EFFICIENT AND EFFECTIVE ALGORITHMS FOR CONTINGENCY RESPONSE IN ELECTRIC POWER SYSTEMS WITH TRANSMISSION SWITCHING

A Dissertation

by

WILLIAM ERIC BROWN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Erick Moreno-Centeno
Committee Members,	Sergiy Butenko
	Alfredo Garcia
	Le Xie
Head of Department,	Lewis Ntaimo

August 2020

Major Subject: Industrial Engineering

ABSTRACT

Optimal operation of the electric power grid is of vital importance to the industrial, medical, and defense entities upon which the population relies. Historically, the power grid has been modeled and operated used a fixed configuration. However, system operators have a method at their disposal, transmission switching, wherein transmission lines are physically switched in and out of the grid. By leveraging this additional flexibility of the grid, system operators can obtain benefits which are otherwise inaccessible.

Unfortunately, because the underlying problem is computationally intractable, transmission switching is currently implemented only in limited capacity. To address this, the bulk of the literature uses either DC-based models or unrealistically small test systems. These facts are problematic because such linearizations have proven highly inaccurate in short term operations and small test cases cloud the impact of models and algorithms. Finally, it is clear in the literature that most approaches cannot satisfy the requirements of real-time operations for large systems, motivating the need for superior techniques.

This research develops two techniques to address the above concerns, adding to the literature and pushing transmission switching towards real-world implementation. Specifically, we develop solution approaches for two variants of the AC optimal transmission switching (ACOTS) model which seek to identify optimal transmission switching actions when a portion of the grid fails.

First, we develop a mixed-integer linear optimization model which accurately reflects the ACOTS and is competitive with the state-of-the-art in solution time. The aim of our model is to identify AC optimal transmission switching actions to prevent load shed caused by contingencies. Our approache accelerates an existing model with three constraint relaxations that dramatically decrease solution time with no decrease in accuracy.

Second, we develop a data mining approach to address a variant of the ACOTS which minimizes post-contingency AC power flow violations. Using a guided undersampling procedure consisting of imbalanced-data classification approaches, we significantly outperform state-of-the-art heuristics. Our approach, which is computationally inexpensive and vetted on real-world AC power system data, addresses three key issues which hamper the practical implementation of transmission switching. These facts, combined with the performance of our approach, demonstrate the strength of our technique and move transmission switching towards practical viability.

DEDICATION

To my baby boy, Jack.

ACKNOWLEDGMENTS

The completion of my Ph.D. culminating in the writing of this dissertation could not have been possible without a myriad of persons.

First, I would like to give the most effusive praise and thanks to my wife, Kate. Without your constant support and encouragement (and agreement to bear this nonsense with me), this absolutely would not have come to fruition. Any Ph.D. will tell you about the dark times they went through. You are and were the light in my life. I love you dearly. Thank you.

Next, I would like to express my deepest gratitude to my advisor, Professor Erick Moreno-Centeno. You bringing me into your team marks an inflection point in my career that I will continue cherish and value. Your guidance and support have been invaluable during this process and I am extraordinarily proud that your name appears at the top of my committee.

I also owe a special thanks to each of the members of my committee. To Dr. Sergiy Butenko, Dr. Alfredo Garcia, and Dr. Le Xie, thank you all for your agreement to participate in this process and the guidance you have provided. In addition, I owe a debt of gratitude to several members of the department past and present for their guidance and advice during my stay at Texas A&M: Jorge Leon, Justin Yates, Greg Graves, Satish Bukkapatnam, Kiavash Kianfar, Yu Ding, Richard Feldman, Natarajan Gautam, and Jose Vazquez.

To the faculty at Abilene Christian University, I owe a debt of thanks to you. The education I received while at ACU laid the basis for my Ph.D. study, and I grow more aware and more thankful of the solid foundation I received. In particular, I must thank Dr. Bo Green for his introducing me to the field of operations research. I must also thank Dr.

David Hendricks for his mentorship and encouragement to pursue postgraduate education.

Jason and Natasha, your support was truly invaluable over these several years. Without the two of you Kate and I almost certainly would not have survived. An enormous thanks to the both of you for the various roles you played in my life while in College Station. I look forward to many years of world travels, delicious food, and brown liquor with the two of you.

To my friends at A&M, thank you for your friendship and support over the years. Specifically to Chris Lourenco, I will miss yelling at you when I was tired of talking to myself. I look forward to many years of continuing to do so remotely. To my friends at TBE, most specifically Jonathan Davis and James Young, this almost certainly would not have happened without you. Without the coffee breaks, post-meeting debriefs, and beers where we (or I) lamented our lack of terminal degrees, I would not have had the wherewithal to pursue this massive undertaking.

To my parents, I cannot thank you enough for the upbringing I received which allowed me to make this achievement. Your constant enforcement of independent, critical thinking played an invaluable role in bringing me to this point. I love you both. To my sister and brother-in-law, thank you for the regular encouragement and (much more) regular writing check-ups. The two of you will never know the degree to which you helped along the way. To my mother-in-law and father-in-law, first and foremost, thank you for bringing Kate to me. Less importantly, thank you for all of the support you've provided along the way. Knowing you two were always there for us has given the peace of mind necessary to survive this journey.

To those of you who have reached this point without receiving a personal acknowledgement, I thank you. There are countless nameless people who have encouraged me, given a kind word, or enacted a simple act of kindness without intention, all of which have helped me reach this point. While I will almost certainly never get the opportunity to give

you proper thanks, I hope your kindness is returned to you in full.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of committee chair, Professor Erick Moreno-Centeno of the Department of Industrial and Systems Engineering, committee members Professors Sergiy Butenko and Alfredo Garcia of the Department of Industrial and Systems Engineering, and Professor Le Xie of the Department of Electrical and Computer Engineering. All work described in this dissertation was completed by the student under the advisement of Erick Moreno-Centeno of the Department of Industrial and Systems Engineering.

Funding Sources

The work in this dissertation was supported in part by the National Science Foundation (NSF) award No. CMMI-1451036 and a seed grant from the Texas A&M Energy Institute.

NOMENCLATURE

Sets

$n \in N$	Set of buses
$g \in G$	Set of generators
$k \in K$	Set of (physical) transmission line indices
$\langle i,j \rangle \in L$	Set of (notational) transmission lines with power flow from bus i to bus j where $i < j$
$\langle i,j \rangle \in L'$	Set of (notational) transmission lines with power flow from bus i to bus j where $i > j$

Decision Variables

V_n	Voltage magnitude at bus n
θ_n	Voltage phase angle at bus n
p_n^g, q_n^g	Active/reactive power output at bus n
ϕ_n	Change in voltage magnitude at bus n
p_{ij}, q_{ij}	Active/reactive power flow from bus i to bus j
q_{ij}^{Δ}	Reactive power flow from bus i to bus j coupled with change in voltage magnitude
$\hat{c}_k^{\ell}, \hat{c}_{ij}^b$	Cosine approximation for transmision line k/bus par $\langle i,j \rangle \in L \cup L'$
z_k	Switch status for transmission line k
l_n	Fraction of demand satisfied at bus n (equal power factor)
l_n^p, l_n^q	Fraction of active/reactive power demand satisfied at bus n
V_n^-, V_n^+	Vilation of lower/uper voltage magnitude bounds at bus n

f_{ij}^i, f_{ij}^j	Branch flow violation at bus i/j corresponding to transmission line $\langle i, j\rangle$
f_{ij}	Total branch flow violation for transmission line $\langle i, j\rangle$

Parameters

$\lvert\tilde{V}_n^t\rvert$	Voltage magnitude setpoint at bus n
$V_n^{\min}, V_n^{\max}$	Min/max voltage magnitude at bus n
p_n^d, q_n^d	Active/reactive power demand at bus n
$p_n^{g\ \min}, p_n^{g\ \max}$	Min/max active power output at bus n
$q_n^{g\ \min}, q_n^{g\ \max}$	Min/max reactive power output at bus n
$\tilde{\theta}_n$	Voltage phase angle setpoint at bus n
$\theta^{\max}$	Maximum voltage phase angle difference
g_k, b_k	Real/imaginary admittance components for transmission line k
S_k	MVA rating for line k
h_c, h_a	Number of hyperplanes used for cosine/apparent power approximation
$M_k^p, M_k^q, M_k^{q^\Delta}$	Big-M values for active power, reactive power, and reactive power coupled with change in voltage magnitude for line k
J	Number of transmission switching actions
d	Spacing of hyperplanes for cosine approximation $d = 2\theta^{\max}/(h_c+1)$
$k[i, j]$	Mapping of node pair $\langle i, j\rangle \in L \cup L'$ to the corresponding line index k

Terminology

ACPF	AC power flow
ACPF-VR-TS	AC power flow optimization model for violation reduction with transmission switching
ACOTS	AC optimal transmission switching

CBVE	Closest branch to the violation element heuristic
CTS	Corrective transmission switching
DCOTS	DC optimal transmission switching
ECDF	Empirical Cumulative Distribution Function
LPAC-s	Linear programming model for AC power flows with reactive power and voltage magnitude constraints used for power system restoration with transmission switching
MINLP	Mixed-integer nonlinear program
MIP	Mixed-integer linear program
MIPAC	Mixed-integer programming model for AC power flows
TS	Transmission switching

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION AND LITERATURE REVIEW [1]

Optimal operation of the electric power grid is of critical importance to maintain the integrity of the industrial, medical, and defense entities upon which the population relies. The application of optimization techniques in the context of power systems can be considered in two veins: long-term planning (i.e, months or years of planning horizon) and immediate real-time operations (i.e., within the next 15 minutes). The primary difference between these two classes of problems is, of course, the limitation of time to obtain solutions. While long-term planning projects are conducted offline and therefore have few constraints regarding solution time, real-time operational projects are required to find solutions quickly. These time constraints naturally shape the research regarding techniques designed to address these two classes of problems.

In the case of long-term planning projects, relatively recent developments of novel optimization techniques allow researchers to address problems of substantial size. For example, recent research efforts such as [1] have tackled optimization problems on power grids such as the Electric Reliability Council of Texas (ERCOT), which contains thousands of buses. Projects such as this address the size of real-world power systems, but typically do so through linearizations which fail to estimate the true complexity of the power grid. Linearizations such as this have proven sufficient for long-term planning scenarios. In contrast, recent research efforts on real-time operations are typically only tested on small-scale networks, consisting of hundreds of buses or fewer, or rely on heuristics in order to satisfy the necessary time requirements at the expense of the quality of solution.

One particularly important scenario within which short-term power grid operations are

[1]A portion of the text within this chapter was excerpted and lightly edited with permission from "Transmission-Line Switching for Load Shed Prevention via an Accelerated Linear Programming Approximation of AC Power Flows," *IEEE Transactions on Power Systems*, Vol. 35, No. 4, pp. 2575-2585, 2020. ©2020 IEEE.

conducted is contingency response. Events such as the failure of power grid components, weather-related accidents, or intentional malicious attacks can cause portions of the power grid to fail. Such failures can result in substantial harm to the grid itself and to the downstream entities which utilize the grid. It is for this reason that contingency analysis is regularly conducted to evaluate the robustness of the power grid and to develop actions to alleviate any deleterious effects stemming from contingencies. One corrective measure which system operators have at their disposal is transmission switching, also known as topology control.

In the past, the power grid has primarily been modeled and operated using a fixed configuration. Under such a framework, system operators exert control over the flow of power through the grid only via unit commitment (i.e., which generators are active) and generation dispatch (i.e., how much power is produced at each generator) decisions. However, system operators have another method of control at their disposal, transmission switching. Transmission switching refers to the technique of physically switching transmission lines in and out of the grid via the use of circuit breakers. Transmission switching has been shown to produce a myriad of benefits which are otherwise inaccessible to system operators. The benefit which we seek to leverage within this dissertation is the ability to more efficiently recover from contingency events.

Unfortunately, independent system operators have yet to implement transmission switching at a large scale. The predominant explanation for this is that the underlying optimization model which guides this process is computationally intractable. Specifically, the AC optimal transmission switching (ACOTS) problem, the optimization model which most accurately models the true behavior of the AC power grid when transmission switching is available, is a mixed-integer nonlinear program which is highly nonconvex. Thus, even using state-of-the-art optimization techniques, solutions to the ACOTS within the inherent time constraints enforced by real time operations remain out of reach, particularly for large-scale power grids. To address this, the bulk of the transmission switching literature uses either DC-based models or only validate their algorithms on unrealistically small test systems. These facts are problematic for several reasons. First, DC-based models have been shown in the literature to be highly inaccurate, especially within the context of short-term operations [2]. In addition, small test cases fail to approximate both the size and complexity of real-world power grids. This renders conclusions regarding the real-world impact of models and algorithms unclear at best. Finally, it is apparent in the literature that, given the performance of most models and algorithms on these small test cases, most approaches simply cannot satisfy the stringent requirements of real-time contingency response operations for real-world systems, motivating the need for more efficient techniques.

The principal objective of this dissertation is to develop models and solution approaches that outperform existing state-of-the-art methods for real-time power systems operations problems. Specifically, we seek to develop solution approaches that leverage the flexibility of the power grid such that grid operations are more efficient and may better respond to contingency events. That is, when a portion of the power grid fails, the approaches described in this dissertation seek AC optimal or near-optimal transmission switching solutions which are obtainable in real time. To that end, we develop two solution approaches – one mixed-integer linear program and one data mining technique – which respectively address two variants of the optimal transmission switching problem specifically within the context of contingency response.

The first problem we consider is the post-contingency AC optimal transmission switching problem for load shed prevention. This problem is an extension of the traditional optimal power flow problem to account for transmission switching in post-contingency operations. This is a particularly challenging problem because of the nonconvexities of AC power flow, the integrality constraints associated with transmission switching, and the limited time associated with post-contingency operations. As discussed herein, previous approaches to this and other similar problems are either unreliable in terms of the selection of topology or far too slow for real-world implementation. Our approach, which is summarized in Section 1.3 and explained in detail in Section 2, develops a mixed-integer linear model which is competitive with the state-of-the-art in solution time, almost always identifies AC optimal transmission switching actions, and shows strong potential to scale to problems of real-world size.

The second problem we consider is the AC power flow optimization model for violation reduction with transmission switching. This model seeks to minimize violations stemming from contingency events within AC power flow solutions via the use of transmission switching. This problem is particularly difficult because of, similar to the first problem we study, the nonconvexities associated with AC power flow, integrality constraints, and tight time constraints associated with the specific problem context. In addition, we study this problem with real-world large-scale power system data, further increasing the need for computationally efficient approaches. Our approach, which is summarized in Section 1.3 and detailed in Chapter 3, develops a data mining heuristic which addresses three key issues from the literature regarding practical implementation of transmission switching. Moreover, our method significantly outperforms existing approaches to the problem and demonstrates the value of data mining approaches to power systems problems.

The remainder of this chapter formalizes the problems upon which our approaches are constructed. While the following sections introduce the reader to much of the important literature on the subject, we note that several excellent studies have been written which detail the progression of optimization modeling and techniques used to model power systems problems such as optimal power flow (OPF), economic dispatch, and optimal transmission switching problems as well as their variants [3, 4, 5, 6, 7, 8, 9, 10, 11]. While there are several surveys on OPF and economic dispatch, to the author's knowledge, [11] is the only literature review on transmission switching which we herein study. The following sections of this chapter discusses specific models, solution techniques, and properties which are relevant to the research developed within this dissertation. We conclude the chapter with a summary of our research contributions.

1.1 Optimal Power Flow

The AC optimal power flow (ACOPF) problem is one of the central problems in power systems optimization. The ACOPF seeks to optimize a given objective (e.g., generation fuel cost or transmission line loss) by routing power flow through an electrical network while satisfying network flow and electrical equipment operational constraints [4]. Since the introduction of the ACOPF in the seminal work by the authors in [12], the ACOPF has been of particular interest to researchers because of its theoretical implications (the ACOPF is highly nonconvex) and its practical implications due to the use of ACOPF solutions in daily power grid operations.

Regarding the formulation of the ACOPF, there exist two equivalent traditional characterizations: one with polar voltage coordinates and another with rectangular voltage. There are merits to each coordinate system which are not held by the other. For example, rectangular coordinates yield constant second partial derivatives, significantly reducing the computational cost required to derive a direction of improvement within the optimization algorithm. Similarly, and as discussed in detail below, using polar voltage coordinates results in a coupling between changes in voltage magnitude and reactive power as well as between changes in voltage phase angle and active power. Given the relative strengths and weaknesses of each formulation, it would appear that a categorization of methods by choice of coordinate system may provide insight into how researchers attempt to leverage the strengths and mitigate the weaknesses. However, this is not the case, as the overwhelming majority of research on the ACOPF utilizes polar voltage coordinates. There are a few notable exceptions to this fact, including [13, 14, 15, 16, 17].

There have been a myriad of other solution approaches applied to the ACOPF. While each of these methods have been used successfully, because of the difficulty of the problem and the known weaknesses of each solution approach, no method has proven superior on all classes of problems. Since the seminal paper on optimal power flow by [12], the ACOPF has been addressed by methods such as gradient methods [18, 19, 20], Newton-based methods [21, 22, 23], trust-region approaches [24, 25], and the broad class of interior-point methods [13, 16, 26]. We note that [4] provides a thorough discussion of each of the above methods, their associated strengths and weaknesses, and noteworthy citations including those cited above.

There is one final class of approaches that merits specific mention because of its leveraging of the polar coordinate system. This method is known as decoupled optimal power flow. Consider the AC power flow equations in polar voltage coordinates for a notional bus i.

$$P_i = V_i \sum_{j \in N} g_{ij} V_j \cos(\theta_i - \theta_j) + b_{ij} V_j \sin(\theta_i - \theta_j) \tag{1.1}$$

$$Q_i = V_i \sum_{j \in N} g_{ij} V_j \sin(\theta_i - \theta_j) - b_{ij} V_j \cos(\theta_i - \theta_j) \tag{1.2}$$

In the polar coordinate system, voltage phase angles of admittance matrix elements are typically near $90°$ or $-90°$. In addition, voltage phase angles of adjacent buses are typically similar. As such, using straightforward substitution, the reader can see that changes in active power are strongly coupled to changes in voltage phase angle. Similarly, changes in reactive power are strongly coupled to changes in voltage magnitude. There have been several classes of optimization methods which have been applied to the solution of the decoupled optimal power flow problem. A few noteworthy examples include [27, 28, 29].

More recent research on optimal power flow has focused on developing convex relaxations to the formulation in an effort to further accelerate the solution process. The most well-studied linear approximation to the ACOPF is the DC approximation [30]. This model, as we discuss in Chapter 2, has been shown to be a crude approximation to the ACOPF and has motivated more accurate linear approaches such as [31], which we herein study. There have been two additional classes of convex relaxations to the ACOPF which are particularly noteworthy: the semidefinite relaxation [32], and the second-order cone programming relaxation [33]. These relaxations are discussed thoroughly in the two-part survey [9, 10].

1.1.1 AC Optimal Power Flow (ACOPF)

Consider a power grid with a set of buses N, a set of buses with electric power generators $G \subseteq N$, and a set of transmission lines L. Further assume that there are electric power demands, in the form of both active and reactive power, at each bus. As discussed previously, the goal of the ACOPF is to optimize a given objective by routing power flow through an electrical network while satisfying network flow and electrical equipment operational constraints [4].

As mentioned previously, there are two traditional equivalent formulations for the ACOPF: one with polar voltage coordinates, consisting of voltage magnitude V and voltage phase angle θ, and another with rectangular voltage coordinates, consisting of real and reactive voltage components e and f, respectively. We can first consider the ACOPF using polar voltage coordinates.

$$\min \sum_{g \in G} C_g(P_g) \tag{1.3}$$

subject to

$$P_i^g - P_i^d = \sum_{j \in N} p_{ij} \qquad i \in N \tag{1.4}$$

$$Q_i^g - Q_i^d = \sum_{j \in N} q_{ij} \qquad i \in N \tag{1.5}$$

$$p_{ij} = g_{ij} V_i^2 - V_i V_j (g_{ij} \cos(\theta_i - \theta_j) + b_{ij} \sin(\theta_i - \theta_j)) \qquad \langle i, j \rangle \in L \tag{1.6}$$

$$q_{ij} = -b_{ij} V_i^2 - V_i V_j (g_{ij} \sin(\theta_i - \theta_j) - b_{ij} \cos(\theta_i - \theta_j)) \qquad \langle i, j \rangle \in L \tag{1.7}$$

$$V_i^{\min} \leq V_i \leq V_i^{\max} \qquad i \in N \tag{1.8}$$

$$P_i^{g\ \min} \leq P_i^g \leq P_i^{g\ \max} \qquad i \in G \tag{1.9}$$

$$Q_i^{g\ \min} \leq Q_i^g \leq Q_i^{g\ \max} \qquad i \in G \tag{1.10}$$

$$p_{ij}^2 + q_{ij}^2 \leq S^2 \qquad \langle i, j \rangle \in L \tag{1.11}$$

Objective (1.3) traditionally minimizes a cost function which is typically quadratic (e.g., generation fuel cost). Constraints (1.4) and (1.5)x enforce active and reactive power balance, respectively. Constraints (1.6) and (1.7) characterize active and reactive power flow, respectively. Constraints (1.8)-(1.11) enforce bounds on voltage magnitude, active power generation, reactive power generation, and apparent power, respectively.

There exists an equivalent OPF formulation using rectangular coordinates, as follows.

$$\min \sum_{g \in G} C_g(P_g) \tag{1.12}$$

subject to

$$p_{ij} = g_{ij}(e_i^2 + f_i^2) - g_{ij}(e_i e_j + f_i f_j) + b_{ij}(e_i f_j - e_j f_i) \qquad \langle i,j \rangle \in L \tag{1.13}$$

$$q_{ij} = -b_{ij}(e_i^2 + f_i^2) + b_{ij}(e_i e_j + f_i f_j) + g_{ij}(e_i f_j - e_j f_i) \qquad \langle i,j \rangle \in L \tag{1.14}$$

$$(V_i^{\min})^2 \leq e_i^2 + f_i^2 \leq (V_i^{\max})^2 \qquad i \in N \tag{1.15}$$

$$(1.4), (1.5), (1.9)\text{-}(1.11)$$

Objective (1.12) is identical to objective (1.3); it minimizes a cost function. Constraints (1.13) and (1.14) characterize active and reactive power flow, respectively, using rectangular coordinates for voltage. Constraint (1.15) bounds the voltage magnitude. As stated previously, there are various strengths and weaknesses of each formulation. As such, that has not to this point been a combination of formulation and solution methodology which dominates all others.

1.1.2 DC Optimal Power Flow (DCOPF)

One important and frequently-studied approximation to the ACOPF is known as the DC optimal power flow (DCOPF) model. The DCOPF can be derived from the ACOPF using several properties of AC power systems. Specifically, there are three assumptions which are inherent to the DCOPF.

1. Resistance of transmission lines is negligible in comparison to reactance (i.e., $g_{ij} \approx 0$)

2. Voltage magnitudes are sufficiently close to 1 (i.e., $V_i \approx 1$)

3. Adjacent phase angle difference are sufficiently close to zero (i.e., $\theta_i - \theta_j \approx 0$). This

assumption leads to the frequently-used trigonometric approximation $\sin(\theta_i - \theta_j) \approx \theta_i - \theta_j$.

Applying these three assumptions to Constraints (1.6) and (1.7), Constraint (1.7) is eliminated altogether and Constraint (1.6) becomes

$$p_{ij} = b_{ij}(\theta_i - \theta_j) \tag{1.16}$$

Because the above assumptions eliminate reactive power flow on all transmission lines, we can also drop Constraints (1.5) and (1.10) which enforce reactive power balance and reactive power generation, respectively. This yields the following model, known as the DCOPF.

$$\min \sum_{g \in G} C_g(P_g) \tag{1.17}$$

subject to

$$p_{ij} = b_{ij}(\theta_i - \theta_j) \qquad \langle i,j \rangle \in L \tag{1.18}$$

$$-S_{ij} \leq p_{ij} \leq Sij \qquad \langle i,j \rangle \in L \tag{1.19}$$

$$(1.4), (1.9)$$

Objective (1.17), similar to previously discussed models, seeks minimize some cost function. However, unlike the models discussed so far, this objective is generally linear. Constraint (1.18) characterizes active power flow. Constraint (1.19) bounds the minimum and maximum active power flow on a given transmission line.

The DCOPF model is widely used in practice. In particular, the DCOPF and other similar models are useful because they are very fast to solve. Moreover, they have shown to be highly useful in long-term decision making. However, because of their inability to accurately approximate the complexities of the ACOPF, the DCOPF and similar models have performed poorly for short-term decision making, particularly for post-contingency transmission switching which we herein study.

1.2 Transmission Switching

Up to this point, each of the discussed models contain the inherent assumption that the power grid has a fixed topology. That is, each of the transmission lines are closed (i.e., operational) and can only remain as such. However, technology exists which allows for the topology of the grid to be altered. Such actions are referred to as transmission switching (TS) or, equivalently, topology control (TC). The seminal work on optimal TS [34] showed that transmission switching can result in significant economic benefit. Several other research efforts support this conclusion [35, 36, 37]. In subsequent years, other benefits which TS may obtain have been found. Such benefits include reliability [38], ability to recover from contingencies [39], and reduction in softer constraints such as thermal and voltage violation [40].

More recent work has focused on variants of the optimal TS problem. One problem of note is the security-constrained ACOTS, which was addressed via decomposition by the authors in [41]. The authors in [42] studied the interplay between TS and flexible AC transmission system (FACTS) devices. Stochastic unit commitment was studied in [43]. In a related work, the authors in [44] studied optimal TS in the presence of stochastic wind generation. The authors in [45] studied how security analysis impacts the physical impacts related to implementation of TS actions. Finally, reliability related to TS was studied in [46].

Other recent research on transmission switching, similar to recent work on the ACOPF problem, has sought to tighten the OTS problem or derive strong relaxations. The authors in [47] and [48] develop novel formulations for the DCOTS and ACOTS, respectively. In

[49], the convex relaxation of the ACOTS is tightened. Several recent works have also studied conic relaxations of the ACOTS, including [50, 51, 52].

As discussed herein, the most accurate model for power systems optimization is the ACOPF. As such, the most accurate transmission switching model is the ACOTS, which is a mixed-integer nonlinear programming (MINLP) model. While the accuracy of the ACOTS is well known, because of the difficulty in solving it, the bulk of transmission switching research focuses on the DC optimal transmission switching (DCOTS) model. Unfortunately, the DCOTS has proven problematic in short-term decision making. This is studied in [2], who showed that solutions identified by the DCOTS may be either AC infeasible or result in negative impact.

Despite its difficulty, there has been some recent work which attempts to address the ACOTS directly. Probably the most effective heuristic in transmission switching, the line profit heuristic [37], was extended from its DC roots to the ACOTS in [53]. The authors in [54] studied a decoupled approach wherein the ACOTS is decoupled into DCOTS and ACOPF models. Two very noteworthy approaches, mixed-integer second order cone programming [48] and decomposition [55] have also proven highly useful in addressing the ACOTS.

1.2.1 DC Optimal Transmission Switching (DCOTS)

The DCOPF model as given in equations 1.17-1.19 can be extended to account for transmission switching. We herein utilize a binary variable z_{ij} which takes on value 1 when the transmission line between buses i and j is operational, and 0 otherwise. We can then characterize the DCOTS as follows.

$$\min \sum_{g \in G} C_g(P_g) \tag{1.20}$$

subject to

$$b_{ij}(\theta_i - \theta_j) - p_{ij} + (1 - z_{ij})M \geq 0 \qquad \langle i, j \rangle \in L \tag{1.21}$$

$$b_{ij}(\theta_i - \theta_j) - p_{ij} - (1 - z_{ij})M \leq 0 \qquad \langle i, j \rangle \in L \tag{1.22}$$

$$- S_{ij}z_{ij} \leq p_{ij} \leq S_{ij}z_{ij} \qquad \langle i, j \rangle \in L \tag{1.23}$$

$$(1.4), (1.9)$$

As in all previously discussed models, (1.20) seeks to minimize some cost function. Constraints (1.21) and (1.22) characterize active power flow through the network using big-M notation. Finally, constraint (1.23) bounds active power flow, accounting for transmission switching.

As discussed above and in subsequent sections, the DCOTS has proven highly useful because of its quick solution time and accuracy for long term planning. However, recent work has shown that transmission switching actions identified by the DCOTS are frequently either infeasible or result in negative implications when implemented in the AC system. As such, transmission switching models with higher fidelity are imperative going forward.

1.2.2 AC Optimal Transmission Switching (ACOTS)

Similar to the DCOPF, the ACOPF can be extended to account for transmission switching. Note that below we utilize polar coordinates for voltage variables. However, similar modifications to the ACOPF with rectangular coordinates may be performed, resulting in an equivalent ACOTS model.

$$\min \sum_{g \in G} C_g(P_g) \tag{1.24}$$

subject to

$$p_{ij} = z_{ij}[g_{ij}V_i^2 - V_iV_j(g_{ij}\cos(\theta_i - \theta_j) + b_{ij}\sin(\theta_i - \theta_j))] \qquad \langle i,j \rangle \in L \tag{1.25}$$

$$q_{ij} = z_{ij}[-b_{ij}V_i^2 - V_iV_j(g_{ij}\sin(\theta_i - \theta_j) - b_{ij}\cos(\theta_i - \theta_j))] \qquad \langle i,j \rangle \in L \tag{1.26}$$

$$p_{ij}^2 + q_{ij}^2 \leq z_{ij}S_{ij}^2 \qquad \langle i,j \rangle \in L \tag{1.27}$$

$$(1.4), (1.5)(1.8)\text{-}(1.10)$$

Objective (1.24) minimizes some notional cost function. Constraints (1.25) and (1.26) model active and reactive power flow accounting for transmission switching. Finally, constraint (1.27) enforces apparent power constraints while accounting for transmission switching.

While the above formulation accurately models the true behavior of transmission switching on the AC system, it is entirely impractical to solve. Specifically, because of the non-convexity and nonlinearity of the ACOPF, when combined with the integrality constraints introduced by transmission switching, the above formulation is simply intractable when addressed directly. As such, it is critical that relaxations are developed which significantly improve solution time while simultaneously having high fidelity with the ACOTS.

1.3 Contributions

The remainder of this chapter provides a brief overview of the remaining chapters in this dissertation.

Chapter 2 discusses our optimization model for AC optimal transmission switching for load shed prevention. In particular, we substantially accelerated an existing model in the literature [56] which is based upon a high-fidelity approximation to the ACOPF [31]. This model gave very strong performance in identifying AC transmission switching actions, but was prohibitively slow in practice in terms of solution time – slower than an exhaustive search of transmission switching actions. We implemented three key constraint modifications (1) a relaxation for the apparent power constraints, (2) a novel relaxation for the cosine approximation, and (3) a relaxation to the implicit power factor constraint. Using these three constraint modifications, we dramatically decreased solution time while increasing accuracy in terms of the ability to identify optimal transmission switching actions. In addition, we note that we are one of the only works on transmission switching which systematically validate the quality of their TS actions.

Chapter 3 discusses our data mining approach to address a different variant of the transmission switching problem: the post-contingency AC power flow model for violation reduction with transmission switching. We herein develop a data mining approach using imbalanced data classification techniques, to address this problem. Our approach, to the author's knowledge, is the first true data mining technique to classify strong transmission switching actions. Moreover, our methodology is unique in that it addresses three key issues related to transmission switching – computational complexity, impact on large-scale systems, and impact on AC systems – and can be directly combined with existing approach to address a key fourth issue: transient stability. Most importantly, our method substantially outperformed existing state-of-the-art approaches applied to this problem and demonstrates the power of data mining approaches when applied to power systems problems.

Section 4 concludes the work by summarizing the primary contributions we develop in this dissertation.

2. TRANSMISSION-LINE SWITCHING FOR LOAD SHED PREVENTION VIA AN ACCELERATED LINEAR PROGRAMMING APPROXIMATION OF AC POWER FLOWS [1]

2.1 Introduction

With increasing emphasis on the robustness of our critical electrical infrastructure, it is imperative that new techniques are developed in order to protect the national power grid. Component failure, weather-related events, and intentional malicious attacks on the grid can cause portions of the system to fail, resulting in potentially substantial harm to the system. For this reason, real-time contingency analysis is executed regularly to evaluate the robustness of the power grid and identify corrective measures should portions of the transmission system fail. One potential corrective measure is topology control, also known as transmission switching (TS).

The power grid has historically been characterized as a static asset with fixed configuration [57]. As such, system operators have traditionally controlled the flow of power only by altering the dispatch at power plants. However, power flow can also be altered by changing the topology of the energy grid; that is, by switching transmission lines in or out of the grid. Previously, the authors in [40] stated that corrective TS (CTS) is currently being used in real-world contingency response operations in limited capacity. Furthermore, [58, 59, 60, 39] have shown that CTS is a valuable technique in recovery operations to prevent load shed. An additional hindrance to the inclusion of TS in recovery operations models is the computational difficulty in solving mixed-integer nonlinear programs (MINLPs), such as the AC optimal transmission switching (ACOTS) problem.

This paper accelerates an existing mixed-integer linear optimization model, the linear programming model for AC power flows with reactive power and voltage magnitude constraints used for power system restoration with TS (LPAC-s), in an effort to more quickly identify AC TS actions to maximize load shed prevention. Specifically, we modify three computationally-costly constraints which allow a significant speedup with no measurable effect on accuracy. The resulting model is henceforth referred to as the mixed-integer programming model for AC power flows (MIPAC). This is an important line of research, as TS actions identified by models such as the DC optimal transmission switching (DCOTS) model have been shown to be AC infeasible or even produce negative outcomes when implemented in the AC system [2]. Moreover, researchers have subtantially more resources to solve mixed-integer linear programs (MIPs) compared to MINLPs. In addition to resources, it is particularly salient to note that linear models in general scale better than nonlinear models because of the computational efficiency of solution approaches [61].

We note to the reader that, in previous works ([39, 57]), the term load shed recovery has been used to characterize minimizing load shed stemming from a contingency event. We argue that such terminology is inaccurate. Specifically, in a practical context, load shed recovery refers to the actions which help recoup load shed after corrective measures have been implemented. On the other hand, the actions described herein and those developed in

[39] and [57] refer to actions which immediately follow a contingency event. We therefore herein utilize the term load shed prevention, which describes actions for minimizing the load which would have to be shed absent corrective mechanisms.

The specific contributions of this work are as follows. First, we take an existing model for AC load shed prevention [56] and, through a series of constraint modifications, significantly decrease the necessary runtime to solve it. Moreover, we do so with no decrease in the accuracy of the transmission switching actions which are identified by the model. Importantly, we remove or reduce the degree of piecewise-linearizations, the benefits of which are two-fold. First, piecewise-linearizations are prone to degeneracy [62]; second, such linearizations also risk potential pitfalls due to computational instability [63]. Our second contribution is that we thoroughly demonstrate the accuracy of our model (and the base model [56]) via a test against an exhaustive search. This more accurately demonstrates the practical implications of our model and shows that our model can be used successfully in realistic contingency response operations. Indeed, excluding [55, 48] we are unaware of any works which approximate the ACOTS that systematically validate the quality of their TS solutions.

The remainder of this paper is organized as follows. Section 2.2 presents a literature review discussing the benefits of TS and techniques designed to address the different optimal TS problem variants. Section 2.3 describes the base model herein studied for load shed prevention, discusses three key constraint modifications, and proposes our new model. Section 2.4 describes the experimental setup and the test case used to study the discussed models. Section 2.5 first motivates our proposed model by showing that solving the LPAC-R-V extension for load shed prevention with TS is slower than an ACOPF-based exhaustive search of switching actions. Second, we show that the constraint modifications on which MIPAC is based cause no measurable decrease in the ability to identify strong TS solutions. Third, we study the impact of our constraint modifications on MIPAC's solution time when solving for a single switching action. Next, we show that MIPAC can be used successfully to identify switching actions in lightly loaded instances. Our final analysis shows the performance of MIPAC when multiple switching actions are allowed. Section 2.6 concludes the work.

2.2 Literature Review

As previously stated, the power grid has traditionally been characterized as a static structure with fixed topology. However, altering the topology of the grid has been shown to produce considerable benefits. The seminal paper on optimal TS [34] showed that such actions can result in substantial cost savings. This finding has been supported by several subsequent works [35, 36, 37]. Since [34], researchers have identified many other benefits stemming from TS. Examples include thermal and voltage violation reduction [40], load shed prevention [39, 57], and system reliability [38].

Recent research regarding TS has focused on several different TS problem variants. The authors in [45] showed the importance of security analysis prior to implementing switching actions. Decomposition was used in in [41] to address security-constrained ACOTS. The authors in [42] studied the interdependence of TS and flexible AC transmission system (FACTS) devices in the context of unit commitment. The authors in [43] studied stochastic unit commitment with TS. A column generation method was developed by the authors in [44] to identify optimal TS actions with stochastic wind generation. The authors in [46] studied the use of TS to maintain N-1-1 reliability. The authors in [64] utilized a decomposition approach to address seasonal TS. Finally, a decentralized TS algorithm was developed in [65] for congestion reduction.

Several relatively recent works have also made efforts to either tighten the OTS problem formulation or to develop relaxations. The authors in [47] and [48] focus on improving the formulation of DCOTS and ACOTS, respectively. The authors in [49] developed a bound strengthening method to tighten the convex relaxation of the problem. Novel conic relaxations have also shown promise such for OTS [50, 51, 52].

The most accurate TS optimization model is ACOTS, which is an MINLP. Because of our inability to efficiently solve MINLPs, the majority of TS research has focused on the DCOTS, which is significantly more computationally tractable. Such models are useful for two reasons: they are very fast to solve, and they have shown usefulness in long-term planning. On the other hand, such models are crude approximations to the ACOTS problem and are prone to errors in short-term applications. Indeed, as previously stated, the authors in [2] showed that DCOTS switching actions may be AC infeasible or produce negative AC outcomes.

Because of the difficulty in solving the ACOTS, there have been few efforts which attempt to address the problem directly. The line profit heuristic [37] was extended in [53] to address the ACOTS with an objective of reducing generation fuel cost. In addition, [48] developed mixed-integer second order cone programming relaxations for the ACOTS problem. Decomposition approaches, such as those in [55], have also proven useful in solving the ACOTS. The authors in [54] decoupled ACOTS into DCOTS and ACOPF. An ACOTS algorithm using sucessive ACOPF solutions was developed in [66]. However, little attention has been applied to AC TS problems for recovery applications such as load shed prevention, which we herein address.

2.3 Load Shed Prevention Optimization Models

In the context of post-contingency load shed prevention, optimization models must be able to provide solutions within tight time constraints. Thus, models which approximate the behavior of the AC power system may be utilized in order to provide solutions quickly. In particular, because of the shortcomings of existing linear approximations, linear models which more closely approximate the AC power system should continue to be developed.

Specific to this work is the need for models which approximate the ACOTS problem for load shed prevention (ACOTS-LS). This need was addressed, in part, by the authors in [31], who developed the linear programming model for AC power flows with reactive generation and voltage magnitude constraints (LPAC-R-V) and its extension LPAC-s. LPAC-R-V, while consisting of linear constraints, models the behavior of the AC power system much more accurately than existing linear models. The remainder of this section is organized as follows. Section 2.3.1 presents LPAC-R-V. Section 2.3.2 presents LPAC-s , the extension of LPAC-R-V for transmission switching. Note that, when binary variables in this model are fixed to one, LPAC-s reduces to LPAC-R-V as presented in Section 2.3.1. Section 2.3.3 discusses three constraint modifications upon which our accelerated model is developed. Finally, Section 2.3.4 presents MIPAC, our accelerated TS model for load shed prevention.

2.3.1 LPAC-R-V

The LPAC-R-V model for load shed prevention as described by [31] is as follows.

$$\max \sum_{n \in N} l_n p_n^d, \sum_{\langle i,j \rangle \in L} \hat{c}_{ij}^b \tag{2.1a}$$

subject to

$$p_i^g - l_i p_i^d = \sum_{\langle i,j \rangle \in L} p_{ij} \qquad i \in N \tag{2.1b}$$

$$q_i^g - l_i q_i^d = \sum_{\langle i,j \rangle \in L} (q_{ij} + q_{ij}^{\Delta}) \qquad i \in N \tag{2.1c}$$

$$p_{ij} = \tilde{V}_i|^2 g_{k[i,j]} - |\tilde{V}_i||\tilde{V}_j|(g_{k[i,j]} \hat{c}_{ij}^b + b_{k[i,j]}(\theta_i - \theta_j)) \qquad \langle i,j \rangle \in L \cup L' \tag{2.1d}$$

$$q_{ij} = -|\tilde{V}_i|^2 b_{k[i,j]} - |\tilde{V}_i||\tilde{V}_j|(g_{k[i,j]}(\theta_i - \theta_j) - b_{k[i,j]} \hat{c}_{ij}^b) \qquad \langle i,j \rangle \in L \cup L' \tag{2.1e}$$

$$q_{ij}^{\Delta} = -|\tilde{V}_i| b_{k[i,j]}(\phi_i - \phi_j) - (|\tilde{V}_i| - |\tilde{V}_j|) b_{k[i,j]} \phi_i \qquad \langle i,j \rangle \in L \cup L' \tag{2.1f}$$

$$V_i^{\min} \leq |\tilde{V}_i| + \phi_i \leq V_i^{\max} \qquad i \in N \tag{2.1g}$$

$$p_n^{g\min} \leq p_n^g \leq p_n^{g\max} \qquad n \in G \tag{2.1h}$$

$$q_n^{g\min} \leq q_n^g \leq q_n^{g\max} \qquad n \in G \tag{2.1i}$$

$$-\theta^{\max} \leq \theta_i - \theta_j \leq \theta^{\max} \qquad \langle i,j \rangle \in L \cup L' \tag{2.1j}$$

$$0 \leq l_i \leq 1 \qquad i \in N \tag{2.1k}$$

$$0 \leq \hat{c}_{ij}^b \leq 1 \qquad \langle i,j \rangle \in L \cup L' \tag{1ℓ}$$

$$\hat{c}_{ij}^b \leq \cos(t \cdot d - \theta^{\max}) - \sin(t \cdot d - \theta^{\max})(\theta_i - \theta_j - t \cdot d + \theta^{\max})$$
$$t \in \{1, \ldots, h_c\}, \ \langle i,j \rangle \in L \cup L' \tag{2.1m}$$

$$p_{ij}^2 + (q_{ij} + q_{ij}^{\Delta})^2 \leq S_{k[i,j]}^2 \qquad \langle i,j \rangle \in L \cup L' \tag{2.1n}$$

Objective (2.1a) is lexicographic; it first maximizes the satisfied active power demand followed by the sum of cosine approximations. Constraints (2.1b) and (2.1c) enforce node balance for active and reactive power at each bus, respectively. Constraint (2.1d) models active power flow. Constraints (2.1e)-(2.1f) model reactive power flow in two separate components. First, q_{ij} is the component of reactive power which is independent to changes in voltage magnitude. Second, q^{Δ}_{ij} is the component of reactive power which is coupled with changes in voltage magnitude. Constraint (2.1f) is derived via a Taylor series expansion; we refer the reader to [31] for a detailed conversation on this derivation. Note that the model accounts for losses inherently by calculating power injection at each of the two buses on a transmission line for each constraint (2.1d)-(2.1f). As an example, for a transmission line between buses i and j, the active power flow injection at bus i is denoted by p_{ij}, the corresponding injection at bus j is p_{ji}, and the loss on this transmission line is $|p_{ij} + p_{ji}|$. Constraints (2.1g)-(1ℓ) enforce limits on voltage magnitude, active and reactive power generation, adjacent phase angle differences, the fraction of unmet demand, and the cosine approximation, respectively. Constraint (2.1m) defines a series of inequalities which characterizes the piecewise-linearization of the cosine function. Note that t denotes the index of the hyperplane in the piecewise-linearization and d is the uniform spacing (distance) between consecutive hyperplanes; specifically, given the desired number of hyperplanes (h_c) and the maximum phase angle differences ($\theta^{\max}$), d is calculated as follows, $d = 2\theta^{\max}/(h_c + 1)$. Finally, constraint (2.2m) enforces thermal constraints for each transmission line as a function of squared apparent power.

2.3.2 LPAC-s for Load Shed Prevention

The LPAC-s model for load shed prevention is described below. We note that one previous work used a similar model to address the restoration ordering problem, which is similar to TS. The following model, LPAC-s, is a straightforward simplification of the LPAC model for the restoration ordering problem proposed in [56]. In a long-term planning context where timing is not a major issue, LPAC-s performs reasonably well. However, to feasibly implement this model in a real-time context such as load shed prevention, LPAC-s requires acceleration; this fact is demonstrated in Section 2.5.

$$\max \sum_{n \in N} l_n p_n^d \tag{2.2a}$$

subject to

(2.1b),(2.1c) (2.2b),(2.2c)

$$p_{ij} - (1 - z_{k[i,j]}) M^p_{k[i,j]} \leq$$
$$|\tilde{V}_i|^2 g_{k[i,j]} - |\tilde{V}_i||\tilde{V}_j|(g_{k[i,j]}\hat{c}^b_{ij} + b_{k[i,j]}(\theta_i - \theta_j))$$
$$\leq p_{ij} + (1 - z_{k[i,j]}) M^p_{k[i,j]} \quad \langle i,j \rangle \in L \cup L' \tag{2.2d}$$

$$q_{ij} - (1 - z_{k[i,j]}) M^q_{k[i,j]} \leq$$
$$-|\tilde{V}_i|^2 b_{k[i,j]} - |\tilde{V}_i||\tilde{V}_j|(g_{k[i,j]}(\theta_i - \theta_j) - b_{k[i,j]}\hat{c}^b_{ij})$$
$$\leq q_{ij} + (1 - z_{k[i,j]}) M^q_{k[i,j]} \quad \langle i,j \rangle \in L \cup L' \tag{2.2e}$$

$$q^\Delta_{ij} - (1 - z_{k[i,j]}) M^{q^\Delta}_{k[i,j]} \leq$$
$$-|\tilde{V}_i| b_{k[i,j]}(\phi_i - \phi_j) - (|\tilde{V}_i| - |\tilde{V}_j|) b_{k[i,j]}\phi_i$$
$$\leq q^\Delta_{ij} + (1 - z_{k[i,j]}) M^{q^\Delta}_{k[i,j]} \quad \langle i,j \rangle \in L \cup L' \tag{2.2f}$$

(2.1g)-(2.1m) (2.2g)-(2.2m)

$$p^2_{ij} + (q_{ij} + q^\Delta_{ij})^2 \leq z_{k[i,j]} S^2_{k[i,j]} \quad \langle i,j \rangle \in L \cup L' \tag{2.2m}$$

Objective (2.2a) maximizes satisfied active power demand. Constraint (2.2d) models active power flow accounting for transmission switching. Constraints (2.2e)-(2.2f) model reactive power flow in two separate components, similar to LPAC-R-V, while accounting for transmission switching. Finally, constraint (2.2m) enforces thermal constraints for each transmission line as a function of squared apparent power. While the authors in [31] did state that apparent power can be modeled "via a polyhedral outer approximation similar to [$\hat{c}_{ij}$]" (which is essentially a piecewise-linear approximation), neither they nor the source they cited discussed how to linearize this constraint. It is particularly important to characterize an appropriate number of hyperplanes to balance solution speed and accuracy. While the authors in [31] fully discuss such matters regarding their cosine approximation, these items remain undiscussed regarding apparent power. Indeed, one of the contributions of this paper appears in Section 2.3.3, where we offer an elegant and efficient linearization which is extremely well-suited for this model.

Constraint (2.2m) constrains power flow as a function of squared apparent power while accounting for TS. As previously stated, the authors in [31] and [56] did not mention how to specifically linearize this constraint. We note that, in p, q space, the apparent power constraint (2.2m) can be modeled as a regular polygon. Therefore, for the purposes of comparison, we linearize constraint (2.2m) according to the piecewise-linear approximation as discussed in [67] as follows

$$
\begin{aligned}
&-\Big(\cos\Big(\frac{2\pi t}{h_A}\Big) - \cos\Big(\frac{2\pi(t-1)}{h_A}\Big)\Big)(q_{ij} + q_{ij}^{\Delta}) \\
&+\Big(\sin\Big(\frac{2\pi t}{h_A}\Big) - \sin\Big(\frac{2\pi(t-1)}{h_A}\Big)\Big)p_{ij} \le S_{k[i,j]} \sin\Big(\frac{2\pi}{h_A}\Big),
\end{aligned}
\tag{2.3}
$$

for each (notational) transmission line $\langle i,j\rangle \in L \cup L'$. Where $t \in \{1, \dots, h_A\}$ is the hyperplane index and h_A is the number of hyperplanes used. We include a brief derivation of this constraint below.

The following derivation is very similar to the derivation from [67], modified for our work. Consider a circle in p, q space centered at the point $(0, 0)$ with radius r. We can approximate this circle using an n-sided regular polygon. Suppose we have a point (p_1, q_1) which lies on the circumference of the circle. Without loss of generality, assume that $(p_1, q_1) = (r\cos(\frac{2\pi}{n}), r\sin(\frac{2\pi}{n}))$. The line connecting (p_1, q_1) with another point on the circle is as follows.

$$y - r\sin(\frac{2\pi}{n}) = \frac{\sin(\frac{2\pi}{n}) - \sin(\frac{2\pi i}{n})}{\cos(\frac{2\pi}{n}) - \cos(\frac{2\pi i}{n})}(x - r\cos(\frac{2\pi}{n})), \tag{2.4}$$

where i is integer-valued and $i \mod n \neq 1$. Then the line connecting any two adjacent points (p_t, q_t) and (p_{t-1}, q_{t-1}) can be characterized as follows

$$y - r\sin(\frac{2\pi t}{n}) = \frac{\sin(\frac{2\pi t}{n}) - \sin(\frac{2\pi(t-1)}{n})}{\cos(\frac{2\pi t}{n}) - \cos(\frac{2\pi(t-1)}{n})}(x - r\cos(\frac{2\pi t}{n})) \tag{2.5}$$

Straightforward algebra yields the following

$$(\sin(\frac{2\pi t}{n}) - \sin(\frac{2\pi(t-1)}{n}))x - (\cos(\frac{2\pi t}{n}) - \cos(\frac{2\pi(t-1)}{n}))y = r\sin(\frac{2\pi}{n}) \tag{2.6}$$

which, using $x = p_{ij}$, $y = q_{ij} + q_{ij}^{\Delta}$, and $r = S_{k[i,j]}$, can be straightforwardly translated to obtain constraint (2.3).

Note that, in the experiments described in Section 2.5, we utilize 13 hyperplanes to construct this approximation. This value was chosen through computational testing as the smallest value which produced reasonably accurate results in terms of TS candidate identification. Additionally, as recommended by the authors of LPAC-R-V [31] and [56], we utilized 20 hyperplanes to develop the piecewise-linear cosine approximation given in constraint (2.1m).

2.3.3 Accelerating LPAC-s

As demonstrated in Section 2.5, in terms of solution time, LPAC-s is inferior to an ACOPF-based exhaustive search of switching actions. To address this, this section proposes several constraint modifications to LPAC-s. Specifically, we replace one piecewise-linear constraint with a less computationally-burdensome approximation, reduce the degree of another, and modify a third constraint which we show is unnecessary for identifying strong TS actions. In particular, the constraint modifications which remove or reduce the piecewise-linear constraints in LPAC-s serve two important purposes: they decrease the chance of degeneracy [62] and potentially enhance numerical stability [63].

2.3.3.1 Apparent Power

Note that LPAC-R-V contains bi-directional components (e.g., p_{ij}, p_{ji}). Because these variables are opposite in sign and have similar magnitudes, they can be constrained utilizing the following linear outer-approximation constraints similar to the constraint described in [68]

$$-S_{k[i,j]} z_{k[i,j]} \leq p_{ij} \leq S_{k[i,j]} z_{k[i,j]} \tag{2.7a}$$

$$-S_{k[i,j]} z_{k[i,j]} \leq q_{ij} + q^{\Delta}_{ij} \leq S_{k[i,j]} z_{k[i,j]} \tag{2.7b}$$

$$-\sqrt{2} S_{k[i,j]} z_{k[i,j]} \leq p_{ij} + q_{ij} + q^{\Delta}_{ij} \leq \sqrt{2} S_{k[i,j]} z_{k[i,j]} \tag{2.7c}$$

$$-\sqrt{2} S_{k[i,j]} z_{k[i,j]} \leq p_{ij} - q_{ij} - q^{\Delta}_{ij} \leq \sqrt{2} S_{k[i,j]} z_{k[i,j]}, \tag{2.7d}$$

for each (notational) transmission line $\langle i, j \rangle \in L \cup L'$. We provide a brief derivation of this constraint below.

Given the traditional form of the apparent power constraint $p^2 + q^2 \leq S^2$, moving everything to the left hand side, we can write the constraint using the function $f(p, q)$ as follows

$$f(p, q) = p^2 + q^2 - S^2 \leq 0$$

The first-order Taylor series expansion is as follows.

$$f(p, q) = 2\hat{p}(p - \hat{p}) + 2\hat{q}(q - \hat{q}) + \hat{p}^2 + \hat{q}^2 - S^2 \leq 0 \tag{2.8}$$

Starting at the top of the circle defined by the apparent power constraint, working clockwise, we then evaluate equation (2.8) at the following evenly-spaced boundary points: $(0, S), (\sqrt{\frac{S^2}{2}}, \sqrt{\frac{S^2}{2}}), (S, 0)$, and $(\sqrt{\frac{S^2}{2}}, -\sqrt{\frac{S^2}{2}}))$

$$
\begin{gathered}
f(0, S) = q - S \leq 0 \\
f(\sqrt{\frac{S^2}{2}}, \sqrt{\frac{S^2}{2}}) = p + q - \sqrt{2}S \leq 0 \\
f(S, 0) = p - S \leq 0 \\
f(\sqrt{\frac{S^2}{2}}, -\sqrt{\frac{S^2}{2}}) = p - q - \sqrt{2}S \leq 0
\end{gathered}
$$

These four constraints correspond to the right hand side of equation set (2.7). The constraints on the left hand side can be straightforwardly obtained by evaluating equation (2.8) at the remaining evenly-spaced points on the circle defined by the apparent power constraint.

2.3.3.2 Cosine Approximation

The piecewise linear formulation of the cosine function defined in constraint (2.1m) can cause a substantial computational burden on the solver. To ease this computational burden, we herein propose a new non-piecewise linear cosine approximation. Note that, in Section 2.3.2, voltage magnitudes are modeled using parameter $|\tilde{V}_i|$ for each bus i. The value of this parameter is obtained from a previous operating state (i.e., pre-contingency operations). Extending this convention, the post contingency voltage phase angle at bus i can be similarly characterized:

$$\theta_i = \tilde{\theta}_i + \theta_i^{\Delta}, \tag{2.9}$$

where $\tilde{\theta}_i$ and θ_i^{Δ} denote the pre-contingency voltage phase angle and the difference between the previous and current voltage phase angle, respectively. Differences in adjacent voltage phase angles $\theta_{ij} \equiv \theta_i - \theta_j$ can be similarly characterized:

$$\theta_{ij} = \tilde{\theta}_i + \theta_i^{\Delta} - (\tilde{\theta}_j + \theta_j^{\Delta}) = \tilde{\theta}_{ij} + \theta_{ij}^{\Delta} \tag{2.10}$$

We can then use the simple identity $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$ to obtain the following.

$$\cos\theta_{ij} = \cos\tilde{\theta}_{ij}\cos\theta_{ij}^{\Delta} - \sin\tilde{\theta}_{ij}\sin\theta_{ij}^{\Delta} \tag{2.11}$$

Similar to the DC approximations, assuming that θ_{ij}^{Δ} is sufficiently small, we can use the following approximations, $\cos\theta_{ij}^{\Delta} = 1$ and $\sin\theta_{ij}^{\Delta} = \theta_{ij}^{\Delta}$, which result in the following linear approximation for cosine:

$$\cos\theta_{ij} = \cos\tilde{\theta}_{ij} - \sin\tilde{\theta}_{ij}\theta_{ij}^{\Delta} \tag{2.12}$$

As in [31], we model our cosine approximation as an inequality rather than an equality. Therefore, the constraint included in the proposed model is

$$\hat{c}_{k[i,j]}^{\ell} \leq \cos\tilde{\theta}_{ij} - \sin\tilde{\theta}_{ij}\theta_{ij}^{\Delta} \tag{2.13}$$

There is an important additional characteristic of our cosine approximation defined in Equation (2.13). Specifically, note that the variable for the cosine approximation defined in Equation (2.13) is parametrized using the line index k. Conversely, the cosine approximation variable used in LPAC-s is parametrized by node pair $\langle i, j\rangle$. Our approximation therefore reduces the number of these variables by half. While this may appear a small change, Section 2.5 demonstrates that it substantially reduces solution time.

In addition to the above characteristic, we note that our proposed cosine approximation removes several piecewise-linear formulated constraints from the model. This is an important note because of two reasons. First, as previously discussed, such formulations may cause issues related to degeneracy and numerical instability. Because the model proposed in Section 2.3.4 contains no such constraints, these issues are less likely to be encountered. Second, as will be demonstrated in Section 2.5, our constraint approximations realize a substantial reduction in runtime.

2.3.3.3 Power Factor

Finally, we modify a key constraint implicit in LPAC-s: that the fraction of satisfied active power demand must be equal to the fraction of satisfied reactive power demand; in [31], this implicit constraint is referred to as an equal power factor. In contrast, our model uses two distinct variables so that the respective fractions of unmet demand for active and reactive power may differ. Section 2.5 demonstrates that such a modification results in a significant reduction in solution time while maintaining the ability to find optimal or near-optimal TS solutions.

Notably, all the constraint modifications described above are applicable to the original LPAC model and all of its variants. This is an important distinction because, as described in [31], the LPAC model variants have several different potential applications. Section 2.5 explores the effects the constraint modifications discussed above on both solution quality and solution time.

2.3.4 A Mixed-Integer Programming Model for AC Power Flows (MIPAC)

Because the ACOTS-LS is nonlinear and highly non-convex, attempts to address ACOTS-LS exist only sparingly in the literature. For this reason, this section proposes MIPAC, which finds TS solutions to the ACOTS-LS problem significantly faster than LPAC-s. MIPAC is an accelerated version of LPAC-s described in Section 2.3.2 which is modified using the constraint modifications described in Section 2.3.3. These modifications allow for a robust, but substantially accelerated, linearization of the ACOTS-LS model, MIPAC:

$$\max \sum_{n \in N} l_n^p p_n^d \tag{2.14a}$$

subject to

$$p_i^g - l_i^p p_i^d = \sum_{\langle i,j \rangle \in L} p_{ij} \qquad i \in N \tag{2.14b}$$

$$q_i^g - l_i^q q_i^d = \sum_{\langle i,j \rangle \in L} (q_{ij} + q_{ij}^{\Delta}) \qquad i \in N \tag{2.14c}$$

$$0 \leq l_i^p \leq 1 \qquad i \in N \tag{2.14d}$$

$$0 \leq l_i^q \leq 1 \qquad i \in N \tag{2.14e}$$

(2.2d)–(2.1j), $(1\ell), (2.7), (2.9), (2.13)$

Similar to LPAC-s, objective (2.14a) maximizes the sum of satisfied active power demand. Constraints (2.14b)-(2.14c) ensure node balance at bus i for active and reactive power, respectively, while allowing for an unequal power factor as Section 2.3.3 discusses. Constraints (2.14d)-(2.14e) enforce limits on the fraction of satisfied active and reactive power demand, respectively. Section 2.5 explores the performance of this model.

2.4 Experimental Setup

2.4.1 Computational Environment

All experiments were performed utilizing computing nodes which have 22 GB of RAM memory shared by two 2.8 GHz quad core Intel Xeon 5560 processors. The operating system was CentOS Linux version 7.6. All code was written in C++ using Concert Technology for CPLEX (version 12.4). Concert and CPLEX are registered trademarks of IBM, Inc. All ACOPF problems were solved in MATLAB using MATPOWER [69].

2.4.2 Test Case Specification

All forthcoming analyses utilized the IEEE 118-bus case [70]. As in [2], we modified the test case to meet AC feasibility requirements: we utilize the generator data from [71] augmented with a subset of generators from the 54-generator test case. This resulted in a total real power generation capacity of 6806.2 MW and reactive power generation capacity of 7029 MVAR. The total real power demand was 4519 MW and the reactive power demand totaled 1438 MVAR.

In [31], the authors tested the LPAC-R-V model using randomly-selected contingency sets from N-3 to N-20. While such experiments are useful in establishing the ability of the model to solve a wide range of scenarios, it does not establish performance in the cases that are likely to occur in practice. Therefore, we explore performance using all *non-trivial* N-1 and N-2 contingencies. Non-trivial contingencies refer to those that will result in nonzero load shed following a generation re-dispatch without implementation of TS actions.

2.5 Results

We begin our analysis with a comparison of the results produced by the DCOPF in comparison with the LPAC-R-V model. It has been established that DC-based models are crude approximations of the true behavior of AC power systems [2]; however, DC models are still used in practice. It is therefore valid to demonstrate that DC-based models such as the DCOPF and its extension the DCOTS are inappropriate for the herein studied application. The authors in [31] investigate the relative performance of the DCOPF and LPAC-R-V. However, they do not investigate the degree to which the DC model fails to accurately estimate loadshed. Moreover, their investigation only includes contingencies which are unlikely to occur: from N-3 to N-20. Therefore, a more appropriate investigation into the ability of these models to estimate load shed has merit.

In this analysis, we seek to assess the ability of the DCOPF and LPAC-R-V to respectively estimate the loadshed associated with a contingency event. Figure 2.1 shows the results of the solution, in terms of the objective value, of the DCOPF and LPAC-R-V for all non-trivial N-1 and N-2 contingencies.

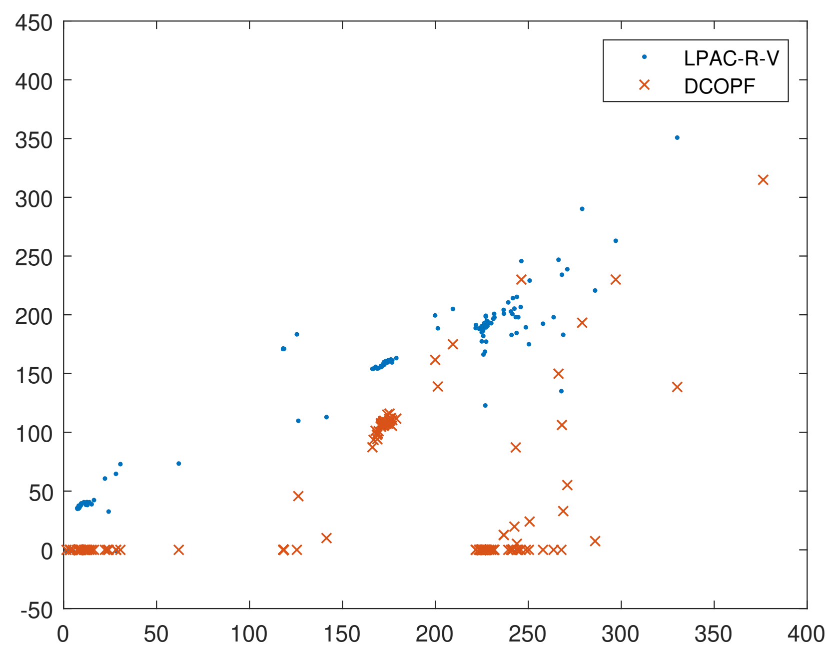

Figure 2.1: The DCOPF model dramatically underestimates the post-contingency loadshed while LPAC-R-V has a strong correlation with the ACOPF

Figure 2.1 shows that the DCOPF estimates zero loadshed in over 73% of contingency scenarios. This results in a scenario which appears trivial, i.e., there is no initial loadshed after a simple generation redispatch without transmission switching. In these scenarios, the results of teh DCOPF cannot be utilized to find any potential measures to prevent loadshed. Moreover, we note that the inability of the DCOPF to estimate loadshed is not dependent on the true loadshed according to the ACOPF. This suggests that the DC model cannot simply be prescribed based on contingency scenarios where loadshed happens to be large.

In addition, these results extend to the DCOTS problem. A contingency scenario where the DCOPF estimates zero loadshed completely destroys the ability of the DCOTS to find any valuable transmission switching actions, as there is no loadshed to be prevented. In the test case explored here, this means that in approximately 73% of scenarios we would find no valuable switching candidates. This renders the model completely useless in these cases, and motivates the development of more accurate models, such as our model, MIPAC.

There exists a strong correlation between the results of LPAC-R-V and the true value from the ACOPF. In fact, the data result in a correlation coefficient of over 0.91. The LPAC-R-V model dramatically outperforms the DCOPF in the ability to approximate the ACOPF to estimate loadshed.

We continue to motivate MIPAC by noting that, while the authors in [56] demonstrate the ability of LPAC-s to find quality solutions for the transmission restoration problem, they do not report the computational effort required for LPAC-s to obtain such solutions. We therefore further motivate the development of MIPAC with a brief analysis of the runtime of LPAC-s. While LPAC-s is certainly an accurate model (we refer the reader to [31, 56] for thorough analyses), the following analysis shows that it requires substantial acceleration in order to achieve feasible implementation to identify TS solutions and other similar applications within reasonable time limits.

Excluding the final analysis, each of the analyses herein described focus on identifying the best switching action for the given contingency. As such, for each of these experiments, we append the constraint

$$\sum_{k \in K} (1 - z_k) = 1 \tag{2.15}$$

to both MIPAC and LPAC-s. This constraint forces the model to identify the best single switching action. The authors in [40] put forth two important practical explanations for why only a single switching action is typically advisable. The first explanation is in regard to transient instability. It is known that transmission line switching creates a substantial disturbance in the grid. In particular, transient oscillations in synchronous generators will follow a TS action on a loaded transmission line. In the case that rotor angle oscillations are not well-damped, transient stability will be potentially difficult to maintain [72]. Indeed, transient stability must be checked after a single line switch and, should multiple switches be implemented, following every switch in the sequence. The second explanation put forth in [40] is that physical implementation of topology control is non-trivial. For example, current practice dictates that for transmission voltages greater than 138kV, circuit breakers are opened at one end of a line, one at a time. Following a brief waiting period, the same procedure is followed at the other end of the transmission line [57]. As such, it places substantial burden on the system operator on how to enact multiple switches. Moreover, if multiple switching actions are deemed advisable, a sequence of switching actions must be obtained. This is primarily because the resulting topology from each TS action must be evaluated to ensure no operational constraints are violated [73]. We discuss the implications of multiple switching actions in Section 2.5.4.

Table 2.1 shows the average and standard deviation of solution time in seconds for LPAC-s compared to an ACOPF-based exhaustive search of TS actions (henceforth referred to as the *exhaustive search*). Specifically, the exhaustive search consists of a sequential solving of ACOPF problems formulated with one given transmission line open. The exhaustive search concludes once all transmission lines have been searched, and is therefore guaranteed to find the optimal solution(s) for the ACOTS-LS problem. Table 2.1 shows the results of this search broken out by contingency type and in aggregate. Single and double generator contingencies are denoted G_1 and G_2. Similarly, transmission line contingencies are denoted L_1 and L_2. Mixed generator and transmission line contingencies are denoted G_1L_1. There were a total of 27 generator contingencies (G_1 and G_2), 519 transmission line contingencies (L_1 and L_2), and 248 mixed contingencies (G_1L_1).

Table 2.1: LPAC-s' solution time is larger than that of an ACOPF-based exhaustive search of switching actions

Contingency Type	LPAC-s (s)		Exhaustive Search (s)	
	Avg	SD	Avg	SD
G_1 & G_2	321.90	78.83	65.29	7.72
L_1 & L_2	234.74	186.55	59.19	3.29
G_1L_1	289.57	111.03	65.75	4.36
All	254.83	166.05	61.44	4.95

Table 2.1 shows that LPAC-s is significantly outperformed by the exhaustive search. In the worst case, for generator-only contingencies, the exhaustive search is almost five times faster than solving LPAC-s. In the best case, for transmission line contingencies, the exhaustive search is approximately four times faster than LPAC-s. These results show that LPAC-s as it stands is inappropriate for CTS and similar applications, and motivates MIPAC. The remainder of this section is organized as follows. Section 2.5.1 analyzes the accuracy of the MIPAC model in terms of the heuristic error. Section 2.5.2 investigates the acceleration of the MIPAC model due to the constraint modifications discussed in Section 2.3.3. Section 2.5.3 analyzes the performance of MIPAC in scenarios where load is decreased. Finally, Section 2.5.4 studies the performance of MIPAC in the event that multiple switching actions are implemented.

2.5.1 Most of the Solutions Found by MIPAC are Optimal Switching Actions for the ACOTS-LS

One of the most important evaluation metrics of a heuristic is the relative difference between the heuristic solution and the optimal solution. We henceforth refer to this value as "heuristic error" and denote it H_E. We calculate the heuristic error as follows

$$H_E\% = \frac{LSP^{AC} - LSP^{H}}{LSP^{AC}} \times 100, \tag{2.16}$$

where LSP^{AC} denotes the AC optimal amount of load shed prevented (obtained from the exhaustive search) and LSP^{H} denotes the amount of load shed prevented as calculated by the ACOPF with the switching action identified by the heuristic method. The amount of load shed prevented is obtained by calculating the amount of unmet demand after a generation re-dispatch without switching and subtracting the amount of unmet demand with switching. We note to the reader that a heuristic error of zero (i.e., $LSP^{H} = LSP^{AC}$) is equivalent to the heuristic method identifying an AC optimal switch. We also implement a tolerance of 10^{-5}.

Table 2.2: MIPAC attains a lower or similar heuristic error compared to LPAC-s

Contingency Type	MIPAC		LPAC-s	
	Avg	SD	Avg	SD
G_1 & G_2	33.34%	133.62%	45.96%	22.33%
L_1 & L_2	1.91%	20.21%	0.98%	13.69%
G_1L_1	25.24%	104.87%	49.03%	146.20%
All	10.27%	66.46%	17.54%	85.72%

Table 2.2 shows summary statistics for the heuristic error attained by both MIPAC and LPAC-s across all scenarios. These statistics are broken out by contingency type and

in aggregate. Table 2.2 shows that, excluding transmission line contingencies, MIPAC actually attains a smaller average heuristic error. We explore these results in greater detail in the following paragraphs.

Figure 2.2 shows a histogram of the heuristic error values of the MIPAC solutions across all instances described in Section 2.4. Note that MIPAC produced an AC infeasible switch a single time. This instance appears in Figure 2.2 in the far right-hand column. LPAC-s produces infeasible switches in three instances and a switch resulting in an undefined heuristic error in a single instance. If we exclude the single case of infeasibility, the overall average heuristic error produced by MIPAC is approximately 10.27%. The corresponding value, with the above four instances excluded, for LPAC-s is 17.54%. More importantly, the main takeaway from Figure 2.2 is that MIPAC identifies the AC optimal switching action or a switching action within a very small tolerance ($< 10^{-5}$) in the vast majority (over 94%) of studied instances. For LPAC-s, this value is only 74%. These results show that the MIPAC model almost always identifies optimal or near-optimal AC switching actions and performs equal or better than LPAC-s.

We also note that there are several instances that fall in the far right-hand column for both MIPAC and LPAC-s. Specifically, there are 34 such instances for MIPAC and 15 such instances for LPAC-s. For MIPAC, these instances occurred exclusively in one of two cases. Either when TS actions could only prevent a very small amount of load shed (less than 0.1% of total load) or when the initial load shed (i.e., after a generation redispatch without switching) was very small (less than 0.2% of total load). LPAC-s exhibited similar behavior in 14 out of 15 instances.

Several other conclusions can be garnered from the data in Figure 2.2. The first conclusion from is that, although we have modified three computationally costly constraints, for a majority of cases, the two models produce identical or near-identical results. In fact, in over 73% of studied instances, the two models produced identical results. Moreover,

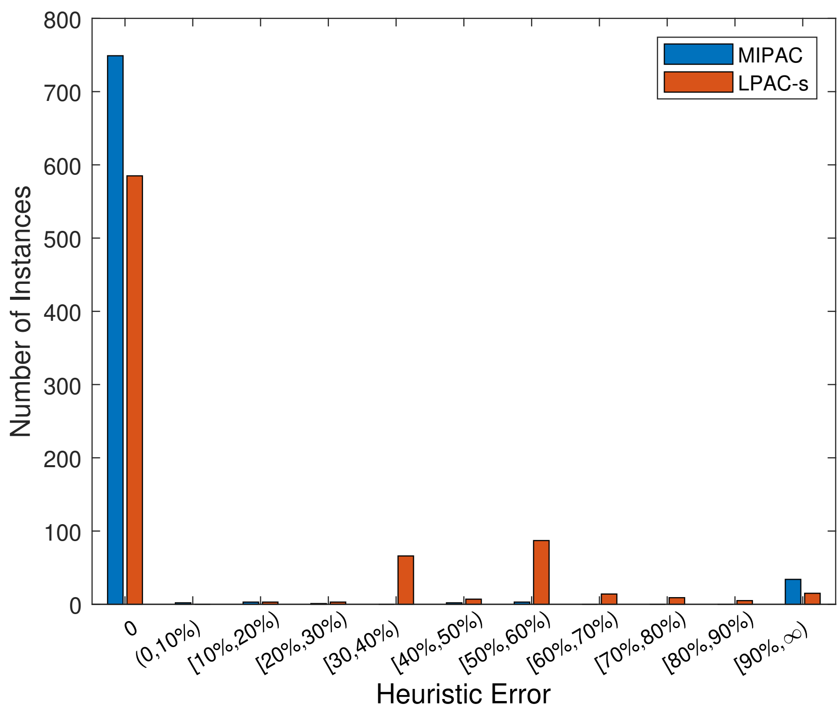

Figure 2.2: MIPAC almost always identifies an AC optimal switching action. Reprinted, with permission, from [75] ©2020 IEEE

for those instances where the heuristic errors for the solutions given by the two models are non-equal, MIPAC is superior for the majority of such instances. Specifically, in approximately 22% of all instances, MIPAC found the better switching action. In contrast, in just over four percent of instances, LPAC-s identified the superior switching action. To further investigate these results, we conducted a paired-sample t-test. Using the Figure 2.2 data, the t-test fails to reject the null hypothesis that the mean difference is zero ($p = 0.3176$). Based on these two analyses, we therefore conclude that MIPAC is, for the test case studied here, at least as accurate as LPAC-s.

2.5.2 MIPAC can be Solved Significantly Faster than LPAC-s

As shown in Section 2.5.1, the accuracy attained MIPAC is quite reasonable for practical implications. This section analyzes the improvement in solution time garnered from the constraint modifications discussed in Section 2.3.3. For this purpose, Table 2.3 shows the solution time broken out by contingency type and in aggregate.

Table 2.3: The Constraint Modifications Substantially Reduce Solution Time

Contingency Type	LPAC-s (s)		MIPAC (s)	
	Avg	SD	Avg	SD
G_1 & G_2	321.90	78.83	66.29	21.40
L_1 & L_2	234.74	186.55	11.64	14.92
G_1L_1	289.57	111.03	55.66	34.71
All	254.83	166.05	27.25	31.62

Table 2.3 shows that the solution time for MIPAC is substantially smaller than that of LPAC-s. Specifically, for generator and mixed contingencies, MIPAC can be solved around five times faster than LPAC-s. For branch contingencies this speedup rises to approximately twenty times. In aggregate, MIPAC can be solved more than nine times faster

than LPAC-s.

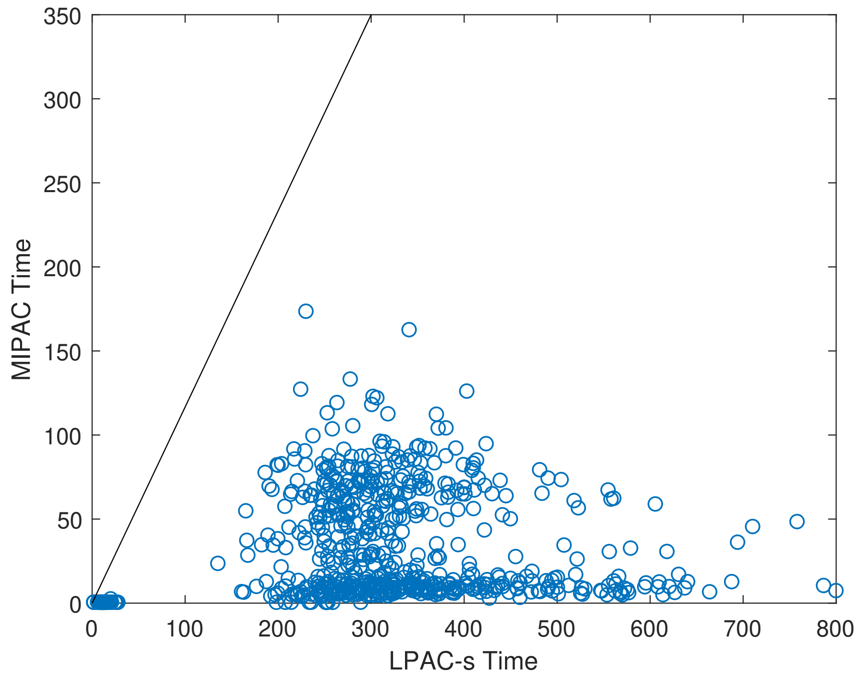

Figure 2.3: The MIPAC model has a smaller solution time that LPAC-s in the overwhelming majority of instances. Reprinted, with permission, from [75] ©2020 IEEE

Figure 2.3 further explores the solution time comparison between MIPAC and LPAC-s. This figure shows a scatter plot of the solution time for each of the two models. Every point below (above) the slanted line in Figure 2.3 denotes a scenario where MIPAC had the faster (slower) time. There are several takeaways from Figure 2.3. First, MIPAC was solved faster than LPAC-s in all instances. Additionally, Figure 2.3 shows that there are many instances where the solution time for LPAC-s is prohibitive. As an example, the worst-case instance for MIPAC requires 173.63 seconds to solve; whereas there are 591 instances where LPAC-s requires a solution time greater than or equal to this. Moreover, there are 334 instances where LPAC-s requires a prohibitive solution time of at least five

Table 2.4: The worst case instances for LPAC-s (left) and MIPAC (right) for solution time in seconds

Contingency	LPAC-s	MIPAC	Ratio	Contingency	LPAC-s	MIPAC	Ratio
L_2	799.79	7.47	107.07	G_1L_1	229.87	173.63	0.76
G_1L_1	786.44	10.52	74.76	G_1L_1	340.79	162.66	0.48
G_1L_1	757.95	48.52	15.62	G_1L_1	277.47	133.29	0.48
L_2	710.17	45.60	15.57	G_1L_1	224.20	127.27	0.57
G_1L_1	693.99	36.27	19.13	G_1L_1	402.79	126.13	0.31

minutes; there are no such instances for MIPAC.

The next analysis described in this subsection investigates the worst-case performance for the two switching models. Table 2.4 shows the worst-case instances for LPAC-s and MIPAC, respectively. In this table, for each of the two models, we list the contingency type, the solution time for each model in seconds, and the ratio of the worst-case time divided by the other model's required solution time. Note that, in the two Ratio columns in Table 2.4, the ratios of LPAC-s to MIPAC in the worst case are significantly larger than the corresponding ratios of MIPAC to LPAC-s. Here, MIPAC may outperform LPAC-s by over one hundred times, where LPAC-s always obtains a ratio below one. Moreover, in each of the five worst cases for MIPAC, MIPAC is actually the faster model. This finding shows that MIPAC outperforms LPAC-s even for instances where MIPAC performs poorly. These findings suggest that MIPAC can be reliably expected to perform reasonably close to LPAC-s in the worst case, whereas LPAC-s is prohibitively slow in the worst case.

2.5.3 MIPAC can be used successfully in lightly loaded instances

We next explore the performance of the MIPAC model in situations where the system has decreased load. It is known that in situations such as those lacking reactive power reserves, voltage issues may be prevalent [74]. As such, we herein explore the performance of MIPAC for nine levels of decreased load ranging from 10% to 90% of the fully-loaded scenario. For each of these nine levels, we developed load profiles by generating uniform random variates at each load bus which range from the specific load levels plus or minus ten percent. As an example, if the load level is set at 50%, each load bus is randomly varied between 40% and 60%. We note that scenarios with a load of 50% or less produced no more than four non-trivial contingencies, and were therefore excluded from the analysis. Figure 2.4 shows the result of this experiment. Specifically, Figure 2.4 shows a histogram of the heuristic error for each of the load groups (60% - 90%).

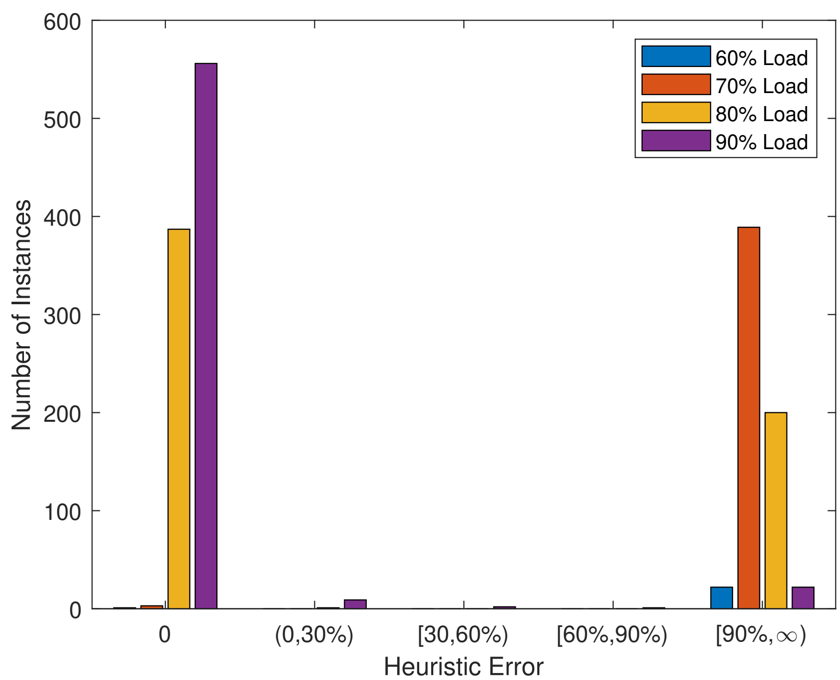

Figure 2.4: MIPAC either identifies an optimal switch or (as explained in the text) determines whether transmission switching has any practical benefit

There are several important findings from Figure 2.4. First, we note that there are two clusters which comprise the majority of instances: one at zero heuristic error and another at greater than or equal to 90% heuristic error. The first cluster, which contains switches within a tolerance of optimal, comprises over 75% of instances displayed in Figure 2.4. The second cluster, which contains instances where MIPAC performed poorly, comprises the remaining roughly 25% of instances. However, in all but 11 instances appearing in the second cluster, there was either very low initial load shed (less than 0.9% of total load) or low load shed prevented by the exhaustive search (less than 0.1% of total load). Therefore, we can conclude that (1) in instances where transmission switching is useful, MIPAC identifies strong switching actions, and (2) in instances where transmission switching is not practical, the model will inform the system operator as such.

2.5.4 MIPAC Outperforms Exhaustive Search, Particularly as the Allowed Number of Switches Increases

We conclude our analysis with a discussion comparing MIPAC to the exhaustive search described in Section 2.5. From Tables 2.1 and 2.3, we can conclude that, when seeking the best single switching action for generator contingencies, MIPAC is approximately equal to the exhaustive search. In the best case, for transmission line contingencies, MIPAC achieves a 5.1 times speedup. In aggregate, using MIPAC as a replacement for the exhaustive search results in an average speedup of approximately 2.3 times.

It is also important to note that, as the number of allowed switches increases, the MIPAC model's comparison vs an exhaustive search improves. To demonstrate this, we append MIPAC with the following constraint.

$$\sum_{k \in K} (1 - z_k) \leq 10 \tag{2.17}$$

Constraint (2.17) dictates that a set of at most ten lines are switched. Within this experiment, we set the CPLEX optimality tolerance, the point at which the relative difference between the incumbent solution and the best lower or upper bound is sufficiently close to optimal, to 0.5%. Table 2.5 summarizes the results of this experiment. We note that in only nine instances (1.1% of all instances), MIPAC could not be solved in less than ten hours of runtime. These instances are excluded from Table 2.5.

Table 2.5: MIPAC scales well for up to 10 switching actions

Contingency Type	MIPAC (min)	
	Avg	SD
G_1 & G_2	28.75	23.35
L_1 & L_2	0.20	0.27
G_1L_1	25.68	49.12
All	9.01	30.32

As the reader can see, across all groups, the solution time has increased substantially. However, these results compare quite favorably to the exhaustive search. From Table 2.1, we can conclude that an exhaustive search with ten switches will require on the order of 10^{11} hours to complete. As such, for a relatively small number of potential switching actions, MIPAC clearly outperforms the exhaustive search. We may therefore conclude that, particularly as the number of allowable switches increases, MIPAC shows strong potential to identify post-contingency switching actions for load shed prevention.

2.6 Discussion and Conclusion

This work accelerated an existing mixed-integer linear programming heuristic for AC optimal transmission switching (TS) for load shed prevention. Specifically, we modified three computationally-costly constraints in the LPAC model for load shed prevention [31, 56]. The resulting mixed-integer linear model, MIPAC, showed substantial promise as a heuristic to identify valuable post-contingency AC TS actions.

In terms of accuracy, MIPAC achieved an heuristic error of approximately 10% across all test instances. More importantly, MIPAC identified an AC optimal or near-optimal switch in over 94% of instances and identified a switching action as good or better than the base model in over 95% of the instances. These results demonstrate that MIPAC is a highly accurate tool which almost always identifies AC optimal TS actions. In terms of solution time, MIPAC achieved a substantial speedup compared both to the base model from [56] and an ACOPF-based exhaustive search of switching actions. Specifically, for a single switching action, the average solution time for MIPAC is approximately nine times faster than the base model; compared to the exhaustive search, MIPAC is approximately 2.3 times faster. MIPAC also proved useful when under decreased load and when multiple switches are allowed. Specifically, over 75% of switching actions identified by MIPAC were optimal, and the bulk of instances where such actions were suboptimal were cases where transmission switching would be impractical. When ten switching actions were allowed, MIPAC significantly outperformed the exhaustive search by several orders of magnitude.

It is important to note that the constraint modifications developed herein are applicable to any variant of the original LPAC model proposed in [31]. These models have already bee applied to several application areas, including power system restoration and capacitor replacement. Moreover, there are several additional application areas where such models have potential. In particular, these include power systems optimization models with discrete variables, such as scheduling for resilience to contingencies, use of FACTS devices, and islanding. As the constraint modifications discussed here have had no adverse affect on accuracy and result in substantial acceleration, these modifications should be considered in future efforts for computational considerations in LPAC-based and other related models.

Despite MIPAC's performance, further improvements are necessary in order to achieve feasibility for large-scale implementation. One particular avenue of research is the use of heuristics such as the line profit heuristic [37], which are applicable to MIPAC and can be used to further accelerate the model at the risk of sacrificing accuracy. In addition, we also plan to pursue two research avenues which are related to this work. We intend to characterize the optimality gap of the MIPAC model with multiple switches. We also intend to seek out ways to characterize the amount of load shed that may be prevented a priori, potentially alleviating many of the inaccurate instances described in Figure 2.2. Regardless, MIPAC shows promise because of its ability to find strong TS solutions.

3. A DATA MINING TRANSMISSION SWITCHING HEURISTIC FOR POST-CONTINGENCY ACPF VIOLATION REDUCTION IN REAL-WORLD, LARGE-SCALE SYSTEMS

3.1 Introduction

The robustness of the electrical power grid is one of the most vital features of our critical infrastructure. Therefore, research efforts which accurately model operation of the grid and validate its robustness are of great importance going forward. In particular, methods concerning post-contingency operations are noteworthy because they mitigate the harm which the grid may undergo following component failure. One notable analytical technique is contingency analysis, which allows operators to study the impacts of various contingencies and develop corrective measures which may be applied should such a failure occur. One example of a post-contingency corrective measure is transmission switching, also known as topology control, which we herein study.

In the past, the power grid has been modeled using a fixed configuration [57]. Using such a modeling paradigm, control on the grid is exerted only by making dispatch decisions. However, transmission switching allows system operators an additional method of control by physically switching transmission lines in and out of the grid. Previous research has demonstrated that transmission switching has a myriad of potential benefits. In one of the initial papers, the authors in [34] showed that transmission switching can produce significant reductions in generation fuel costs. Other works have echoed this conclusion with focuses on sensitivity analysis ([35, 37]) and contingency analysis [36]. In addition to cost, transmission switching has demonstrated usefulness in preventing loadshed ([39, 57, 75]), improving system reliability [38], and, as studied herein, reducing post-contingency violations [40].

This paper develops a data mining method which identifies transmission switching actions to reduce post-contingency voltage magnitude and branch flow violations. Specifically, we apply a guided undersampling method proposed by [76] and then utilize logistic regression to identify post-contingency transmission switching candidates to reduce AC power flow (ACPF) violations. Notably, these data mining methods are computationally inexpensive and can be quickly executed on real-world AC power system data, providing greater certainty regarding *both* AC system performance and large-scale implementation. This is notable because it addresses three of the four explanations mentioned in [40] as to why transmission switching is currently being used only in limited capacity: computational complexity, uncertainty of impact on real-world large-scale power systems, uncertainty of impact when moving from DC to AC, and transient stability. These four items are substantiated in detail in the following paragraphs.

1 – Computational complexity: The first explanation for lack of widespread implementation of transmission switching is in regard to computational complexity or, equivalently, the scalability of algorithms. Mixed-integer nonlinear programs (MINLPs), such as the AC optimal transmission switching (ACOTS) problem, are notoriously difficult to solve. Because of this, researchers have primarily tested algorithms on small-scale networks and/or used linearizations to reduce the difficulty of the problem, which cast uncertainty over scalability and solution accuracy, respectively. These two approaches are discussed in greater detail in the following two paragraphs.

2 – Impact on large-scale systems: The second item is that the overwhelming majority of research on transmission switching studies small-scale systems. To substantiate this fact, we note that, excluding ([40, 77, 78]), all papers herein cited study small-scale systems such as the IEEE 118-bus test case which appear in test case archives such as [70]. In contrast, real-world systems have up to tens of thousands of buses. Given that most algorithms are tested on small systems, it remains unclear if these algorithms can be applied in the real world.

3 – Impact on AC systems: The third item addresses the concern that benefits derived from DC-based models may not translate to AC models. Because of the issues associated with solving problems such as the ACOTS, the majority of transmission switching research focuses on DC-based models. Indeed, excluding [53, 55, 40, 77, 48, 79], all papers herein cited use linearizations to address the various transmission switching problem variants. This is worrisome because, as demonstrated in [2], transmission switching actions identified by DC models may be infeasible or result in negative outcomes when implemented in AC systems. Moreover, as aforesaid, most if not all of these models are only tested on unrealistically small networks, further clouding the issue of whether conclusions derived from DC models can be applied to AC systems.

4 – Transient stability Finally, we note that relatively recent research has shown that transmission line switching may introduce disturbances into large power systems that may cause stability problems [80]. Despite this fact, the overwhelming majority of research in this area fails to account for transient stability. While some works tested for transient stability after the fact, to the authors knowledge, the method described by the authors in [57] is the only one to systematically verify stability within the algorithm itself. Moreover, the framework proposed in [57] can be combined with virtually any approach.

This work makes three primary contributions. First, to the authors' knowledge, our proposed heuristic is the first true data mining technique to classify strong transmission-switching actions using power flow information rather than a simple history of useful switching actions. Second, our method is unique in the literature in that it directly addresses three of the four issues regarding transmission switching implementation. Moreover, it can be directly combined with with the framework in [57] to address the fourth issue. Finally, we make a much stronger empirical analysis than the analysis described in [40], studying the impact of an exhaustive set of switching actions on real-world, large-scale power system data, rather than a small subset of the actions or actions across a small system. A secondary contribution of this work is that we formalize the problem characterized in [40].

We note that, in recent years, several papers claim to use “previous knowledge” in the context of a data mining framework, but fail to do so in the way data mining is understood. The authors in [77] propose a heuristic where a lookup table containing contingency-specific switching solutions is used to select potential switching candidates. The authors in [36] propose a similar heuristic where previously-successful switches are considered as potential switching actions. Finally, the authors in [40] extend the approach in [36] by splitting previously-successful switches into training and testing sets; these switches are subsequently validated accordingly. However, this method does not incorporate any additional information other than the switching action itself. In contrast with the above methods, the herein proposed method incorporates information about the current status of the grid into a sophisticated data mining method to predict the *impact* of a given candidate switching action. Succinctly, the resulting method is a true data mining approach which exploits the set of available information to identify candidate switching actions.

The remainder of this paper is organized as follows. Section 3.2 presents and describes two optimization models for post-contingency ACPF violation reduction. Section 3.3 proposes our data mining heuristic using guided undersampling and logistic regression. Section 3.4 discusses the experimental setup and several implementation issues which are inherent to the proposed methodology and the herein studied problem. Section 3.5 presents our analysis and demonstrates that our proposed method identifies optimal and near-optimal switching solutions and is superior to existing heuristics based on distance to violation elements. Section 3.6 concludes the work.

3.2 ACPF Violation Reduction Optimization Model with Transmission Switching

During post-contingency operations, operators must be provided with corrective actions which are both quick to identify and can be implemented with confidence. One manner to derive actions is through the solution of optimization models, such as those described in this section. We herein describe the mathematical formulation of the ACPF optimization model for violation reduction with transmission switching (ACPF-VR-TS). Note that these problems were described qualitatively by the authors in [40]. However, this work is the first to formally characterize these problems as optimization models.

3.2.1 ACPF-VR-TS for Voltage Magnitude Violations

The ACPF-VR-TS model for voltage magnitude violation reduction is characterized as follows.

$$\min \sum_{n \in N} V_n^+ + V_n^- \tag{3.1a}$$

subject to

$$P_i^g - P_i^d = \sum_{\langle i,j \rangle \in K} p_{ij} \qquad i \in N \tag{3.1b}$$

$$Q_i^g - Q_i^d = \sum_{\langle i,j \rangle \in K} q_{ij} \qquad i \in N \tag{3.1c}$$

$$p_{ij} = Z_{ij}(g_{ij}V_i^2 - V_iV_j(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij})) \qquad \langle i,j \rangle \in K \tag{3.1d}$$

$$q_{ij} = Z_{ij}(b_{ij}V_i^2 - V_iV_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij})) \qquad \langle i,j \rangle \in K \tag{3.1e}$$

$$V_n^- \geq V_n^{\min} - V_n \qquad n \in N \tag{3.1f}$$

$$V_n^- \geq 0 \qquad n \in N \tag{3.1g}$$

$$V_n^+ \geq V_n - V_n^{\max} \qquad n \in N \tag{3.1h}$$

$$V_n^+ \geq 0 \qquad n \in N \tag{3.1i}$$

$$\sum_{\langle i,j \rangle \in K} 1 - Z_{ij} = 1 \tag{3.1j}$$

Objective (3.1a) minimizes the sum of voltage magnitude violations. Constraints (3.1b) and (3.1c) model active and reactive power injections at bus i, respectively. Constraints (3.1d) and (3.1e) model the flow of active and reactive power from bus i to bus j, respectively, while accounting for transmission switching. Note that fixing all binary variables Z_{ij} to 1 reduces constraints (3.1b)-(3.1e) to the constraints which characterize the ACPF. Constraints (3.1f) and (3.1g) model the violation of the voltage magnitude lower bounds. Constraints (3.1h) and (3.1i) model the violation of the voltage magnitude upper bounds.

Finally, Constraint (3.1j) dictates that the number of allowable transmission line switches is equal to one. There are two primary practical explanations for this constraint. First, significant concerns exist regarding transient stability with multiple switching actions. Second, physical implementation of transmission switching requires substantial effort. This effort magnifies as the number of switches increases. We refer the reader to [75] for a detailed discussion on these two topics.

3.2.2 ACPF-VR-TS for Branch flow Violations

The ACPF-VR-TS model for branch flow violation reduction is characterized as follows.

$$\min \sum_{\langle i,j\rangle \in K} f_{ij} \tag{3.2a}$$

subject to

(3.1b)–(3.1e), (3.1j) (2b)–(2e),(2j)

$$f^i_{ij} \geq p^2_{ij} + q^2_{ij} - S^2_k \qquad \langle i,j\rangle \in K \tag{3.2f}$$

$$f^j_{ij} \geq p^2_{ji} + q^2_{ji} - S^2_k \qquad \langle i,j\rangle \in K \tag{3.2g}$$

$$f_{ij} \geq f^i_{ij} \qquad \langle i,j\rangle \in K \tag{3.2h}$$

$$f_{ij} \geq 0 \qquad \langle i,j\rangle \in K \tag{3.2i}$$

Objective (3.2a) minimizes the sum of branch flow violations. Constraints (3.2f) and (3.2g) model violation of thermal limits as a function of apparent power. Constraint (3.2h)-(3.2i) dictates that the flow violation for transmission line k is equal to either zero or the larger of the two violations at each end of the line.

3.3 Methodology

It is important to note that, as in [40], we do not solve the models outlined in Section 3.2 via the use of an optimization solver. The primary reason for this is, as discussed in Section 3.1, MINLPs are prohibitively difficult to solve. As such, most works addressing problems similar to those in Section 3.2 either use linearizations or solve across small systems. In contrast, this work seeks solutions to these models via a data mining-based heuristic in a manner tractable at a large scale without linearization. The following two paragraphs discuss an important feature of transmission switching which contextualizes the problem within a data mining framework – imbalanced data. Following this, we present our dual-component guided undersampling method, introduce the classification methodology, and formalize the proposed method.

It is well known that, regardless of objective (e.g., violation reduction or cost savings), there are typically only a small number of transmission switching actions which result in a substantial objective improvement. Thus, in a data mining context, identifying strong transmission switching actions should be viewed as an imbalanced-data classification problem. In a two-class setting, imbalanced data refers to a dataset where one class (the majority class) has cardinality substantially larger than that of the opposing (minority) class. In such settings, classification algorithms can suffer poor performance because they can simply classify all instances as majority and still achieve a high classification rate. Therefore, when developing a data mining approach for transmission switching, one must take care to adequately handle the problems inherent to imbalanced data.

The proposed methodology addresses the imbalanced data problem as follows. First, we utilize a guided undersampling procedure [76] which undersamples the majority class using two instance-selecting techniques. The first technique, ensemble outlier-filtering, utilizes a unique ensemble classifier to remove both majority and minority outliers from the training data. The second technique, normalized-cut sampling, undersamples the majority class such that the density distribution of the majority class is preserved. After our guided undersampling method, we apply logistic regression to predict which transmission lines are strong candidates. The following subsections detail each of the components of the proposed methodology.

3.3.1 Ensemble Outlier-Filtering

The first component of the guided undersampling method used herein is ensemble outlier-filtering. For the purposes of this work, we consider an outlier as a data instance whose class (i.e., whether or not it is a strong transmission switching action) differs from the majority of data instances with similar features. From a data-analytic perspective, outlier removal is a critical component of building an effective classifier. However, when the training data is imbalanced, many outlier filtering techniques demonstrate poor performance. One potential explanation for such performance is that, as many conventional outlier removal approaches seek to minimize the number of misclassified instances. this goal may prove faulty when the ratio of majority to minority instances is very high because many minority instances, even those strongly indicative of the minority region, may be identified as outliers. To address this, the authors in [76] proposed an ensemble outlier-filtering technique which utilizes the power of an ensemble classifier in which each training set is balanced. The ensemble outlier-filtering technique is described as follows. Note that, throughout the remainder of this section, an important value is the ratio of majority instances to minority instances, known as the imbalance ratio.

Given a set of data $X = X_{\text{maj}} \cup X_{\text{min}}$, with imbalance ratio $r = \frac{|X_{\text{maj}}|}{|X_{\text{min}}|}$, ensemble outlier-filtering proceeds as follows.

1. Partition the majority class X_{maj} into r subsets of equal size, where each subset has cardinality $|X_{\text{min}}|$. Construct r distinct training subsets: X_{min} and one subset of X_{maj}.
2. Train r logistic regression classifiers, one for each subset of the training data.
3. Predict the class of every instance in X using the majority voting scheme: instance X_i is assigned to a class if it is predicted as such by at least $\frac{r}{2}$ classifiers.
4. Remove from the training data the outliers – instances whose predicted class differs from their true class.

One salient characteristic of the method described above is that it seeks to remove both minority and majority from the training data. While the removal of minority data may appear counterintuitive, as shown in [76], it is critical to strong imbalanced data classification performance. In particular, in the context of imbalanced data classification approaches, minority outliers should be viewed as high-leverage instances which may exhibit undue influence on the classifier. It is therefore critical that they be identified and removed to achieve strong classification performance.

3.3.2 Normalized-Cut Sampling

One traditional approach to imbalanced data classification is majority subsampling, in which a subset of the instances from the majority class is selected from the training data in an effort to improve or eliminate class imbalance. However, as described in [76], the user must take great care to construct their majority subsample, as they risk an inaccurate decision boundary if there are regions where the density distribution of the greater majority class does not persist. To address this, the authors in [76] proposed a new technique, normalized-cut sampling. This method utilizes the normalized-cut segmentation technique proposed in [81] to iteratively cluster the majority class such that instances within clusters are similar and instances in distinct clusters are different. The medoids of the clusters are then selected as the majority class subsample.

Given training data $X = X_{\text{maj}} \cup X_{\text{min}}$, with $k = |X_{\text{min}}|$, the normalized-cut sampling procedure is described as follows.

1. Construct graph $G_1 = (V, E)$, where V denotes all majority instances and E contains edges between all node pairs in V. Set edge weights between node pair (i, j) as

$$\mathcal{L}_{ij} = \exp(||\mathbf{X}_i - \mathbf{X}_j||^2) \tag{3.3}$$

2. Initialize an empty set C.

3. For $i = 1 \ldots k$

   (a) Utilizing the procedure from [81], bipartition G_i into clusters C_i^1 and C_i^2. Add these two clusters to C.

   (b) Construct G_{i+1} using the instances from the cluster in C with maximum cardinality ($C_{\max}$).

   (c) Update $C = C \setminus C_{\max}$.

4. Form the majority subsample with the medoids of the clusters in C.

Note that in step (3), the iterative procedure is conducted k times, where k is equal to the cardinality of the minority class. That is, the above procedure will create k clusters, and the medoids of the clusters will form the new majority subsample. Moreover, the procedure will indeed preserve the density distribution of the original majority class because in regions with higher density, normalized-cut sampling will create a higher number of clusters and in regions with low density, fewer clusters will be created.

3.3.3 Logistic Regression

Logistic regression is a classification model which estimates the probability that a given categorical dependent variable belongs to a particular category [82]. Given two categories, denoted numerically as 0 and 1 and an n-dimensional datum **x**, the logistic regression model estimates the probability that the datum belongs to category 1 as

$$P(y|\mathbf{x}) = \frac{e^{\beta_0 + \sum_{j=1}^{n} \beta_j x_j}}{1 + e^{\beta_0 + \sum_{j=1}^{n} \beta_j x_j}}, \tag{3.4}$$

where β_0 is the intercept of the regression function and β_j is the multiplicative regression coefficient associated with the j-th dimension of the data. The regression coefficients are those which maximize the log-likelihood function

$$\begin{aligned} \ln L(\beta|Y) = &\sum_{i=1}^{m} Y_i \ln P(Y_i|\mathbf{X}_i) \\ &+ \sum_{i=1}^{m} (1 - Y_i) \ln(1 - P(Y_i|\mathbf{X}_i)), \end{aligned} \tag{3.5}$$

where **X** is an $m \times n$-dimensional training data set and $Y \in \{0, 1\}^m$ encodes the category of each training datum.

Note that the output of the logistic regression model from Equation (3.4) is the probability that a given datum **x** belongs to the class of focus.

3.3.4 Proposed Methodology

The previous subsections detailed the two components of the guided undersampling method as well as a specific classification methodology. Given training data $X = X_{\text{maj}} \cup X_{\text{min}}$, and target variable Y, the proposed method is formalized as follows.

1. Apply ensemble outlier-filtering to the training data (X, Y) to obtain the clean subsample $X' = X'_{\text{maj}} \cup X'_{\text{min}}$

2. Apply normalized-cut sampling to X'_{maj} to further subset the majority class and obtain X''_{maj}, where $|X''_{\text{maj}}| = |X'_{\text{min}}|$

3. Train a logistic regression model on the clean and balanced training data $X''_{\text{maj}} \cup X'_{\text{min}}$ to derive the classification model

The logistic regression model which is trained in Step (3) can then be utilized as follows. Given a transmission line, its characteristics, the post-contingency ACPF information, and the relationship among these items, the model trained in Step (3) can be used to predict whether the transmission line is a strong switching candidate. In our case, Equation (3.4) can be interpreted as the probability that a transmission switching action results in a substantial reduction in post-contingency violations. Using this procedure for each transmission line in the system, we can rank each line according to its expected impact. We elaborate on this process in the following section.

3.4 Experimental Setup

3.4.1 Test Case

The test case utilized for analysis herein consists of emergency management system snapshots from the Pennsylvania New Jersey Maryland Interconnection (PJM). These snapshots span 167 hours, with a total of 8064 critical contingencies across the entire test case. For a detailed description on how these contingencies were identified, we refer the reader to [40]. The PJM system consists of approximately 15,500 buses, 2,800 generators, and 20,500 transmission lines. The total active power load is approximately 139 GW and the total reactive power load is approximately 22 GW.

3.4.2 Data Development

3.4.2.1 Exhaustive Search of Switching Actions

To develop the training data which drives the predictive model, we performed an exhaustive search (henceforth referred to as the "exhaustive search") of non-radial transmission switching actions across all critical contingencies from [40] as described in the following paragraph. Because it is indeed exhaustive, this search guarantees that, for each contingency, we will identify the switching action that reduces the most violations (i.e., the optimal switching action). This is a salient point which allows this work to be the first to conduct a thorough analysis using the true optimality gap (as defined in Section 3.5.1) for the herein studied heuristics and for a problem of this size.

We perform the exhaustive search for each hour and critical contingency. For contingencies which did not include generator failure, all generator outputs remained at the pre-contingency level with the exception of the slack bus. For contingencies which simulate generator failure, re-dispatch was performed using a participation factor as outlined in [83]. Finally, only a single corrective line switch is implemented. The procedure utilized to generate the data is described in the following steps. Given a power system instance with n non-radial lines, the exhaustive search proceeds as follows.

1. Run ACPF using given input data
2. Simulate contingency
3. Calculate re-dispatch (if necessary)
4. Run ACPF
5. Record system information
6. For $i = 1 \ldots n$

(a) Perform line switch

(b) Run ACPF

(c) Record remaining violation magnitude (if any)

Note that, in Steps (5) and (6c), information gathered includes both voltage magnitude and branch flow violations. As such, the exhaustive search generates data for both problems outlined in Section 3.2. The exhaustive search outlined above was run systematically to include all critical contingencies identified in [40] and all feasible line switches. Doing so yields a rich data set which can be exploited via the use of data mining techniques.

3.4.2.2 Training Data

For each of our objectives (i.e., voltage magnitude and branch flow violations, respectively), the training data was created as follows. For each hour and critical contingency, we took only the top 100 switching actions; we did so for three reasons. First, this allowed us to satisfy the memory requirements of the computational setup used to train the model. Second, we note that it has been established in the literature that one successful approach to majority subsampling is to focus on regions where the majority and minority classes overlap [84]. Finally, we note that even using the top 100 switching actions results in an imbalanced data set. Specifically, for voltage magnitude violations, approximately 2.5% of the top 100 switches result in an optimality gap of less than ten percent, the threshold for which the categories discussed in Section 3.3.3 were coded. For branch flow violations, only 5.7% switches result in such an optimality gap.

3.4.2.3 Data Features

There are several important pieces of data collected during step (5) of the procedure outlined in Section 3.4.2.1 which drive the prediction methodology. The first set of features describes the switching element. This set includes branch resistance, reactance, susceptance, thermal rating, and active and reactive power flow. The second set of features is composed of the magnitude of all violations within a set of distances (calculated using the unweighted, undirected graph distance) around the switching element. The third feature set includes characteristics of the violation elements including bus type, active and reactive power demand, node degree, resistance, reactance, MVA rating, active power flow, and reactive power flow. The fourth set comprises the undirected distance from the switching element to each violation element, the distance from each violation element to the contingency element, and the distance from the switching element to the contingency element. The final feature set contains distance data in a directed fashion, using the flow of active power to construct a directed graph. From this graph we identified distances identical to those described above.

We note that the model herein developed is sensitive to the set of features used to train it. As such, we conducted feature selection prior to training the model. We utilized the first 83 hours (approximately on half of the data set), to conduct a forward-stepwise procedure to select a set of features which minimized the optimality gap. As such, all forthcoming experiments were conducted on the remaining 84 hours of data described in Section 3.4.1.

3.4.3 Cross-Validation

Cross-validation is imperative for the development of any predictive model to properly assess the model's strength. However, it bears mentioning that cross-validation in this setting is different than in traditional contexts. In an effort to minimize bias from similar instances, each fold used during cross-validation consisted of all contingencies within a single hour out of the 84 hours used for testing as described in Section 3.4.2.3. The remaining data used to train the model consists of all contingencies within the remaining 83 hours. This resulted in an 84-fold cross-validation procedure. In addition, each contingency in the test fold was tested individually for the performance metric described in Section 3.5.1. That is, rather than a single data point for each fold, testing generates a number of data points equal to the number of contingencies within the given hour.

3.4.4 Computational Environment

All computational experiments herein described were conducted in a distributed environment on computing nodes which had 22GB of RAM shared by two 2.8 GHz quad core Intel 503 Xeon 5560 processors using Texas A&M University supercomputing resources. The operating system was CentOS Linux version 7.6.1810. All power flows were solved using the MATPOWER toolbox [69]. Ensemble outlier-filtering, normalized-cut sampling, and logistic regression models were implemented using the LIBLINEAR package [85] within MATLAB.

3.5 Results

3.5.1 Performance Metric

To measure the effectiveness of our data mining heuristic, we utilize the optimality gap. This measure describes the relative difference between the percentage of violations reduced by the optimal switching action and the corresponding value from the best action chosen by the heuristic. We utilize the percentage of violations reduced because it dampens the impact of instances with a particularly large or small initial violation magnitude. As mentioned in Section 3.4.2.1, we note that the exhaustive search guarantees the identification of the optimal solution. Therefore, the forthcoming analysis using the optimality gap provides context on how closely a heuristic performs relative to the optimal solution. The optimality gap is calculated as follows

$$Gap\% = \frac{V^* - V^H}{V^*} \times 100, \tag{3.6}$$

where V^* denotes the percentage of violation reduction stemming from the optimal switching action as identified by the exhaustive search, and V^H denotes corresponding value stemming from the switching action proposed by the heuristic. Note that we calculate this value separately for voltage magnitude violations and branch flow violations.

3.5.2 The Data Mininig Heuristic Consistently Identifies Optimal or Near-Optimal Transmission Switching Actions

Table 3.1 shows the average and standard deviation of the optimality gap obtained for the top switching action according to Equation (3.4), fixed all other lines as closed, and solved the ACPF problem associated with that topology. The results are broken out by violation type.

Table 3.1: Summary statistics for optimality gap

Violation Type	Mean	St. Dev.
Voltage Magnitude	5.51%	17.60%
Branch Flow	5.89%	93.71%

Table 3.1 shows that the data mining heuristic produces strong performance in terms of optimality gap. Specifically, for both voltage magnitude and branch flow violations, the average optimality gap of the top switch identified by the data mining heuristic is less than six percent, which is a very reasonable result in practice for real-world, large-scale systems. However, the branch flow violations produce a relatively large standard deviation. This is largely because of three instances, where the data mining heuristic produced optimality gaps substantially larger than 100%. Removing these three instances, the standard deviation drops to 17.5%, a much more reasonable value. The following paragraphs explore these results in greater detail.

Figure 3.1 shows a histogram of the optimality gaps attained by the data mining heuristic for the two types of violations. The first finding shown in Figure 3.1 is that the bulk of instances fall either at zero or less than 0.1%. This means that, for the bulk of instances herein studied, the data mining heuristic identified optimal or near-optimal switching actions.

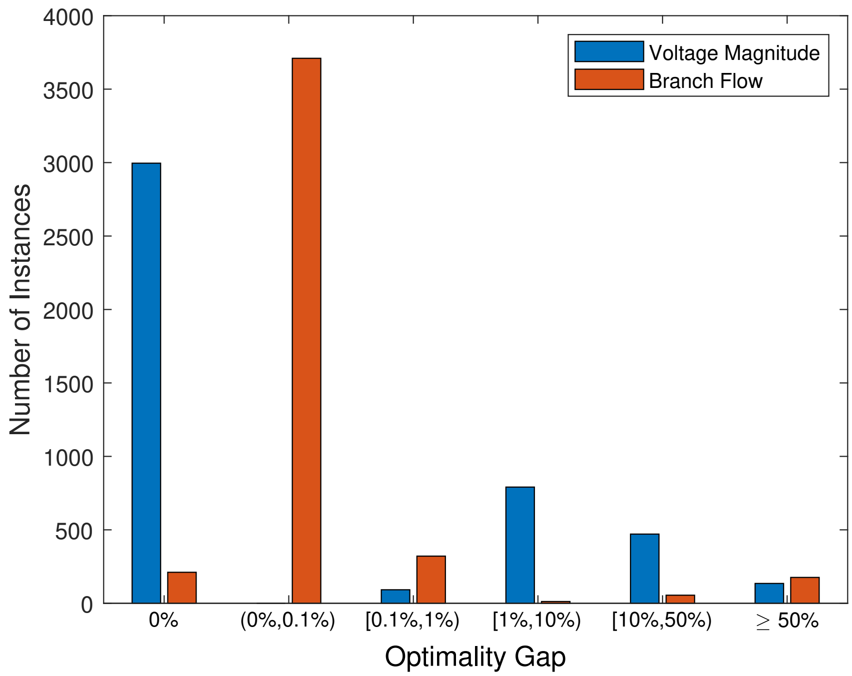

Figure 3.1: The data mining heuristic attains optimality gaps at or near zero in the overwhelming majority of instances

One additional important finding from Figure 3.1 is the number of scenarios with optimality gaps greater than 50%. Specifically there are 135 such scenarios for voltage magnitude violations and 176 such scenarios for branch flow violations. We note that these scenarios make up only approximately 1.7% and 2.2% of studied instances, respectively, further demonstrating the strength of the method.

3.5.3 The Data Mining Hueristic Attains an Optimality Gap Substantially Smaller than that of a Distance-based Heuristic

While Section 3.5.2 showed that the data mining heuristic produced strong performance in its own right, it is also important to compare against the strongest-performing existing methods. The authors in [40] developed distance-based heuristics which constitute the strongest-performing methods which are viable at a large scale for the problems outlined in Section 3.2. Specifically, the authors in [40] developed a heuristic (CBVE) which selects as candidate switching branches those closest to the violation element(s). However, because there are typically multiple violation elements and multiple branches "closest" to each violation element, there is no such thing as the single highest-ranked switch according to CBVE. Therefore, in order to form a fair basis for comparing the proposed method to CBVE, the forthcoming experiments utilize a candidate pool – a group of switching candidates from which the best switch is selected. Accordingly, for the proposed method, we re-conducted feature selection using a candidate pool of ten. In the interest of full disclosure, we note that the authors in [40] also developed a heuristic using the distance from the switch to the contingency element. However, in our experiments, this method was effectively dominated by CBVE. We therefore excluded it from our analysis.

Table 3.2 summarizes the performance of the two heuristic methods in terms of the optimality gap using a candidate pool of size ten. Specifically, Table 3.2 shows the mean and standard deviation of the optimality gap for the proposed data mining method, denoted DM-10, against that of the distance-based metric, CBVE, when using a candidate pool of size ten. These results are broken out by the type of violation.

Table 3.2 shows that the data mining heuristic outperforms the distance based heuristic across the board. In regard to voltage magnitude violations, the proposed heuristic attains an average optimality gap over eight times smaller than that of the distance-based heuris-

Table 3.2: Summary statistics for optimality gap using a candidate pool of size ten

Violation Type	Method	Mean	St. Dev.
Voltage Magnitude	DM-10	3.41%	15.91%
	CBVE	28.94%	8.91%
Branch Flow	DM-10	0.04%	0.81%
	CBVE	1.58%	10.04%

tic. Regarding branch flow violations, the average optimality gap using the data mining heuristic is almost forty times smaller than that of the distance-based heuristic; the standard deviation is 12 times smaller using the same comparison. These results show that, at a high level, the data mining heuristic has extremely strong performance against distance-based heuristics in terms of the optimality gap. The remainder of this section explores these results in greater detail.

Figure 3.2 plots the empirical cumulative distribution function (ECDF) of the optimality gaps obtained for voltage magnitude violation reduction by both the proposed data mining heuristic and the distance based heuristic. An ECDF plots the fraction of data points that are less than or equal to a certain value for all possible values of the metric of interest. We use this plot because it characterizes the fraction of instances for which each heuristic achieves a certain performance.

There are two findings from Figure 3.2. First, for the data mining heuristic, there is a vertical line at zero which reaches almost 92% of instances. This means that, in approximately 92% of studied instances, the proposed heuristic identified the optimal switch. In contrast, the distance-based heuristic had no such instances where the optimal switch was found within the candidate pool. This can be seen when the blue line diverges from the red line. These results show that, for the test case studied here, the data mining heuristic dramatically outperforms the distance-based heuristic in identifying optimal or near-optimal solutions.

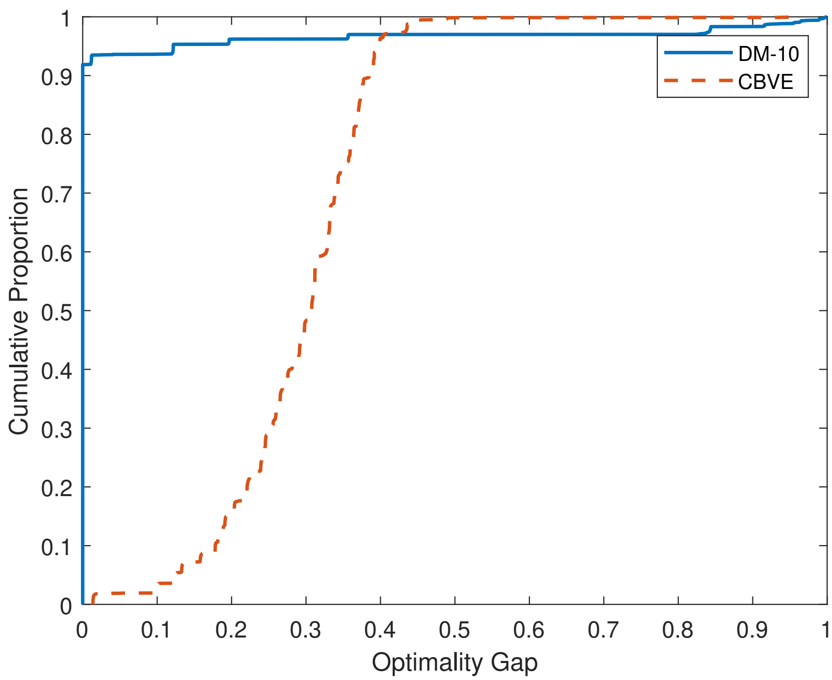

Figure 3.2: DM-10 dramatically outperforms CBVE for voltage violation reduction in terms of the optimality gap

The second finding is the relative performance between the data mining heuristic and the distance-based heuristic. Specifically, the proposed data mining heuristic attains an acceptable optimality gap of less than ten percent in over 93% of instances and and optimality gap of less than 25 percent in over 96% of instances, respectively. In contrast, the distance-based heuristic only attains such performance in and 3.3% and 29% of instances, respectively. We can therefore conclude that, for the test case studied here, the data mining heuristic dramatically outperforms the distance-based heuristic in the reduction of voltage magnitude violations.

Next, we conducted an identical analysis to the one described above studying the impact of the two heuristics on branch flow violations. Figure 3.3 shows the ECDF for both the data mining heuristic and the distance-based heuristic.

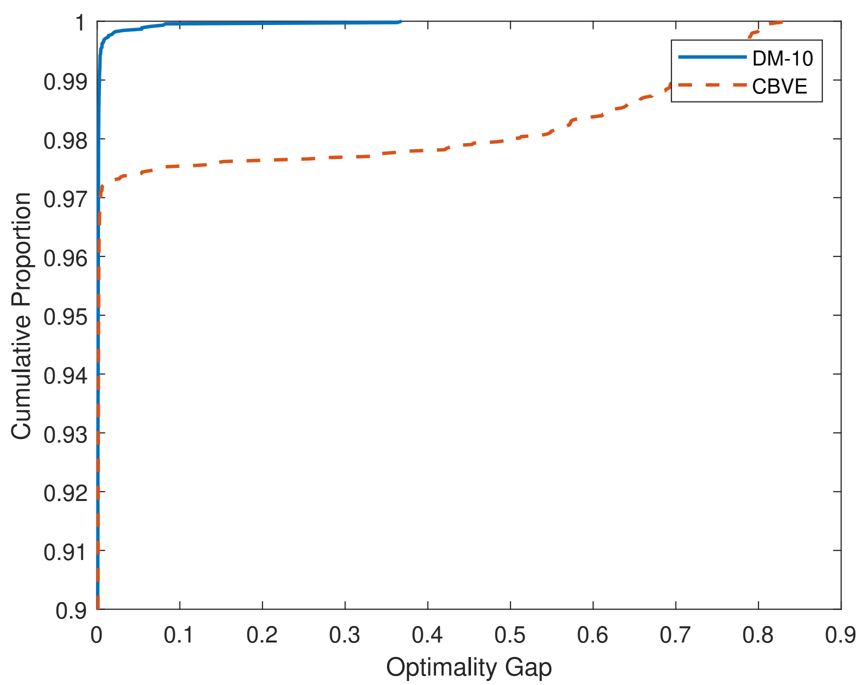

Figure 3.3: DM-10 performs more strongly in branch flow violations than CBVE

Figure 3.3 shows that, while the two heuristics are much more competitive in the case of branch flow violation reduction, the data mining heuristic is still the superior method. The first finding from Figure 3.3 is that both heuristics have a long vertical line at or near zero. This shows that both methods have strong performance in this case. Specifically, the data mining heuristic attains an optimality gap of less than 0.1% in over 92.8% of instances. The distance-based heuristic attains such a performance in over 89.5% of instances. If we increase the threshold to an optimality gap of one percent, the data mining heuristic identifies such a switch in over 99.6% of solution and the distance-based heuristic identifies such a switch in over 97.2% of solutions. These results, and those described by Figure 3.2, show that the proposed data mining heuristic definitively outperforms the distance based heuristic for the test case studied here.

3.6 Conclusion

This work developed a data mining heuristic to identify transmission switching candidates to reduce post-contingency voltage magnitude and branch flow violations. We used real-world, large-scale AC power system data to generate a robust data set to feed into our logistic regression model with guided undersampling. The resulting heuristic demonstrated considerable performance in identifying strong transmission switching solutions, even given the substantial size of the PJM system. Our methodology shows the ability of data mining methods to substantially reduce the workload associated with identifying strong transmission switching candidates for post-contingency violation reduction. While the specific predictive model (i.e., the features chosen during feature selection and the regression coefficients) may not be specifically applicable to every system, the methodology herein proposed should generate a model which exhibits strong performance on alternate data sets.

We first showed that the data mining heuristic has strong performance in terms of accuracy. Specifically, using only the top switch identified by the data mining heuristic, the proposed methodology attained an average optimality gap of 5.5% for voltage magnitude violations and 5.9% for branch flow violations. Moreover, for both violation types, the bulk of studied instances (92% and 93%, respectively) result in optimality gaps of less than 0.1%. We therefore conclude that the data mining heuristic can regularly identify optimal or near-optimal solutions. We also showed that the data mining heuristic substantially outperforms distance-based heuristics in terms of accuracy. Using a candidate pool of only ten switches, the proposed heuristic outperformed an existing distance-based heuristic in terms of average optimality gap by over eight times for voltage magnitude violations and over 35 times for branch flow violations. These findings show that data mining methods such as the data mining heuristic with guided undersampling developed herein are noteworthy techniques which can identify optimal and near-optimal candidate switching actions for the herein studied problem with extremely high regularity.

Our method demonstrates the strong performance that data mining methods can achieve in regard to power systems operation. Specifically, our proposed method is one of few which uses data mining techniques to address power systems operations. Moreover, ours is the first true data mining technique applied to transmission switching. As mentioned previously, transmission switching is only implemented in limited capacity because of concerns over computational complexity, uncertainty of AC performance, and scalability to real-world systems. Because our data mining heuristic is computationally inexpensive, addresses an AC system problem directly, and has been rigorously tested on real-world large-scale data, it addresses these three issues directly. Given the performance of our model, it should be strongly considered in the use of post-contingency violation reduction. More importantly, it should motivate the study and development of new data mining techniques to address this and similar power systems operations problems.

4. SUMMARY AND CONCLUSIONS

The principal objective of this research was to develop models and solution approaches that outperform the existing state-of-the-art methods for real-time power systems operations problems. Specifically, we sought to develop approaches that leverage the flexibility of the power grid such that grid operations are more efficient and may better respond to contingency events. Towards that end, we developed two solution approaches – one mixed-integer linear program and one data mining approach – which address two variants of the optimal transmission switching problem.

The first work discussed in this dissertation developed a relaxation to the AC optimal transmission switching (ACOTS) problem. The ACOTS is a mixed-integer nonlinear program which models the full complexity of utilizing transmission switching within the AC power grid. In the literature, the ACOTS has been addressed either via DC-based approaches, other linear models, or conic optimization. DC-based models have been shown to be highly inaccurate for transmission switching, particularly within real-time response applications. Other linear approaches prior to this work have proven either to be too slow or to contain drawbacks similar to those of DC-based models. Conic optimization approaches have shown the most promise thus far, but mixed-integer programming has stronger potential for scaling to problems of real-world size. Our model, which we named the mixed-integer programming model for AC power flows (MIPAC) substantially accelerated an existing transmission switching model from the literature which was slow to such a degree that it was outperformed by an exhaustive search. We accelerated this model by developing three key constraint modifications: an apparent power constraint relaxation, a novel cosine approximation, and finally a relaxation for the implicit power factor constraint. From a performance perspective, our constraint modifications took the existing model from a place of total impracticality to a level comparable with the state-of-the-art in terms of solution time. Moreover, our methodology improved upon the base model's ability to identify AC optimal and near optimal transmission switching actions. This work motivates the development of other mixed-integer linear approaches to the AC optimal transmission switching problem, which show stronger potential than conic optimization approaches to scale to problems of real-world size.

The second work in this dissertation introduced our data mining approach to address another variant of the optimal transmission switching problem. Specifically, we seek to utilize imbalanced-data classification techniques to identify transmission switching actions that minimize AC power flow violations stemming from contingency events. Our data mining approach is computationally inexpensive and was tested thoroughly on real-world AC power system data; thus, it directly addresses three issues identified in the transmission switching literature: computational complexity, impact on AC systems, and impact of problems of real-world size. Moreover, our methodology can be directly combined with existing approaches to address a fourth and final concern from the literature, transient stability. It is noteworthy that our methodology is the first true data mining approach to address a transmission switching problem. Our approach significantly outperformed existing state-of-the-art approaches used to address this problem, and because of its low computational cost, shows strong potential to be utilized in practice.

REFERENCES

[1] M. Majidi-Qadikolai and R. Baldick, “Integration of $n-1$ contingency analysis with systematic transmission capacity expansion planning: Ercot case study,” *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2234–2245, 2016.

[2] T. Potluri and K. W. Hedman, “Impacts of topology control on the acopf,” in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–7, IEEE, 2012.

[3] B. H. Chowdhury and S. Rahman, “A review of recent advances in economic dispatch,” *IEEE transactions on power systems*, vol. 5, no. 4, pp. 1248–1259, 1990.

[4] S. Frank, I. Steponavice, and S. Rebennack, “Optimal power flow: a bibliographic survey i,” *Energy Systems*, vol. 3, no. 3, pp. 221–258, 2012.

[5] S. Frank, I. Steponavice, and S. Rebennack, “Optimal power flow: a bibliographic survey ii,” *Energy Systems*, vol. 3, no. 3, pp. 259–289, 2012.

[6] M. Huneault and F. Galiana, “A survey of the optimal power flow literature,” *IEEE transactions on Power Systems*, vol. 6, no. 2, pp. 762–770, 1991.

[7] J. A. Momoh, R. Adapa, and M. El-Hawary, “A review of selected optimal power flow literature to 1993. i. nonlinear and quadratic programming approaches,” *IEEE transactions on power systems*, vol. 14, no. 1, pp. 96–104, 1999.

[8] J. A. Momoh, M. El-Hawary, and R. Adapa, “A review of selected optimal power flow literature to 1993. ii. newton, linear programming and interior point methods,” *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 105–111, 1999.

[9] S. H. Low, “Convex relaxation of optimal power flow—part i: Formulations and equivalence,” *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.

[10] S. H. Low, "Convex relaxation of optimal power flow—part ii: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, 2014.

[11] K. W. Hedman, S. S. Oren, and R. P. O'Neill, "A review of transmission switching and network topology optimization," in *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–7, IEEE, 2011.

[12] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin de la Societe Francoise des Electriciens*, vol. 3, no. 8, pp. 431–447, 1962.

[13] F. Capitanescu, M. Glavic, D. Ernst, and L. Wehenkel, "Interior-point based algorithms for the solution of optimal power flow problems," *Electric Power systems research*, vol. 77, no. 5-6, pp. 508–517, 2007.

[14] Q. Jiang, H.-D. Chiang, C. Guo, and Y. Cao, "Power–current hybrid rectangular formulation for interior-point optimal power flow," *IET generation, transmission & distribution*, vol. 3, no. 8, pp. 748–756, 2009.

[15] T. Saha and A. Maitra, "Optimal power flow using the reduced newton approach in rectangular coordinates," *International Journal of Electrical Power & Energy Systems*, vol. 20, no. 6, pp. 383–389, 1998.

[16] G. L. Torres and V. H. Quintana, "An interior-point method for nonlinear optimal power flow using voltage rectangular coordinates," *IEEE transactions on Power Systems*, vol. 13, no. 4, pp. 1211–1218, 1998.

[17] W. Yan, J. Yu, D. Yu, and K. Bhattarai, "A new optimal reactive power flow model in rectangular form and its solution by predictor corrector primal dual interior point method," *IEEE transactions on power systems*, vol. 21, no. 1, pp. 61–67, 2006.

[18] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on power apparatus and systems*, no. 10, pp. 1866–1876, 1968.

[19] J. Peschon, D. W. Bree, and L. P. Hajdu, "Optimal power-flow solutions for power system planning," *Proceedings of the IEEE*, vol. 60, no. 1, pp. 64–70, 1972.

[20] R. Burchett, H. Happ, and K. Wirgau, "Large scale optimal power flow," *IEEE Transactions on Power Apparatus and Systems*, no. 10, pp. 3722–3732, 1982.

[21] D. I. Sun, B. Ashley, B. Brewer, A. Hughes, and W. F. Tinney, "Optimal power flow by newton approach," *IEEE Transactions on Power Apparatus and systems*, no. 10, pp. 2864–2880, 1984.

[22] A. Santos, S. Deckmann, and S. Soares, "A dual augmented lagrangian approach for optimal power flow," *IEEE transactions on power systems*, vol. 3, no. 3, pp. 1020–1025, 1988.

[23] Y. Hong, "Enhanced newton optimal power flow approach: experiences in taiwan power system," in *IEE Proceedings C (Generation, Transmission and Distribution)*, vol. 139, pp. 205–210, IET, 1992.

[24] W. Min and L. Shengsong, "A trust region interior point algorithm for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 27, no. 4, pp. 293–300, 2005.

[25] A. A. Sousa and G. L. Torres, "Globally convergent optimal power flow by trust-region interior-point methods," in *2007 IEEE Lausanne Power Tech*, pp. 1386–1391, IEEE, 2007.

[26] H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185–1193, 2007.

[27] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *IEEE transactions on power apparatus and systems*, no. 3, pp. 745–751, 1974.

[28] A. D. Papalexopoulos, C. F. Imparato, and F. F. Wu, "Large-scale optimal power flow: effects of initialization, decoupling and discretization," *IEEE Transactions on Power Systems*, vol. 4, no. 2, pp. 748–759, 1989.

[29] G. Contaxis, C. Delkis, and G. Korres, "Decoupled optimal load flow using linear or quadratic programming," *IEEE Transactions on Power Systems*, vol. 1, no. 2, pp. 1–7, 1986.

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

[31] C. Coffrin and P. Van Hentenryck, "A linear-programming approximation of ac power flows," *INFORMS Journal on Computing*, vol. 26, no. 4, pp. 718–734, 2014.

[32] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 383–392, 2008.

[33] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE transactions on power systems*, vol. 21, no. 3, pp. 1458–1459, 2006.

[34] E. B. Fisher, R. P. Neill, and M. C. Ferris, "Optimal transmission switching," *Power Systems, IEEE Transactions on*, vol. 23, no. 3, pp. 1346–1355, 2008.

[35] K. W. Hedman, R. P. O'Neill, E. B. Fisher, and S. S. Oren, "Optimal transmission switching—sensitivity analysis and extensions," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1469–1479, 2008.

[36] K. W. Hedman, R. P. O'Neill, E. B. Fisher, and S. S. Oren, "Optimal transmission switching with contingency analysis," *Power Systems, IEEE Transactions on*, vol. 24, no. 3, pp. 1577–1586, 2009.

[37] P. A. Ruiz, J. M. Foster, A. Rudkevich, and M. C. Caramanis, "Tractable transmission topology control using sensitivity analysis," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1550–1559, 2012.

[38] A. S. Korad and K. W. Hedman, "Robust corrective topology control for system reliability," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4042–4051, 2013.

[39] A. R. Escobedo, E. Moreno-Centeno, and K. W. Hedman, "Topology control for load shed recovery," *Power Systems, IEEE Transactions on*, vol. 29, no. 2, pp. 908–916, 2014.

[40] X. Li, P. Balasubramanian, M. Sahraei-Ardakani, M. Abdi-Khorsand, K. W. Hedman, and R. Podmore, "Real-time contingency analysis with corrective transmission switching," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2604–2617, 2017.

[41] M. Khanabadi, H. Ghasemi, and M. Doostizadeh, "Optimal transmission switching considering voltage security and n-1 contingency analysis," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 542–550, 2013.

[42] Y. Sang and M. Sahraei-Ardakani, "The interdependence between transmission switching and variable-impedance series facts devices," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2792–2803, 2018.

[43] J. Shi and S. S. Oren, "Stochastic unit commitment with topology control recourse for power systems with large-scale renewable integration," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3315–3324, 2018.

[44] J. C. Villumsen, G. Brønmo, and A. B. Philpott, "Line capacity expansion and transmission switching in power systems with large-scale wind power," *IEEE Transac-*

tions on Power Systems, vol. 28, no. 2, pp. 731–739, 2013.

[45] P. Henneaux and D. S. Kirschen, "Probabilistic security analysis of optimal transmission switching," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 508–517, 2016.

[46] M. Abdi-Khorsand, M. Sahraei-Ardakani, and Y. M. Al-Abdullah, "Corrective transmission switching forn-1-1 contingency analysis," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1606–1615, 2017.

[47] B. Kocuk, H. Jeon, S. S. Dey, J. Linderoth, J. Luedtke, and X. A. Sun, "A cycle-based formulation and valid inequalities for dc power transmission problems with switching," *Operations Research*, vol. 64, no. 4, pp. 922–938, 2016.

[48] B. Kocuk, S. S. Dey, and X. A. Sun, "New formulation and strong misocp relaxations for ac optimal transmission switching problem," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4161–4170, 2017.

[49] S. Fattahi, J. Lavaei, and A. Atamtürk, "A bound strengthening method for optimal transmission switching in power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 280–291, 2019.

[50] S. Fattahi, J. Lavaei, and A. Atamtürk, "Promises of conic relaxations in optimal transmission switching of power systems," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 3238–3245, IEEE, 2017.

[51] R. A. Jabr, R. Singh, and B. C. Pal, "Minimum loss network reconfiguration using mixed-integer convex programming," *IEEE Transactions on Power systems*, vol. 27, no. 2, pp. 1106–1115, 2012.

[52] Y. Bai, H. Zhong, Q. Xia, and Y. Wang, "A conic programming approach to optimal transmission switching considering reactive power and voltage security," in *2015*

IEEE Power & Energy Society General Meeting, pp. 1–5, IEEE, 2015.

[53] M. Soroush and J. D. Fuller, "Accuracies of optimal transmission switching heuristics based on dcopf and acopf," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 924–932, 2014.

[54] M. Khanabadi and H. Ghasemi, "Transmission congestion management through optimal transmission switching," in *2011 IEEE Power and Energy Society General Meeting*, pp. 1–5, IEEE, 2011.

[55] Y. Bai, H. Zhong, Q. Xia, and C. Kang, "A two-level approach to ac optimal transmission switching with an accelerating technique," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1616–1625, 2017.

[56] C. Coffrin and P. Van Hentenryck, "Transmission system restoration with co-optimization of repairs, load pickups, and generation dispatch," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 144–154, 2015.

[57] P. Dehghanian, Y. Wang, G. Gurrala, E. Moreno-Centeno, and M. Kezunovic, "Flexible implementation of power system corrective topology control," *Electric Power Systems Research*, vol. 128, pp. 79–89, 2015.

[58] A. S. Korad, P. Balasubramanian, and K. W. Hedman, "Robust corrective topology control," *Handbook of Clean Energy Systems*, 2015.

[59] P. Balasubramanian and K. W. Hedman, "Real-time corrective switching in response to simultaneous contingencies," *Journal of Energy Engineering*, vol. 141, no. 1, p. B4014003, 2014.

[60] M. A. Khorsand and K. W. Hedman, "Day-ahead corrective transmission topology control," in *PES General Meeting| Conference & Exposition, 2014 IEEE*, pp. 1–5, IEEE, 2014.

[61] R. A. Sarker and C. S. Newton, *Optimization modelling: a practical approach*. CRC press, 2007.

[62] R. Fourer, "A simplex algorithm for piecewise-linear programming ii: Finiteness, feasibility and degeneracy," *Mathematical Programming*, vol. 41, no. 1-3, pp. 281–315, 1988.

[63] M. Johansson, *Piecewise linear control systems*. PhD thesis, Ph. D. Thesis, Lund Institute of Technology, Sweden, 1999.

[64] M. Jabarnejad, J. Wang, and J. Valenzuela, "A decomposition approach for solving seasonal transmission switching," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1203–1211, 2015.

[65] M. Khanabadi, Y. Fu, and C. Liu, "Decentralized transmission line switching for congestion management of interconnected power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5902–5912, 2018.

[66] F. Capitanescu and L. Wehenkel, "An ac opf-based heuristic algorithm for optimal transmission switching," in *2014 Power Systems Computation Conference*, pp. 1–6, IEEE, 2014.

[67] T. Akbari and M. T. Bina, "A linearized formulation of ac multi-year transmission expansion planning: A mixed-integer linear programming approach," *Electric Power Systems Research*, vol. 114, pp. 93–100, 2014.

[68] J. A. Taylor, *Convex optimization of power systems*. Cambridge University Press, 2015.

[69] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *Power Systems, IEEE Transactions on*, vol. 26, no. 1, pp. 12–19, 2011.

[70] R. Christie, "Power systems test case archive," *Electrical Engineering dept., University of Washington*, 2000.

[71] C. P. Barrows, *Flexible network topologies and the smart grid in electric power systems*. The Pennsylvania State University, 2013.

[72] H. Ahmadi, M. Khanabadi, and H. Ghasemi, "Transmission system reconfiguration for congestion management ensuring transient and voltage stability," in *2013 13th International Conference on Environment and Electrical Engineering (EEEIC)*, pp. 22–26, IEEE, 2013.

[73] Q. Zhou, D. Shirmohammadi, and W.-H. Liu, "Distribution feeder reconfiguration for service restoration and load balancing," *IEEE Transactions on Power Systems*, vol. 12, no. 2, pp. 724–729, 1997.

[74] X. Geng, L. Xie, and D. Obadina, "Chance constrained optimal reactive power dispatch," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2018.

[75] W. E. Brown and E. Moreno-Centeno, "Transmission-line switching for load shed prevention via an accelerated linear programming approximation of ac power flows," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2575–2585, 2020.

[76] K. Sung, W. E. Brown, E. Moreno-Centeno, and Y. Ding, "Gum: an imbalanced data preprocessing method using majority and minority outlier elimination and distribution-preserving majority sampling," *Submitted*, 2019.

[77] X. Li, P. Balasubramanian, and K. Hedman, "A data-driven heuristic for corrective transmission switching," in *2016 North American Power Symposium (NAPS)*, pp. 1–5, IEEE, 2016.

[78] M. Sahraei-Ardakani, A. Korad, K. W. Hedman, P. Lipka, and S. Oren, "Performance of ac and dc based transmission switching heuristics on a large-scale polish system," in *2014 IEEE PES General Meeting| Conference & Exposition*, pp. 1–5, IEEE, 2014.

[79] J. Shi and S. S. Oren, "A data mining approach for real-time corrective switching," in *2015 IEEE Power & Energy Society General Meeting*, pp. 1–5, IEEE, 2015.

[80] J. Liu, G. M. Huang, Z. Ma, and Y. Geng, "A novel smart high-voltage circuit breaker for smart grid applications," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 254–264, 2011.

[81] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Departmental Papers (CIS)*, p. 107, 2000.

[82] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*, vol. 112. Springer, 2013.

[83] X. LI, P. BALASUBRAMANIAN, M. ABDI-KHORSAND, A. Korad, and K. W. HEDMAN, "Effect of topology control on system reliability: Tva test case," in *2014 Cigre US National Committee Grid of the Future Symposium*, vol. 19.

[84] N. Japkowicz, "The class imbalance problem: Significance and strategies," in *Proc. of the Int'l Conf. on Artificial Intelligence*, 2000.

[85] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *Journal of Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.