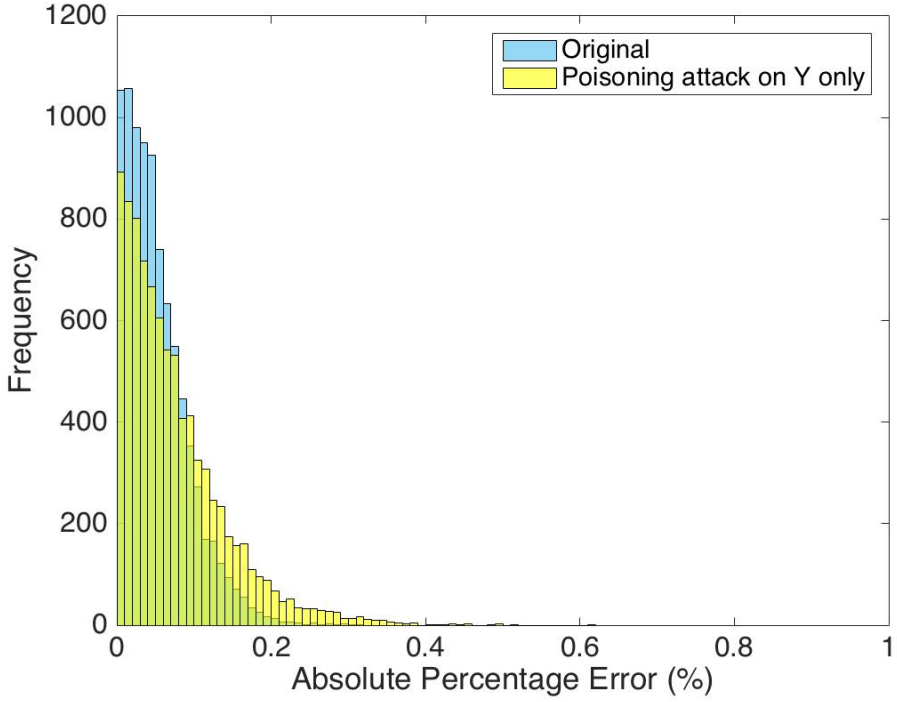


Figure 6.4: Histogram of APE for strategic attack on load

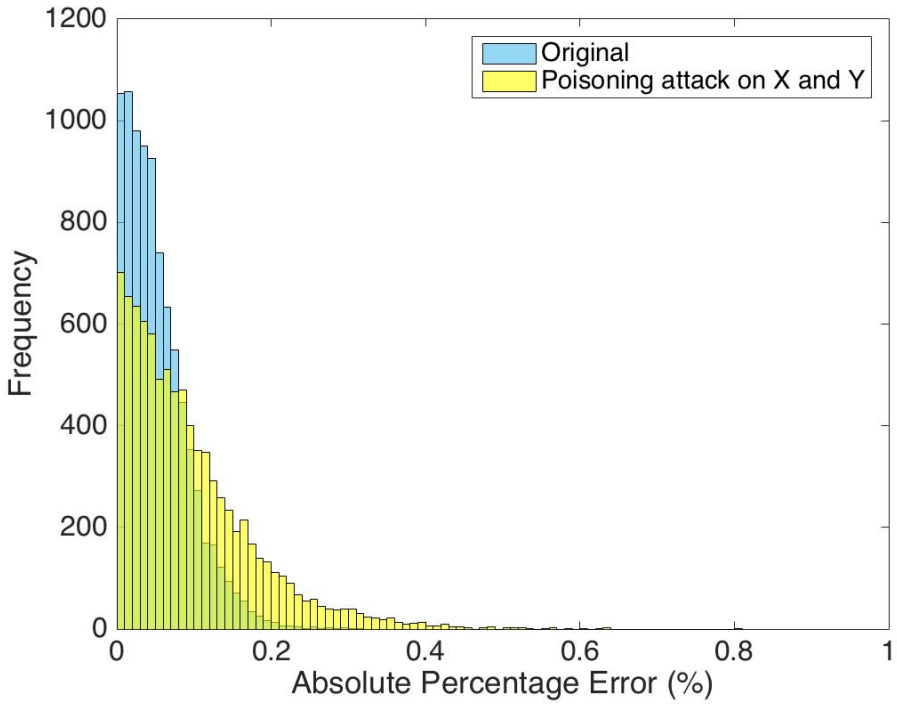


Figure 6.5: Histogram of APE for strategic attack on both temperature and load

these points trigger outlier detection, while our algorithms always limit the attack within the undetectable range to avoid detection. We trained our NN, with the same architecture as that of [80], using the deep learning toolbox Keras [90]. We found that the solutions and attacking performance for NN are almost identical to those for MLR, so we will not re-present a similar set of results here. Considering that both MLR and NN models achieve very similar forecasting accuracy, it is reasonable to expect that the same set of poisoned data resulted in similar outcomes from these two models.

6.5 Conclusion

Short-term load forecasting (STLF) models for power grids are vulnerable to training data poisoning attacks. In this chapter, we developed a poisoning strategy that can corrupt energy load forecasting model even in the presence of outlier detection. Our closed-form model estimation technique, line search and simulated annealing based methods have been shown to be effective in the

presence of an outlier detector. In the future, we would like to explore a wider variety of strategies to generate attacking samples that are more effective and harder to detect. Deep learning based methods such as variational auto-encoder and generative adversarial networks (GANs) might shed light on our study. We are also interested in applying the techniques to a wider variety of forecasting models.

CHAPTER 7

CONCLUSION AND FUTURE WORK

In this dissertation, we solved several key challenges in smart grids to improve their efficiency and reliability from different angles. The approaches include a variety of optimization methods, high-performance computing techniques and machine learning algorithms. Furthermore, we take into account the scalability of our approaches for the applicability to large-scale smart grids, which has been shown in the experimental results.

In Chapters 2 and 3, we introduced the fast ACOPF algorithm and the ClusRed algorithm that can solve ACOPF and POPF much faster while maintaining high accuracy. They help improve the efficiency and reliability of smart grids. In Chapter 2, we presented the fast ACOPF algorithm that uses power system network reduction to speed up the computation of ACOPF problems. Experimental results show that this approach can achieve 1.32× to 7.01× speedup over full ACOPF while introducing just 0.54% error on average. The computation time of the proposed algorithm grows almost linearly. We demonstrated its robustness by showing that as long as the original ACOPF can converge, our method can also converge. Compared to the widely used DCOPF, we reduce the error by 77.6% on average. In Chapter 3, we presented a fast clustering and NR based cumulant method, the ClusRed algorithm, that can solve POPF much faster and more accurately. A new linear mapping matrix based on NR has shown better performance than previous approaches. Experimental results demonstrate that ClusRed can achieve several thousandfold speedup compared to MCS and up to 4.57X compared to previous CMs for large-scale smart grids. On average, we improve the worst estimation accuracy of mean value by 32.08% and standard deviation value by 36.76%. The fast OPF and ClusRed methods can be used to solve ACOPF and POPF problems for large-scale smart grid in many applications, such as market management and reliability analysis.

In Chapters 4 and 5, we designed and implemented the BC and SA methods that can solve OBP problems, which mitigate the effects of GICs. We first show that the BC method can provide optimal solution for the basic OBP problem. Then we study several new scenarios, including the basic OBP problem with per-transformer GIC constraint, the block by system (BBS) scenario where the impact on neighboring systems is limited, and the time-varying geoelectric field (TVGF) scenario where a more realistic geoelectric field, which varies over time, is used in GMD modeling to better estimate and mitigate the real GMDs impact. None of them has ever been studied before. In addition, the mathematical complexity of the OBP problem increased significantly for these new scenarios and it becomes much more difficult to solve. For the first time, we solve the OBP problem for these scenarios by introducing the novel SA method. For the basic OBP problem, we demonstrated that the solution quality of the SA method is close to that of the BC method and is much better than that of the relaxed SOCP method. In terms of running time, the SA method is much faster than the BC method and is even faster than the relaxed SOCP method when the number of BDs is large. More importantly, we demonstrate that the SA method can handle the three new scenarios mentioned above. Covering these new scenarios represents a major milestone for mitigating the effects of GICs as they are crucial in realistic and practical applications. Furthermore, the SA method offers a framework that solves not only the specific OBP problems formulated in this dissertation, but also OBP-related problems with various objective functions and constraints. Further investigations are recommended for the following directions: (1) Accurately modeling of the damages from GMDs and the selection of objective function in the OBP problems; (2) Low cost and flexible devices for mitigating the effects of GMDs.

In Chapter 6, we studied poisoning attack on short-term load forecasting (STLF) systems in power grids. We built a first-of-its-kind poisoning strategy that can corrupt the energy load forecasting model even in the presence of outlier detection. Our closed-form model estimation technique, line search and simulated annealing based methods have been shown to be effective in the presence of an outlier detector. As more and more data-driven methods are used in smart grids, it is extremely critical to protect data integrity and enhance system reliability. It would be desirable to explore a wider variety of strategies to poison training data that are more effective and harder to

detect. Deep learning based methods such as variational auto-encoder and generative adversarial networks (GANs) might shine light. It would also be interesting to apply the techniques to a wider variety of forecasting models. Furthermore, it would be interesting to generate a defensive algorithm for such attacks.

In conclusion, this dissertation contributed to the efficiency and reliability of smart grids by introducing novel high-performance algorithms to solve complex optimization problems. Furthermore, it applied machine learning algorithms to poisoning attacks on STLF systems and demonstrated the necessity of data integrity protection in smart grids.

REFERENCES

[1] "Annual energy review," Energy Information Administration, U.S. Department of Energy, 2011. [Online]. Available: http://www.eia.gov/totalenergy/data/annual/

[2] A. Castillo and R. P. O'Neill, "International energy statistics," U.S. Department of Energy, Energy Information Administration (EIA), Tech. Rep., 2012.

[3] M. B. Cain and R. P. O'Neill, "History of optimal power flow and formulations," *Federal Energy Regulatory Commission Staff Technical Paper*, 2012.

[4] A. Castillo and R. P. O'Neill, "Computational performance of solution techniques applied to the ACOPF," *Federal Energy Regulatory Commission Staff Technical Paper*, 2013.

[5] K. Aoki, A. Nishikori, and R. Yokoyama, "Constrained load flow using recursive quadratic programming," *Power Systems, IEEE Transactions on*, vol. 2, no. 1, pp. 8–16, 1987.

[6] L. Lai, J. Ma, R. Yokoyama, and M. Zhao, "Improved genetic algorithms for optimal power flow under both normal and contingent operation states," *International Journal of Electrical Power and Energy Systems*, vol. 19, no. 5, pp. 287–292, 1997.

[7] M. Abido, "Optimal power flow using particle swarm optimization," *International Journal of Electrical Power and Energy Systems*, vol. 24, no. 7, pp. 563–571, 2002.

[8] O. Alsac, J. Bright, M. Prais, and B. Stott, "Further developments in LP-based optimal power flow," *Power Systems, IEEE Transactions on*, vol. 5, no. 3, pp. 697–711, 1990.

[9] X. Yan and V. Quintana, "Improving an interior-point-based OPF by dynamic adjustments of step sizes and tolerances," *Power Systems, IEEE Transactions on*, vol. 14, no. 2, pp. 709–717, 1999.

[10] Y.-C. Wu, A. Debs, and R. Marsten, "A direct nonlinear predictor-corrector primal-dual interior point algorithm for optimal power flows," *Power Systems, IEEE Transactions on*, vol. 9, no. 2, pp. 876–883, 1994.

[11] J. Yuryevich and K.-P. Wong, "Evolutionary programming based optimal power flow algorithm," *Power Systems, IEEE Transactions on*, vol. 14, no. 4, pp. 1245–1250, 1999.

[12] D. Forner, T. Erseghe, S. Tomasin, and P. Tenti, "On efficient use of local sources in smart grids with power quality constraints," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, 2010, pp. 555–560.

[13] J. B. Ward, "Equivalent circuits for power-flow studies," *Electrical Engineering*, vol. 68, no. 9, pp. 794–794, 1949.

[14] "A vision for the smart grid," National Energy Technology Laboratory, Tech. Rep., 2009.

[15] B. Borkowska, "Probabilistic load flow," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-93, no. 3, pp. 752–759, May 1974.

[16] X. Cheng and T. Overbye, "PTDF-based power system equivalents," *Power Systems, IEEE Transactions on*, vol. 20, no. 4, pp. 1868–1876, 2005.

[17] H. Oh, "A new network reduction methodology for power system planning studies," *Power Systems, IEEE Transactions on*, vol. 25, no. 2, pp. 677–684, 2010.

[18] A. Schellenberg, W. Rosehart, and J. Aguado, "Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and Gamma distributions," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 773–781, May 2005.

[19] G. Verbic, A. Schellenberg, W. Rosehart, and C. A. Canizares, "Probabilistic optimal power flow applications to electricity markets," in *2006 International Conference on Probabilistic Methods Applied to Power Systems*, June 2006, pp. 1–6.

[20] H. Cramer, *Numerical Methods of Statistics*. Princeton, NJ: Princeton Univ. Press, 1946.

[21] V. D. Albertson, J. M. Thorson, R. E. Clayton, and S. C. Tripathy, "Solar-induced-currents in power systems: Cause and effects," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 2, pp. 471–477, March 1973.

[22] V. Albertson, J. Thorson, and S. Miske, "The effects of geomagnetic storms on electrical power systems," *IEEE Transactions on Power Apparatus and Systems*, no. 4, pp. 1031–1044, 1974.

[23] "Geomagnetic disturbance reliability standard; reliability standard for transmission system planned performance for geomagnetic disturbance events," Federal Energy Regulatory Commission, Tech. Rep., Nov 2018.

[24] "Effects of geomagnetic disturbances on the bulk power system," North American Electric Reliability Corporation (NERC), Tech. Rep., Feb 2012.

[25] V. D. Albertson, J. G. Kappenman, N. Mohan, and G. A. Skarbakka, "Load-flow studies in the presence of geomagnetically-induced currents," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, no. 2, pp. 594–607, Feb 1981.

[26] D. H. Boteler and R. J. Pirjola, "Modelling geomagnetically induced currents produced by realistic and uniform electric fields," *IEEE Transactions on Power Delivery*, vol. 13, no. 4, pp. 1303–1308, Oct 1998.

[27] R. Pirjola, "Properties of matrices included in the calculation of geomagnetically induced currents (GICs) in power systems and introduction of a test model for GIC computation algorithms," *Earth Planets and Space*, vol. 61, pp. 263–272, 02 2009.

[28] T. J. Overbye, K. S. Shetye, T. R. Hutchins, Q. Qiu, and J. D. Weber, "Power grid sensitivity analysis of geomagnetically induced currents," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4821–4828, Nov 2013.

[29] K. Zheng, D. Boteler, R. J. Pirjola, L.-g. Liu, R. Becker, L. Marti, S. Boutilier, and S. Guillon, "Effects of system characteristics on geomagnetically induced currents," *IEEE Transactions on Power Delivery*, vol. 29, no. 2, pp. 890–898, 2014.

[30] R. Horton, D. Boteler, T. J. Overbye, R. Pirjola, and R. C. Dugan, "A test case for the calculation of geomagnetically induced currents," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 2368–2373, Oct 2012.

[31] I. A. Erinmez, J. G. Kappenman, and W. A. Radasky, "Management of the geomagnetically induced current risks on the national grid company's electric power transmission system," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, no. 5-6, pp. 743–756, 2002.

[32] E. Arajärvi, R. Pirjola, and A. Viljanen, "Effects of neutral point reactors and series capacitors on geomagnetically induced currents in a high-voltage electric power transmission system," *Space Weather*, vol. 9, no. 11, 2011.

[33] L. Bolduc, M. Granger, G. Pare, J. Saintonge, and L. Brophy, "Development of a dc current-blocking device for transformer neutrals," *IEEE Transactions on Power Delivery*, vol. 20, no. 1, pp. 163–168, Jan 2005.

[34] J. S. Foster Jr, E. Gjelde, W. R. Graham, R. J. Hermann, H. M. Kluepfel, R. L. Lawson, G. K. Soper, L. L. Wood Jr, and J. B. Woodard, "Report of the commission to assess the threat to the united states from electromagnetic pulse (emp) attack." National Research Council Washington DC Committee on Electromagnetic Pulse Environment, Tech. Rep., 2004.

[35] A. H. Etemadi and A. Rezaei-Zare, "Optimal placement of GIC blocking devices for geomagnetic disturbance mitigation," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2753–2762, 2014.

[36] A. Rezaei-Zare and A. H. Etemadi, "Optimal placement of GIC blocking devices considering equipment thermal limits and power system operation constraints," *IEEE Transactions on Power Delivery*, vol. 33, no. 1, pp. 200–208, 2018.

[37] H. Zhu and T. J. Overbye, "Blocking device placement for mitigating the effects of geomagnetically induced currents," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2081–2089, July 2015.

[38] M. Tawarmalani and N. V. Sahinidis, "A polyhedral branch-and-cut approach to global optimization," *Mathematical Programming*, vol. 103, no. 2, pp. 225–249, Jun 2005. [Online]. Available: https://doi.org/10.1007/s10107-005-0581-8

[39] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[40] V. Černỳ, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*, vol. 45, no. 1, pp. 41–51, 1985.

[41] A. Sadeghi, C. Wachsmann, and M. Waidner, "Security and privacy challenges in industrial internet of things," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, June 2015.

[42] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, Dec 1987.

[43] Y. Liang and D. Chen, "Fast large-scale optimal power flow analysis for smart grid through network reduction," in *2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC)*, Jan 2014, pp. 373–378.

[44] Y. Liang and D. Chen, "ClusRed: Clustering and network reduction based probabilistic optimal power flow analysis for large-scale smart grids," in *2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, June 2014, pp. 1–6.

[45] Y. Liang, H. Zhu, and D. Chen, "Optimal blocker placement for mitigating the effects of geomagnetic induced currents using branch and cut algorithm," in *2015 North American Power Symposium (NAPS)*, Oct 2015, pp. 1–6.

[46] Y. Liang, D. He, and D. Chen, "Poisoning attack on load forecasting," in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, May 2019.

[47] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *Power Systems, IEEE Transactions on*, vol. 24, no. 3, pp. 1290–1300, 2009.

[48] R. Baldick, B. Kim, C. Chase, and Y. Luo, "A fast distributed implementation of optimal power flow," *Power Systems, IEEE Transactions on*, vol. 14, no. 3, pp. 858–864, 1999.

[49] U. Leeton, D. Uthitsunthorn, U. Kwannetr, N. Sinsuphun, and T. Kulworawanichpong, "Power loss minimization using optimal power flow based on particle swarm optimization," in *Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010 International Conference on*, 2010, pp. 440–444.

[50] W. W. Price et al., "Bibliography on load models for power flow and dynamic performance simulation," *Power Systems, IEEE Transactions on*, vol. 10, no. 1, pp. 523–538, 1995.

[51] "IEEE guide for synchronous generator modeling practices and applications in power system stability analyses," *IEEE Std 1110-2002 (Revision of IEEE Std 1110-1991)*, pp. 1–72, 2003.

[52] H. Li, Z. Chen, and L. Han, "Comparison and evaluation of induction generator models in wind turbine systems for transient stability of power system," in *Power System Technology, 2006. PowerCon 2006. International Conference on*, 2006, pp. 1–6.

[53] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.

[54] R. Zimmerman, C. Murillo-Sanchez, and R. Thomas, "Matpower's extensible optimal power flow architecture," in *Power Energy Society General Meeting, 2009. PES '09. IEEE*, 2009, pp. 1–7.

[55] P. Zhang and S. T. Lee, "Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 676–682, Feb 2004.

[56] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1594–1601, Nov 2007.

[57] M. Madrigal, K. Ponnambalam, and V. H. Quintana, "Probabilistic optimal power flow," in *Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No.98TH8341)*, vol. 1, May 1998, pp. 385–388 vol.1.

[58] G. Verbic and C. A. Canizares, "Probabilistic optimal power flow in electricity markets based on a two-point estimate method," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1883–1893, Nov 2006.

[59] F. Faxvog, G. Fuchs, W. J. D. Wojtczak, M. Marz, S. Dahman, and W. Pewaukee, "HV power transformer neutral blocking device (NBD) operating experience in Wisconsin," in *MIPSYCON Conference November*, vol. 7, 2017.

[60] PowerWorld Corporation, "Powerworld simulator 18," 2016. [Online]. Available: https://www.powerworld.com/

[61] Y. Liang, H. Zhu, and D. Chen, "Optimal blocker placement for mitigating the effects of geomagnetic induced currents using branch and cut algorithm," in *2015 North American Power Symposium (NAPS)*, Oct 2015, pp. 1–6.

[62] "Benchmark geomagnetic disturbance event description," North American Electric Reliability Corporation (NERC), Tech. Rep., May 2016.

[63] "TPL-007-2 transmission system planned performance for geomagnetic disturbance events," North American Electric Reliability Corporation, Tech. Rep., Nov 2017.

[64] V. Granville, M. Krivanek, and J. P. Rasson, "Simulated annealing: a proof of convergence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 6, pp. 652–656, June 1994.

[65] "Screening criterion for transformer thermal impact assessment white paper," North American Electric Reliability Corporation (NERC), Tech. Rep., October 2017.

[66] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," pp. 97–109, 1970.

[67] N. V. Sahinidis, *BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual, 2017.

[68] "Supplemental geomagnetic disturbance event description," North American Electric Reliability Corporation (NERC), Tech. Rep., Oct 2017.

[69] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 210–224, Jan 2012.

[70] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1535–1547, Nov 1990.

[71] R. Ramanathan, R. Engle, C. W. Granger, F. Vahid-Araghi, and C. Brace, "Short-run forecasts of electricity loads and peaks," *International Journal of Forecasting*, vol. 13, no. 2, pp. 161 – 174, 1997.

[72] T. Hong, "Short term electric load forecasting," Ph.D. dissertation, North Carolina State University, 2010.

[73] N. Charlton and C. Singleton, "A refined parametric model for short term load forecasting," *International Journal of Forecasting*, vol. 30, no. 2, pp. 364 – 368, 2014.

[74] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *International Journal of Forecasting*, vol. 30, no. 2, pp. 357 – 363, 2014.

[75] D. C. Park, M. A. El-Sharkawi, R. J. Marks, L. E. Atlas, and M. J. Damborg, "Electric load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 442–449, May 1991.

[76] A. Khotanzad, R. Afkhami-Rohani, and D. Maratukulam, "ANNSTLF-artificial neural network short-term load forecastergeneration three," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1413–1422, Nov 1998.

[77] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, Feb 2001.

[78] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector machines: a study on EUNITE competition 2001," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1821–1830, Nov 2004.

[79] S. U. Hideo Tanaka and K. Asai, "Linear regression analysis with fuzzy model," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 12, no. 6, pp. 903–907, Nov 1982.

[80] J. Luo, T. Hong, and S.-C. Fang, "Benchmarking robustness of load forecasting models under data integrity attacks," *International Journal of Forecasting*, vol. 34, no. 1, pp. 89 – 104, 2018.

[81] P. Wang, B. Liu, and T. Hong, "Electric load forecasting with recency effect: A big data approach," *International Journal of Forecasting*, vol. 32, no. 3, pp. 585 – 597, 2016.

[82] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*, Second Edition, 2012.

[83] A. J. Hodge, V. J., "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, p. 85126, 2004.

[84] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*. John Wiley & Sons, 1987.

[85] J. Xie and T. Hong, "Gefcom2014 probabilistic electric load forecasting: An integrated solution with forecast combination and residual simulation," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1012 – 1016, 2016.

[86] B. Biggio, I. Corona, B. Nelson, B. Rubinstein, D. Maiorca, G. Fumera, G. Giacinto, and F. Roli., "Security evaluation of support vector machines in adversarial environments," *Support Vector Machines Applications*, p. 105153, 2014.

[87] M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, and B. Li, "Manipulating machine learning: Poisoning attacks and countermeasures for regression learning," in *2018 IEEE Symposium on Security and Privacy (SP)*, May 2018, pp. 19–35.

[88] M. J. Powell, "A fortran subroutine for solving systems of nonlinear algebraic equations," United Kingdom, Nov. 1968.

[89] T. Hong, P. Wang, and L. White, "Weather station selection for electric load forecasting," *International Journal of Forecasting*, 2015.

[90] F. Chollet, "Keras," https://github.com/fchollet/keras, 2015.