2017

Synthetic Modeling of Power Grids Based on Statistical Analysis

Seyyed Hamid Elyas 8045266

Synthetic Modeling of Power Grids Based on Statistical Analysis

In partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

Seyyed Hamid Elyas

Director: Dr. Zhifang Wang

Assistant professor, Electrical and Computer Engineering

Virginia Commonwealth University

Richmond, Virginia

May, 2017

Acknowledgments

I would like to express my gratitude to all those who helped me with various aspects in conducting and writing up this dissertation. First and foremost, I am thankful to Professor Zhifang Wang for her advice and support throughout my Ph.D. work. Special thanks to my defense/candidacy committee members for their valuable insights and contributions to the content and the improvement of this thesis. Further, I am thankful for the collaboration opportunities and the financial support from the Electric Power and Energy System (EPES) Lab, and the Electrical and computer department at Virginia Commonwealth University. I would also like to thank my friends and colleagues. They were a great comfort and encouragement throughout my graduate academic career and thesis writing process.

And last, but certainly not least, I am eternally grateful to my parents who have always loved me and encouraged me to constantly distinguish myself and be the best I can be. Without them, I would have never had the educational opportunities I had. And a big shout-out to my brother and sister who have been with me from day one, and to my fiancé, who has patiently supported me through this journey from start to finish.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

SYNTHETIC MODELING OF POWER GRIDS BASED ON STATISTICAL ANALYSIS

By Seyyed Hamid Elyas, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, at Virginia Commonwealth University.

Virginia Commonwealth University, 2017.

Director: Zhifang Wang, Ph.D.

The development of new concepts and methods for improving the efficiency of power networks needs performance evaluation with realistic grid topology. However, much of the realistic grid data needed by researchers cannot be shared publicly due to the security and privacy challenges. With this in mind, power researchers studied statistical properties of power grids and introduced synthetic power grid topology as appropriate methodology to provide enough realistic power grid case studies. If the synthetic networks are truly representative and if the concepts or methods test well in this environment they would test well on any instance of such a network as the IEEE model systems or other existing grid models.

In the past, power researchers proposed a synthetic grid model, called RT-nested-smallworld, based on the findings from a comprehensive study of the topology properties of a number of realistic grids. This model can be used to produce a sufficiently large number of power grid test cases with scalable network size featuring the same kind of small-world topology and electrical characteristics found in realistic grids. However, in the proposed RT-nested-smallworld model the approaches to address some electrical and topological settings such as (1) bus types assignment,

(2) generation and load settings, and (3) transmission line capacity assignments, are not sufficient enough to apply to realistic simulations. In fact, such drawbacks may possibly cause deviation in the grid settings therefore give misleading results in the following evaluation and analysis.

To address this challenges, the first part of this thesis proposes a statistical methodology to solve the bus type assignment problem. This method includes a novel measure, called the Bus Type Entropy, the derivation of scaling property, and the optimized search algorithm. The second part of this work includes a comprehensive study on generation/Load settings based on both topology metrics and electrical characteristics. In this section a set of approaches has been developed to generate a statistically correct random set of generation capacities and assign them to the generation buses in a grid. Then we determine the generation dispatch of each generation unit according to its capacity and the dispatch ratio statistics, which we collected and derived from a number of realistic grid test cases. The proposed approaches is readily applied to determining the load settings in a synthetic grid model and to studying the statistics of the flow distribution and to estimating the transmission constraint settings. Considering the results from the first two sections, the third part of this thesis will expand earlier works on the RT-nested-smallworld model and develop a new methodology to appropriately characterize the line capacity assignment and improve the synthetic power grid modeling.

Chapter 1

Introduction

It is widely agreed that electric power grid is one of the most critical infrastructures whose reliable and efficient operation plays a vital role for the essential functioning of our society [1]. For this reason the access of realistic power system data is highly restricted due to many security concerns. On the other hand, grid data of the right type and fidelity is crucial for the advance of any power system related researches to test new concepts and methods. In the past, power engineers and researchers mainly depended on small number of historical test systems such as IEEE case studies (see Figure 1.1) or other existing realistic grid models [2]. Obviously, being able to generate a large number of power grid test cases with realistic topologies could be an essential step towards improving the verification and validation of new concepts. With this in mind, researchers studied statistical properties of power grids and introduced synthetic grid network topology as appropriate methodology to provide enough realistic power grid case studies. In synthetic network research area the main idea is that by studying the statistical properties of real networks a method can be constructed to generate fictional networks that have all the properties of a real network. It should be mentioned here that the generated fictional network is completely different from the reduced abstraction of largescale power grids [3]. Technically, although the dynamic reduction methods change the size of the network, the generated network cannot be considered as a synthetic representation for the original power grid [3]-[6]. The main reason comes

from the fact that in synthetic power grid modeling the physical and electrical characteristics of the original network are not fully available. Besides, network reduction techniques extract a reduced abstraction of large-scale power grids, and cannot generate a larger network like we can observe in synthetic power grid modeling.

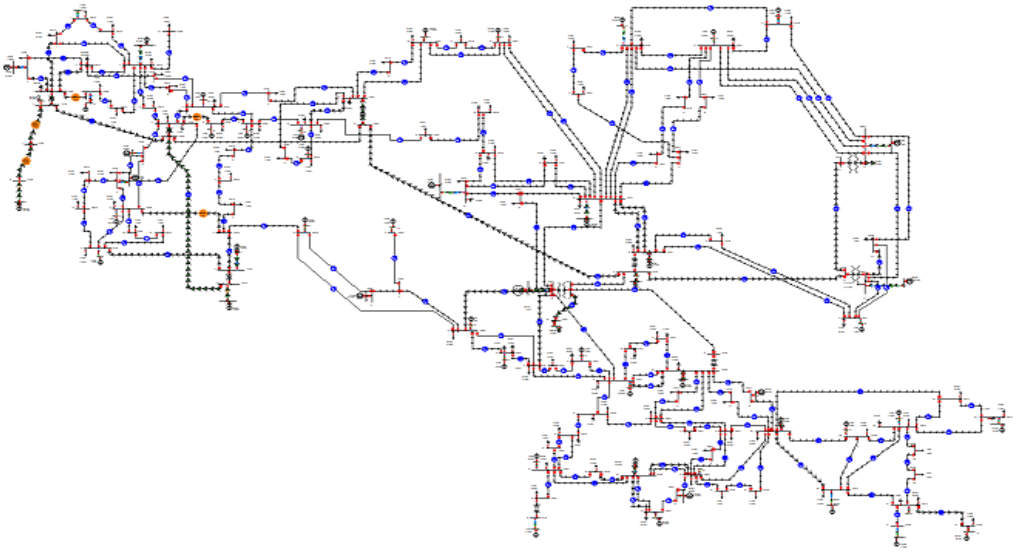


Figure 1.1: IEEE-118 bus system (source: www.ee.washington.edu)

1.1 Power system model

The overall purpose of power systems is to supply electricity to consumers in a safe, reliable, and economic way. The primary structure of traditional power systems comprises power generation, transmission and distribution to consumers, or loads (Figure 1.2). A so-called hierarchical, vertical structure is based upon a limited number of large, central power plants delivering electricity to a large number of loads [7]. Power flows from generation into high-voltage

transmission networks and then into medium- and low-voltage distribution networks, hence only in a top-down, 'vertical' direction. The advantages of interconnected, vertically integrated power systems include economies of scale in power generation, increased reliability, a reduction of reserve margins and aggregation of load variations. Presently, increasing amounts of distributed generation are connected to the low-voltage networks. This trend increasingly leads to bi-directional power flows in the distribution system [8]. In observing the primary structure of power systems, it is important to note that electrical energy as such cannot be stored in significant amounts. Electrical power is consumed at the same moment it is generated. For a reliable power supply it is therefore essential to maintain a precise balance between demand (total system load including transmission and distribution losses) and generation. It is in principle possible to maintain the power balance by adjusting both generation and demand, but historically, mostly the central generation units have been used to follow the demand at all times. The operation of power systems is therefore critically dependent on the capabilities of generators for balancing the load.

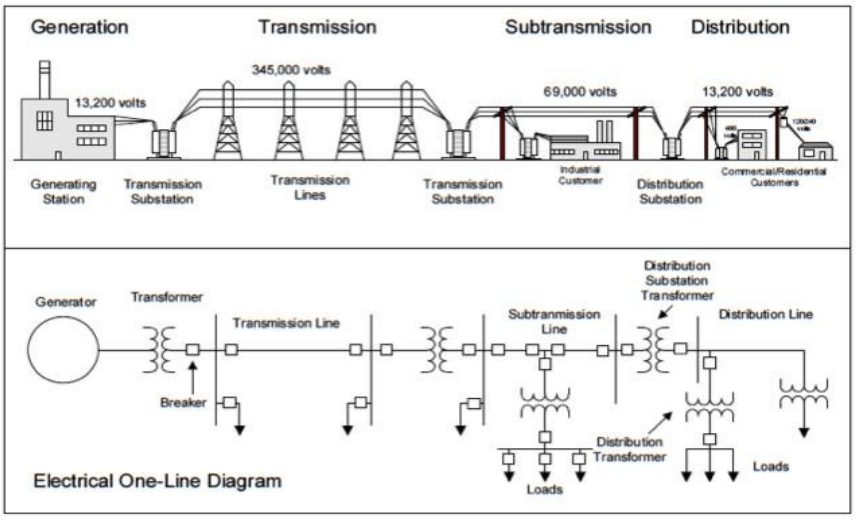


Figure 1.2: Overview of a traditional power system structure [9].

One of the mathematical objects allowing a complete definition of the power system topology is the Adjacency (A) and Laplacian (L) Matrixes. The connecting topology of an undirected graph network with n nodes (representing buses in the grid) and *m* links (representing transmission lines in the grid) is fully described by a Laplacian matrix which can be obtained as

$$L = A^T A \tag{1.1}$$

where *A* is the adjacency matrix. For a given power network with undirected and unitary links, it is defined as $A_{ij} = 1$ if nodes *i* and *j* are connected, 0 otherwise.

On the other hand, a power grid is more than a graph network. Therefore its power-transmission topology can be represented by an admittance matrix $Y_{n\times n}$ which is defined as $Y = A^T \Lambda^{-1}(z_l)A$. Where $\Lambda^{-1}(.)$ denotes the diagonal inverse matrix with a specified vector and z_l the vector of branch impedances in a grid. Comparing the mentioned above equations, one may find that the admittance matrix of a power grid can be viewed as a complex weighted Laplacian. Obviously, the definition of admittance matrix *Y* neglects the shunts in a grid and can be corrected by revising its diagonal entries. However, ignoring the shunts will not cause big errors in some grid analysis like the AC power flow or those only concerning real power transmission such as the DC power flow approximation, a standard approach widely used in optimizing flow dispatch and for assessing line overloads [10-12].

Considering a power grid transmission network with n buses interconnected by *m* branches, the flow distribution follows the network constraints as

$$P(t) = B'(t)\theta(t) \tag{1.2}$$

$$F(t) = \Lambda(y_l)A\theta(t) \tag{1.3}$$

where $P(t) = [P_G(t), -P_L(t), P_C]^T$ represents the vector of injected real power from generation, load and connection buses and obviously the power injection from connection buses equals zero, i.e., $P_C = 0$. $\theta(t)$ is the vector of phase angles, and $F(t)$ the vector of real-power delivered along the branches. The matrix $B'(t)$ is defined as $A^T \Lambda(y_l) A$, where $y_l = 1/x_l$ with $xl$the branch series reactance and $\Lambda(y_l)$ represents a diagonal matrix with entries of $\{y_l, l = 1,2, \dots, m\}$. Alternatively we can have $B' = -Img(Y)$ where the shunts of the grid neglected in the admittance matrix. Note that in the vector injected power as above, we assume that buses in a grid have been reordered so that the generation, load, and connection buses could be grouped as listed. Clearly the location of generators and loads plays a vital role in grid operation. The topology location of generators and loads in a grid can be fully determined by a vector of bus type assignments of generation, load, or connection.

Besides the network constraints, grid operation also needs to account for the constraints of generation capacity, load settings, and transmission capacity, such as

$$P_G^{Min} \leq P_G \leq P_G^{Max} \tag{1.4}$$

$$P_L^{Min} \leq P_L \leq P_L^{Max} \tag{1.5}$$

$$F^{Min} \leq F_l \leq F^{Max} \tag{1.6}$$

From what is presented above it is clear that the dynamics of a power grid not only depend on the "electrical" topology but also the generation and load settings including their locations. The location setting of generation and loads is equivalent to the bus type assignment in power grid modeling. Generally speaking, all the buses in a grid can be grouped into three categories as follows with minor overlaps since it is possible that a small portion of buses may belong to more than one categories: (G) the generation buses which connect generators, (L) the load buses which

support custom demands, (C) the connection buses which form the transmission network. From the recent study [13], it is found that here exists nontrivial correlation between some network topology metrics and the three bus types in a grid. The study results in [13] imply that random permutation of the bus type assignment in a grid may dramatically changes the grid dynamics and makes it behave no longer like a power grid. This discovery has led to the efforts to define an effective measure to characterize the bus type assignments of realistic power grids.

1.2 The scaling property of grid connecting topology

The connecting topology of a power grid can be fully described by its admittance matrix which contains both graph topology and electric parameters. Compared with the topology of other natural or man-made networks, power grid topology is very different in many ways [14]-[25], such as the salient small-world properties characterized by shorter average path length and higher clustering coefficients than those of an Erdos-Renyi random graphs [25] with the same network size and total number of links. The average node degree of a typical power grid does not scale with the network size but remains within a very strict range. Besides, the node degrees approximates a statistical distribution of a truncated geometric random variable with some mixture of an irregular discrete. The algebraic connectivity of a grid also exhibits some special scaling property [17]. Another important property of power grid is its heavy-tailed distribution of line impedances, which is well-fitted by a clipped double-ParetologNormal (dPlN) distribution [26].

The node degree of bus i in a grid equals the total number of branches it connects and can be obtained from the ith diagonal entry of the Laplacian matrix, $i.e.,\ k_i = L(i,i)$. Then the average nodal degree of the grid is

$$< k >= \frac{1}{n}\sum_{i=1}^{n} L(i,i) \tag{1.7}$$

Given the connecting topology of a grid, we can run the Dijkstra's algorithm to calculate the shortest path length measured in hops between any two buses i and j, i.e., l_{ij}. Then the average shortest path length of a grid is

$$< l >= \frac{2\sum_{i,j} l_{ij}}{n(n-1)} \tag{1.8}$$

Another important topology measure is the second smallest eigenvalue of the Laplacian matrix, $\lambda_2(L)$, called the algebraic connectivity, with

$$[\lambda_1, \lambda_2, \dots, \lambda_n] = Eigen\ (L) \tag{1.9}$$

As a fact the smallest eigenvalue of the Laplacian is always zero, i.e., $\lambda_1(L) \equiv 0$ and the number of times that 0 appears as an eigenvalue in the Laplacian is the total number of islanded components in the network. $\lambda_2(L)$ reflects the overall connectivity of a network and how fast information data can be broadcast across it. The eigenvalue $\lambda_2(L)$is greater than 0 if and only if network is a connected graph. If the algebraic connectivity $\lambda_2(L)$is close to zero, the network is close to being disconnected. Otherwise, if $\lambda_2(L)/n$ gets close to 1, where N is the network size, the grid tends to be a fully connected topology.

Table 1.1 presents some topology measures evaluated on the IEEE test cases and other real-world grids of different network size, where the PEGASE systems represent some European nation's grid at different levels of network reduction, the NYSIO system partially represents the

New York interconnect in the US, and the RTE system is an equivalent of the French Grid. Figure 1.1 shows the average node degree (1.7) of each grid system versus its network size. Obviously the average node degree of most sample grids does not scale but staying within a very stable region as $\langle k \rangle \in [2.0; 3.5]$. Two exceptions are the PEGASE 89-bus system with $\langle k \rangle = 4.72$ and the NYISO 2935-bus system with $\langle k \rangle = 4.47$, which are unusually higher than the average node degree observed in the rest grid systems. This may be caused by the original grid's denser connecting topology or by the equalization approaches used in the network reduction.

Figure 1.4 depicts the average path length in hops, as defined in (10), versus the network size for each sample grid in Table 1.1, where the dashed line represents an approximate fitting curve of the observed scaling property as $\langle l \rangle \propto 6.205\ log\ n$. Note that for the purpose of simplicity, the logarithm in this section is with base 10.

Table1.1: Topology measures of real-world power grids

	$(n,\ m)$	$\langle k \rangle$	$\langle l \rangle$	λ_2
IEEE-24	(24,38)	3.17	3.21	0.2132
IEEE-30	(30,41)	2.73	3.31	0.2121
NE-39	(39,46)	2.36	4.75	0.0762
IEEE-57	(57,80)	2.81	4.95	0.0882
PEGASE-89	(89,210)	4.72	3.87	0.1537
IEEE-118	(118,186)	3.15	6.31	2.71e-2
IEEE-300	(300,409)	2.73	9.94	9.38e-3
PEGASE-1354	(1354,1991)	2.90	11.20	6.59e-3
PEGASE-2869	(2869,4582)	3.19	20.00	6.23e-4
NYISO-2935	(2935,6567)	4.47	16.43	1.42e-3
RTE-6515	(6515,9037)	2.77	14.95	1.93e-3

Figure 1.5 plots the algebraic connectivity, as defined in (1.9), scaling curve of power grid versus network size. We can compares it with that of 1-Dimensional and 2-Dimensional lattices where 1D-lattice is a ring structured topology, with nodes connected with most adjacent neighbors

on both sides. 2D-lattice is a regular two-dimension meshed grid with each boundary side merging with the other side and each node connected to the most adjacent neighbors around it. For 1- D lattice, its connectivity scales as $\lambda_2(L) \propto n^{-2}$; for 2-D lattice, its connectivity grows as $\lambda_2(L) \propto n^{-1}$; interestingly, for power grids, its connectivity grows as $\lambda_2(L) \propto n^{-1.041}$, lying between those of 1-D lattice and 2-D lattice.

Figure 1.6 presents a scatter plot of the algebraic connectivity and the average path length of tested power grids, which exhibits strong correlation between the two measures. In fact there exists an approximate fitting function as $log(\lambda_2) \propto -0.1678\langle l \rangle$.

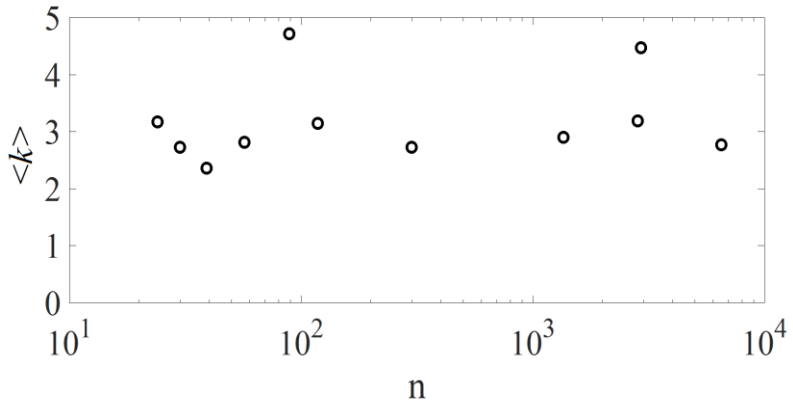


Figure 1.3: Average node degree versus the network size

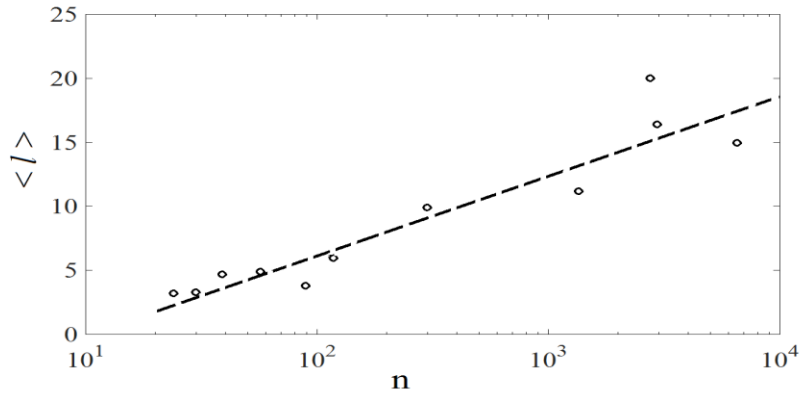


Figure 1.4: Average path length versus the network size

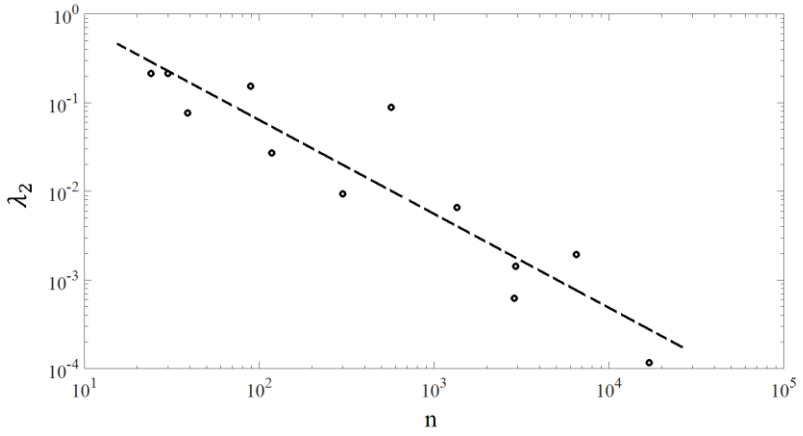


Figure 1.5: Algebraic connectivity versus the network size

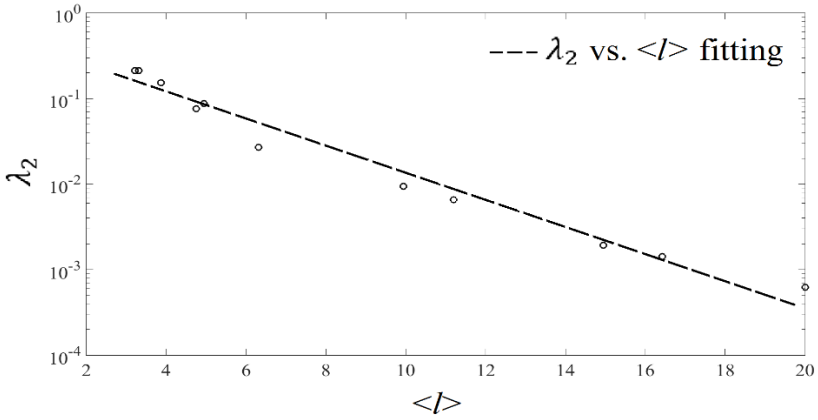


Figure 1.6: Algebraic connectivity versus average path length

1.3 Synthetic power grid modeling

Several emerging issues, including the resiliency of electric power delivery during extreme weather events, expanding use of distributed generation, the rapid growth of renewable generation [27-29] and the economic benefits of improved grid efficiency and flexibility, are challenging the way electricity is delivered from suppliers to consumers. This grid of the future requires advances in transmission and distribution system management with algorithms to control and optimize how

power is transmitted and distributed on the grid. However, the development of these systems has been hindered because the research community lacks high-fidelity, public, large-scale power system models that realistically represent current and evolving grid characteristics. Due to security and privacy concerns, much of the real data needed to test and validate new tools and techniques is restricted. To help drive additional innovation in the electric power industry, there is a need for grid models that mimic the characteristics of the actual grid, but do not disclose sensitive information. These models, say synthetic power grid models, will have the detail required to allow the successful development and testing of transformational power system optimization and control algorithms, including new Optimal Power Flow (OPF) algorithms [30-32]. Synthetic power grids are created in order to form randomly generated but realistic power grids. Currently we model only high-voltage transmission networks. The system topology and some electrical settings are created with specific random distribution functions and the generated topology preserves both simpleness and connectedness.

The development of algorithms for generating an efficient synthetic power grid requires comprehensive study on electrical and topological characteristics of real world power grids. For example, references [33] and [34] proposed a Tree-topology power grid model to study power grid robustness and to detect critical points and transitions in transmission flows to cause cascading failure blackouts. Reference [35] used Ring-structured power grid topologies to study the pattern and speed of contingency or disturbance propagation. Reference [14] proposed the first statistically modeling a power grid as a small-world network. Reference [16] used a small-world graph model to study the intrinsic spreading mechanism of the chain failure in a large-scale grid. Reference [17] provided a statistical model for power networks in an effort to grasp what class of communication network topologies need to match an underlying power network. All these models provide useful

perspectives of power grid characteristics. However, the topology of the generated power grids, such as the ring or tree-like structures and the small-world graph networks, fails to accurately or fully represent a realistic power system, especially its distinct sparse connectivity and scaling property versus the grid size. The worked mentioned above provide a very useful foundation to investigate the topological structure of power networks. However, power networks are much more than a graph topology and we need to consider realistic electrical and topological parameter setting to fully represent a realistic power system.

The network models that are probably most popular in the literature are random graphs based on the algorithms of Erdo"s-Re'nyi [25] or Gilbert [36], the small-world model from Watts and Strogatz [14], and finally the scale-free networks based on Baraba'si and Albert's preferential attachment [37]. The simplest approach to create spatially embedded networks are so-called random geometric networks [38] where node locations are drawn randomly from the unit square and nodes *i, j* are linked if their distance *dij* is below a threshold ε. However, such graphs are not necessarily connected, their average node degree grows as $\approx 4N\varepsilon 2$ instead of being constant, their node degree distribution is Poissonian, thus decaying faster than exponentially, and their clustering coefficient is larger than those of power grids. A model with similar properties is the Waxman model [39], where two nodes are linked with an exponential probability.

Scale-free networks on the other hand have much more heavy-tailed, power-law degree distributions, and their mean degree is not continuously adjustable. Baraba'si and Albert [37] show however that, in the case of equal (non-preferential) connection probabilities, their growth model produces exponentially decaying degree distributions. The Watts-Strogatz small-world algorithm [14] with its rewiring parameter *p* is often used in studies that focus on the transition from regular to random topologies, but it is not suited to generate power grid topologies.

Any desired degree distribution, including an exponentially decaying one, can be generated using the so-called configuration model [40-41] in which node degrees are prescribed endogenously for all nodes and the links are generated respecting these degrees but in an otherwise random manner. Still, this does not ensure the correct behavior of other statistical network measures and the construction mechanism does not seem a plausible assumption for the case of power grids where the exponential decay is very likely not prescribed by design but rather emerges endogenously.

The "RT-nestedsmallworld"-model [17] is to our knowledge the first attempt to design a model that especially matches the statistical properties of power grid topologies, by combining and modifying standard components of existing network models. This model constructs a largescale power grid using a hierarchical way: first form connected subnetworks with size limited by the connectivity requirement; then connect the subnetworks through lattice connections; finally, generate the line impedances from some specific distribution and assign them to the links in the topology network. The hierarchy in the model arises from observation of real-world power grids: usually a large-scale system consists of a number of smaller-size subsystem (e.g., control zones), which are interconnected by sparse and important tie lines. RT-nestedSmallworld model mainly contains three components: a) clusterSmallWorld subnetwork; b) lattice connections; and c) generation and assignment of line impedances [42].

ClusterSmallWorld subnetwork : Power grid topology has small-world characteristics; it is sparsely connected with a low average nodal degree not scaling with the network size. On the other side, in order for a small-world model to generate a connected topology, the network size has to be limited. In this model, different mechanisms from that of Watts–Strogatz small-world model have been adopted to form a power grid subnetwork in order to improve its resulting connectivity,

as shown in the following paragraphs. Consequently, the connectivity limitation on the network size can be expanded from what is indicated by (23). The experiments have shown that: for$\langle k \rangle = 2 \sim 3$, the network size should be limited no greater than 30; and for$\langle k \rangle = 4 \sim 5,300$. Therefore, the first step of this new model is to select the size of subnetworks according to connectivity limitation. Then a topology is built up through a modified small-world model, called clusterSmallWorld.

Lattice connections: In this step lattice connections are selected at random from neighboring subnetworks to form a whole large-scale power grid network. The number of lattice connections between neighboring subnetworks is chosen to be an integer around $\langle k \rangle$.

Generation and assignment of line impedances: In this part a number of line impedances are generated from a specified heavy-tailed distribution, and then sorted by magnitude and group into local links, rewire links, and lattice connection links according to corresponding portions. Finally, line impedances in each group are assigned at random to the corresponding group of links in the topology.

1.4 The big picture

Statistics is the science of collecting, analyzing and making an inference from data. Statistics is a particularly useful branch of mathematics that is not only studied theoretically by advanced mathematicians but one that is used by researchers in many fields to organize and analyze, and data. Statistical methods and analyses are often used to communicate research

findings and to support hypotheses and give credibility to research methodology and conclusions. The major purpose of statistics is to help us understand and describe phenomena and to help us draw reliable conclusions about the phenomena which are not able to resolve through classical theories. With this in mind, this thesis applies the statistical approaches to random topology power grid modeling in order to improve the existing methods.

Synthetic power grid modeling is one of the most significant fields of studies. It should be noted here that the classic approaches of research towards this topic don't always end with a satisfactory output. Therefore, majority of researchers utilize the standard statistical methods to analyze and develop the synthetic power grid models. The idea is that by studying the statistical properties of real networks a method can be constructed that would be capable of generating fictional networks that have all the properties of a real power grid. This method can be used to generate approximate electric topology, to determine the correlated locations of generation, load and connection buses, and further to determine some electrical parameters such as generation/load settings and transmission line capacities.

The RT-nestedsmallworld model, to our best knowledge, is the most comprehensive and appropriate model in the literature to formulate a small-world connecting topology. This model could be used to produce any needed number of power grid test cases with scalable network size featuring the same kind of smallworld "*electrical*" topology of real-world power transmission networks. This model offers a helpful procedure to properly address the needs of power network researches. However, there are still a number of drawbacks within the current power grid modeling, such as bus type assignment, generation/load setting and transmission line capacity assignment.

This thesis presents a three-part statistical approach for improving the existing RT-nestedsmallworld model. The proposed improved model will provide a basis for developing optimal power flow competitions, further incentivizing future progress. This model promises to enable increased grid resilience, flexibility and improved energy efficiency while helping deliver the benefits of integrating renewable generation technologies into the electric power system .

1.5 Thesis overview and contributions

The main body of this thesis is divided into three chapters. In this section, we will summarize each, and highlight the contributions of this thesis.

1.5.1 Correlated bus type assignment in synthetic power grids

The first part of this thesis tries to solve the bus type assignment problem associated with current methodology of the RT-nestedsmallworld model. In this part, and for the first time we define a numerical measure, called the Bus Type Entropy, to characterize correlated bus type assignment of realistic power grids. This measure incorporates both bus type ratios and the link type ratios. Therefore it can effectively capture the correlated characteristics of realistic grids' bus type assignments. The proposed measure has an acceptable numerical stability since it follows more strict entropy definition, therefore it will simplify our analysis of the scaling property if the entropy value versus the network size. With the derived scaling function of correlated bus type assignment versus network size, a more efficient search algorithm based on clonal selection

procedure is developed to present more accurate bus type assignments of generation, load, and connection buses in existing synthetic power grid modeling.

1.5.2 Statistical settings of generation capacity, generation dispatch and load

Examining and interpreting structural setting of power systems enable the development of an appropriate synthetic modeling that could be utilized to produce power grid test cases with accurate grid topology parameters. However, power grid networks are much more than a graph topology and we need to consider realistic electrical parameter setting to fully represent a realistic power system. In recent years, various approaches have been proposed to model synthetic power grids in the literature. However, one of the most significant shortcoming of existing synthetic power grid models lies in the generation/load setting.

> *What is important here is that in synthetic power grid modeling it is impossible to apply the conventional methods into some electrical setting problems such as generation/load settings and line capacity assignment. To address this issue, we have to extract the statistical behavior of realistic power grid in the hope that these discoveries can be useful to design a practical methodology to solve the electrical setting problems.*

The thesis presents our recent study results on the statistics of generation/load capacities and settings in a synthetic grid modeling. A set of approaches has been developed to generate a statistically correct random set of generation capacities and assign them to the generation buses in a grid. Then we determine the generation dispatch of each generation unit according to its capacity

and the dispatch ratio statistics, which we collected and derived from a number of realistic grid test cases. The proposed approaches is readily applied to determining the load settings in a synthetic grid model and to studying the statistics of the flow distribution and to estimating the transmission constraint settings.

1.5.3 Transmission line capacity assignment

In the third part of this thesis, we investigate the electrical setting of power grids and the relationship between transmission line capacity and some network topology metrics. It is found that the capacity of transmission lines follows a well-known distribution and it can be fully defined by mathematical definition. The obtained distribution provides a pattern to generate reasonable capacity line values for a given synthetic power grid. Also, with respect to a new measure, called Neighboring Capacity Ratio (NCR), our experiments reveals some useful and interesting relations presented between the transmission capacity of a specific line and that of neighboring lines in the grid. Our statistical analysis on real-world power grid provides potential insights to propose a practical method to find the best transmission line capacity assignment in synthetic power grid modeling. Therefore, the presented statistical experiments can be used in the last part of this thesis to develop a new methodology to appropriately characterize the line capacity assignment and improve the RT-nestedsmallworld model.

Chapter 2

Correlated Bus Type Assignment in Synthetic Power Grids

2.1 Introduction

This section presents our study results on the correlated assignment of generation, load, and connection buses in a given grid topology and the development of an optimized search algorithm to improve the existing synthetic grid modeling. The previous works proposed a random-topology power grid model, called RT-nestedSmallWorld, which could be used to produce any needed number of power grid test cases with scalable network size featuring the same kind of small world "electrical" topology of real-world power transmission networks. This model offers a helpful procedure to properly address the needs of power network researches. However, there still are a number of drawbacks within the current power grid modeling, among which is the randomized assignment of generation (G), load (L), and connection (C) buses done purely according to a given set of bus type ratios. That is, in a typical power grid, 20-40% of the buses are generation buses, 40-60% load buses, and about 20% connection buses. Although a small number of buses may belong to more than one categories, this can usually be clarified by examining the net active power

injection at the bus location therefore it will not cause a big difference in the evaluation results. Reference [13] verified that there exists non-trivial correlation between the three bus types and other topology metrics such as node degrees and clustering coefficients in a real-world power grid. From [13], it is also found that random permutation of the original bus type assignment in a grid may cause big differences in system dynamics such as its vulnerability analysis to cascading failures. Utilizing a stochastic model of cascading failures proposed in [8], one may estimate the transition probabilities of a grid's connecting state from a stochastic model of flow redistribution and determine the time margin, called the expected safety time that is left to perform corrective action on each line.

Figure 2.1 plots the comparison results of the first 60 critical lines with shortest expected safety time. In this figure the blue solid line shows the safety analysis, presented above, on the original grid of IEEE 300 buses. The purple and green dashed lines plot the results obtained from two test cases of IEEE 300 bus system with random bus type assignment. And finally the red solid line with "x" is shows the average expected life time from the results of 10 random assignment cases. From Figure 2.1 and what presented in reference [13], experiments on the IEEE 300 bus system show that the expected safety time of the grid, after randomizing the grid bus type assignments, could result in a 150% increase compared with the realistic values. In other words, if we utilize random bus type assignment in a given random topology power grid model, i.e. RT-nestedsmallworld, although the topology of the generated test cases is consistent and comparable to that of a real-world grid, the inappropriate bus type assignment may still possibly cause deviation in the grid settings therefore give misleading results in the following evaluation and analysis.

The changes in bus type assignment in a grid may cause big differences in system dynamics such as the grid vulnerability to cascading failures. In other words, if we utilize random bus type assignment in the RT-nested smallworld model although the topology of the generated test cases is consistent and comparable to that of a real-world grid, the inappropriate bus type assignment may still possibly cause deviation in the grid settings therefore give misleading results in the following evaluation and analysis. Therefore, the random assignment of bus types in the RT-nested smallworld model should be improved by using a more accurate assignment which is consistent with that of realistic power grids.

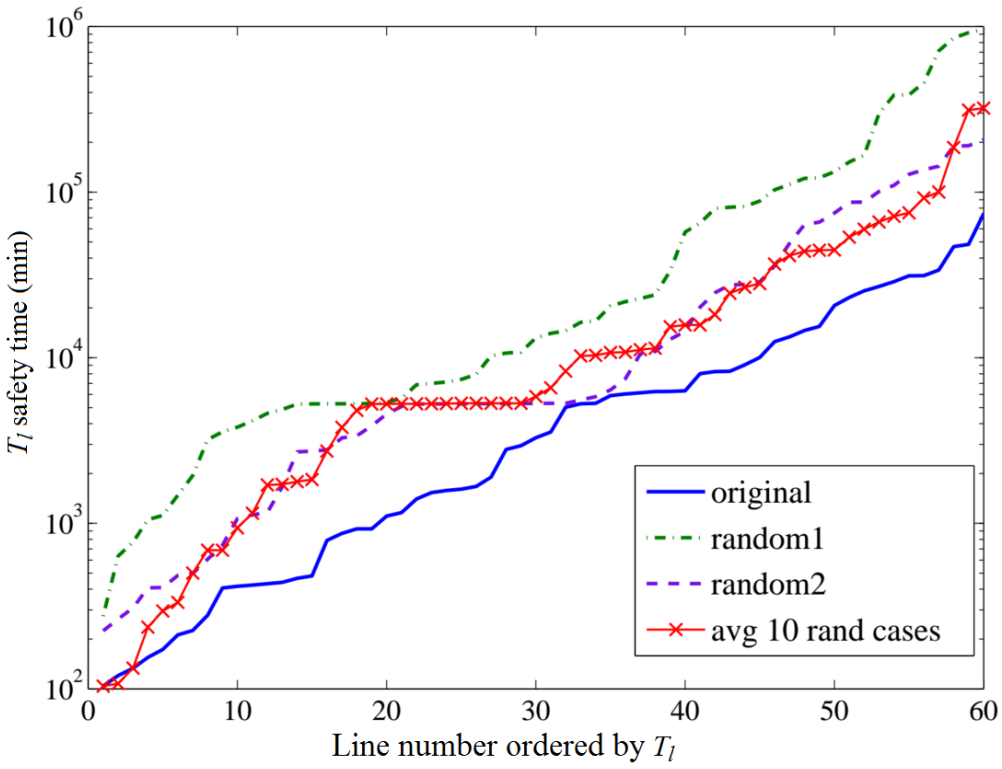


Figure 2.1: Safety time of grid lines of IEEE300 bus system: Original vs random bus type assignment [13]

2.2 Power system topology and definitions

Given a power grid topology with N buses and M branches, the setting of generation and load locations can be described by three indicator vectors: the Bus Type Vector ($\mathbb{T}$), Bus Type Ratio Vector (r), and Link Type Ratio Vector (R). The detailed discussions on these indicators are as follow:

Bus type vector: $\mathbb{T} = [\mathbb{T}_i]_{1\times N}$ where the type of bus i, $\mathbb{T}_i$= 1, 2 or 3 for a G, L or C bus, respectively.

Bus type ratio vector: $r = [r_k]_{1\times 3} = [r_1, r_2, r_3]$ where $r_1 = N_G/N$, $r_2 = N_L/N$ and $r_2 = N_C/N$ are the ratios of G, L and C buses in a grid, respectively, with NG, NL, NC representing the total number of G/L/C buses respectively.

Link type ratio vector: $R = [R_k]_{1\times 6} = [R_1, R_2, R_3, R_4, R_5, R_6]$ where R1−6 represent the ratios of the six types of branches in a grid, i.e. GG, LL, CC, GL, GC, LC. That is, the link type of a branch is determined by the bus types of its end buses. And $R_k = M_k/M$, for k = 1, · · · , 6 with Mk being the total number of branches of a specific link type in the grid.

Obviously the two vectors of bus type ratios and link type ratios can be fully determined by a bus type assignment vector T with respect to a given grid topology. In Figure 2.2, IEEE 30-bus system is considered as example to illustrate how the above-mentioned vectors can be dertmined in a given power grid. In this graphical representation of a grid network we find that the Bus and Link type ratio vectors are: r = [0.2000, 0.6000, 0.2000] and R = [0.0500, 0.1000, 0.1000, 0.4000, 0.2500, 0.1000] respectively. Table 2.1, presents the two indicator vectors for the IEEE test systems and some realistic power grids.

Table 2.1: Ratio of bus type and link types in real-world power networks

	Bus type ratio % G/L/C	**Link type ratio % GG/LL/CC/GL/GC/LC**
IEEE-30	20/60/20	5/40/10/10/10/25
IEEE-57	12/62/26	6/34/6/16/6/32
IEEE-118	46/46/8	27/14/2/43/11/3
IEEE-300	23/55/22	2/36/11/22/7/22
NYISO-2935	33/44/23	16/28/8/29/6/13
ERCOT-5633	9/54/37	1/27/26/2/6/38
WECC-16994	20/40/40	1/13/40/2/7/37

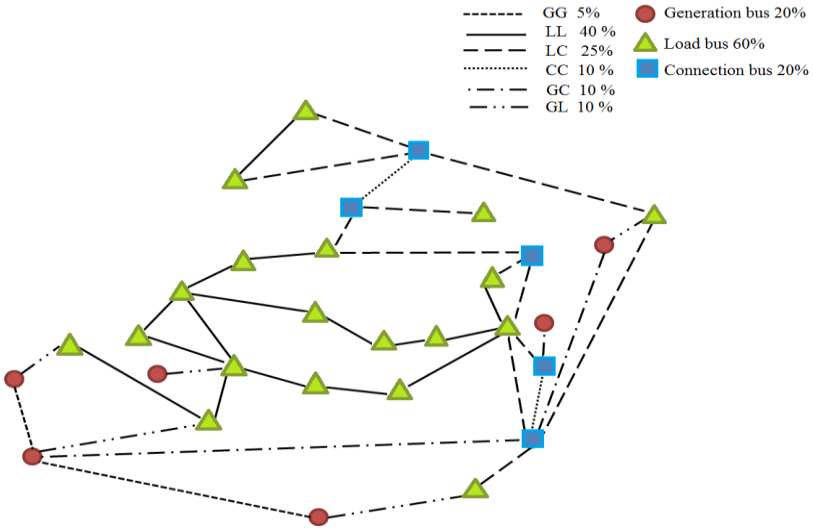


Figure 2.2: Graphical representation of the IEEE-30 bus system.

2.3 The Bus Type Entropy

As mentioned in [13], there exists non-trivial correlation between the three bus types and other topology metrics such as node degrees and clustering coefficients in a real-world power grid; and a random permutation of the grid's original bus type assignments, i.e. to make it a random one, will cause significant deviation in the grid dynamics and makes the resulting network behave no longer like a power grid.

This fact has led to define a measure, which should be simple and convenient, and at the same time able to effectively quantify and characterize the "correlated" bus type assignments in realistic power grids, i.e., distinguishing them from the randomized ones. The proposed measure is called Bus Type Entropy, with three various definitions. The statistical analysis based on empirical probability density function (PDF) based on the IEEE 300-bus system and the NYISO 2935-bus system [43] has roughly verified the performance of the proposed entropy measures.

After obtaining additional data of some realistic grids through the support from the CERTS initiative, we find it would be ideal to derive a scaling property of the correlated bus type assignment versus the network size of a grid. Therefore, the optimization algorithm to search for the best correlated bus type assignment for a given random topology power grid model will be enhanced with a direct target entropy value of the correlated bus type assignments determined by the given network size, saving the mandatory, but most the time unattainable, requirement of a set of realistic grid data with a comparable network size for identifying the search target.

In order to capture a good scaling function of the bus type entropy we examine our proposed entropy definitions using the IEEE 30, 57, 118, 300-bus systems and the NYISO- 2935, the ERCOT-5633 [44], and the WECC-16994 [45] systems. We extract an empirical probability density functions (PDF) of the bus type entropy value from randomized bus type assignments over the grid original topology, estimate the distribution parameters, measured the relative location of a target entropy value within the PDF curve, and teste different curve fitting approaches to detect an appropriate scaling function. Therefore, in the first step we propose the definition of the bus type entropy, with two variations, each with its own special advantages:

$$W_0(\mathbb{T}) = -\sum_{k=1}^{3} r_k \times \log(r_k) - \sum_{k=1}^{6} R_k \times \log(R_k) \quad (2.1)$$

$$W_1(\mathbb{T}) = -\sum_{k=1}^{3} log(r_k) \times N_k - \sum_{k=1}^{6} log(R_k) \times M_k \quad (2.2)$$

Obviously, (2.1) is a typical entropy definition of statistical variables, based on which the derived entropy values fall within a very stable or restricted numerical region, as shown with the results in section 2.4.1. This property provides the advantage to simplify our design of the optimization procedure seeking for the best bus type assignments. While the second definition (2.2) can be viewed as a more "generalized" entropy, which tends to "magnify" the scaling impact of the entropy value versus the grid network size. Therefore it has the advantage to simplify the approximation procedure of the scaling function. We choose above entropy definitions because they are simple and convenient to evaluate and ready for empirical PDF analysis. And most importantly, the definitions incorporate both bus and link type ratios in a grid, with the grid topology information embedded, therefore they are able to recognize the "correlated" bus type assignments consistent with those in realistic power grids.

2.4 Statistical analysis

In this section, we conduct statistical analysis and investigate the effectiveness of the proposed bus type entropy in realistic power grid systems. We aim at assessing the relative difference or distance between the original bus type assignment $\mathbb{T}^*$in a realistic power grid and other randomized bus type assignments $\widetilde{\mathbb{T}} = \mathrm{P}(\mathbb{T}^*)$ with the latter obtained from random permutation of the former.

In order to investigate the benefits offered by (2.1) and (2.2), we will randomly permute the location of G/L/C buses in given realistic grids, and collect the resulting entropy values$W_{0-1}(\widetilde{\mathbb{T}})$, which are computed using corresponding link and bus type ratios. Although the permutation

process does not change the bus type ratios in a grid, that is, the number of G/L/C buses is maintained, we will have different link type ratio vectors (R) with respect to the random location of generation and load buses. Then the empirical probability density function (PDF) will be formulated based on the collected random samples of $W_{0-1}(\widetilde{\mathbb{T}})$and utilized to investigate the relationship between $\mathbb{T}^*$ and $\widetilde{\mathbb{T}}$'s.

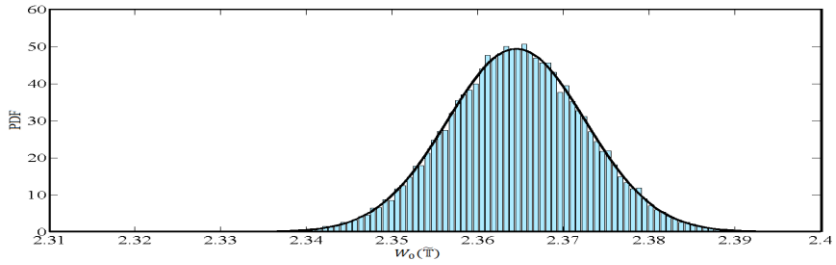
It is worth noting that by the Central Limit Theorem (CLT), the empirical PDF of $W(\widetilde{\mathbb{T}})$converges to a normal or Gaussian distribution. And this can be verified by the results in the following session of this thesis. Then the estimate distribution parameters of (μ, σ) will be extracted to measure the relative location of $\mathbb{T}^*$ among the randomized $\widetilde{\mathbb{T}}$'s based on their corresponding characteristic entropy values.

The quality of generated PDF directly depends on the size of sample set. If the sample size is sufficiently large, the following empirical PDF will have appropriate statistical accuracy. Theoretically, in our permutation process the total number of all possible samples of $W_{0-1}(\widetilde{\mathbb{T}})$ will be no greater than $\hat{n} = \frac{N!}{N_G!N_L!N_C!}$where N_G, N_L, and N_C represent the total number of generation, load and connection buses in a grid, respectively.

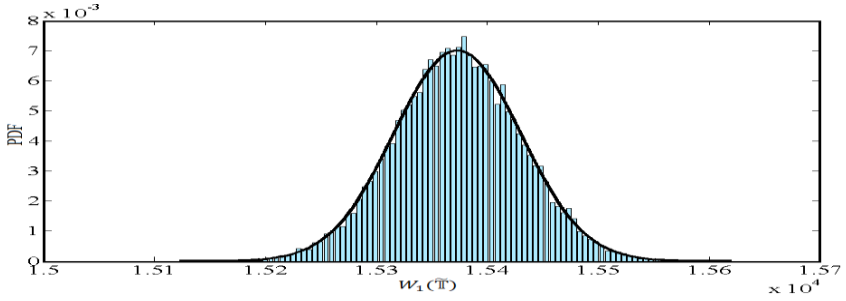
In our study an appropriate sampling size k^{max} ($k^{max} < \hat{n}$) is considered as that can achieve, at a reasonable computational cost, the required statistical accuracy which means a recognizable normal distribution with feasible estimates of distribution parameters. For example, our experiments on realistic power grids show that for a test case with a network size (i.e. the total number of buses) of $N \leq 4000$, we may consider $k^{max} = 25,000$. For larger case studies the sampling size will be increased to $k^{max} = 40,000$. The difference between two proposed numbers comes from the fact that increasing in the number of buses (network size) causes an extraordinary

growth in the number of possible scenarios. The important thing to note here is that our goal is not to utilize the complex mathematical sampling-based methods to determine the optimal sample size. Instead, we want to shift focus to the fact that in order to have a desired quality for empirical PDF we do not need to consider the entire sample set. Because considering $\hat{n}$=k^{max} dramatically increases the computation cost without a tangible change in the statistical accuracy [46]-[48].

Figure 2.3 displays the empirical PDF that results from the numerical simulation of the ERCOT system with respect to the entropy definitions of $W_0(\mathbb{T})$ and $W_1(\mathbb{T})$. These empirical PDF curves (denoted as bar plots) are generated with the sample size of k ^{max} = 40, 000 and include a fitting Normal Distribution (black solid line). The comparison of the generated empirical PDF with the normal distribution shows that the deviation is very small.



(a)



(b)

Figure 2.3 Empirical PDF of the bus type entropies and the fitting normal distribution evaluated in the ERCOT system: (a) $\boldsymbol{W_0}(\widetilde{\mathbb{T}})$, and (b) $\boldsymbol{W_1}(\widetilde{\mathbb{T}})$.

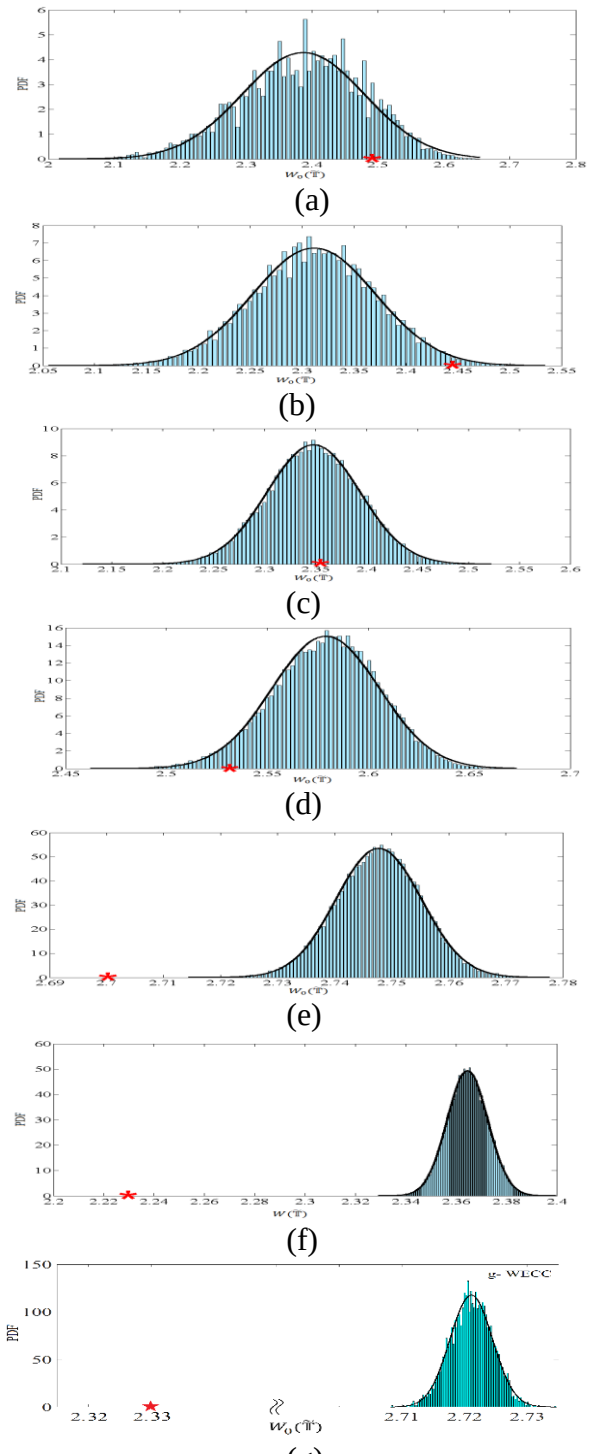


Figure 2.4: The Empirical PDF of the bus type entropies of $\boldsymbol{W_0}(\widetilde{\mathbb{T}})$ and the fitting normal distribution: (a) IEEE-30, (b) IEEE-57, (c)IEEE-118, (d)IEEE- 300, (e)NYISO, (f)ERCOT, and (g)WECC. In each sub-figure the realistic bus type entropy $W^* = W_0(\mathbb{T}^*)$. is marked by a red 'star'

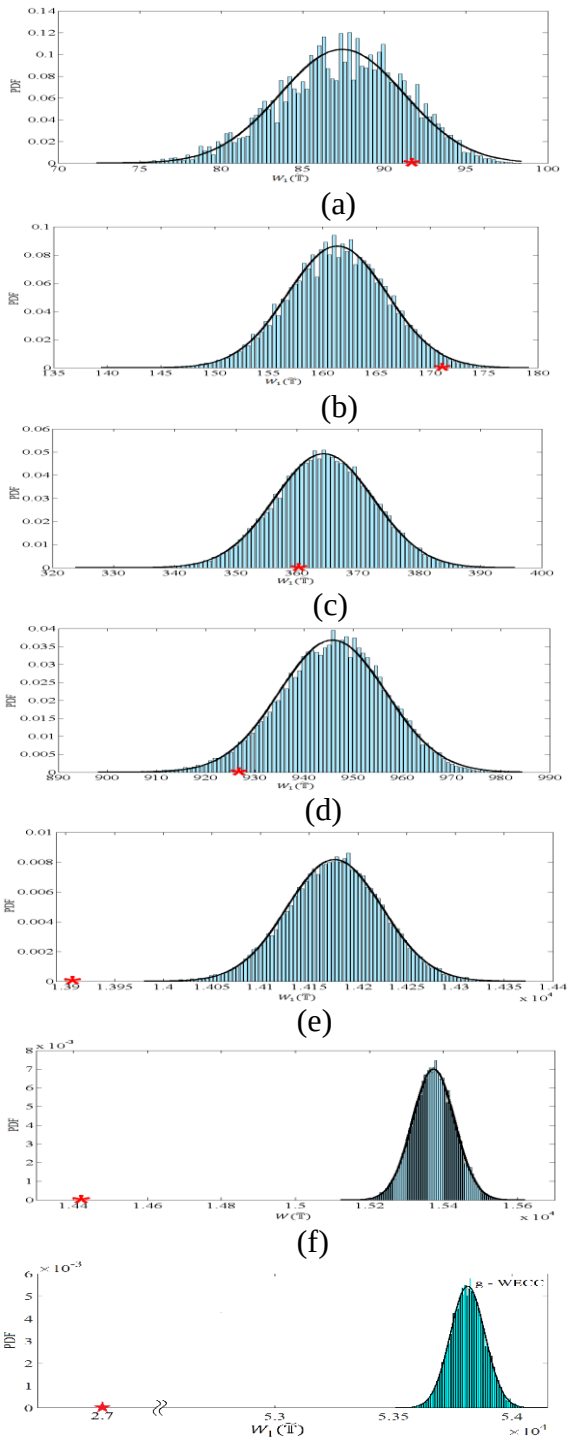


Figure 2.5: The Empirical PDF of the bus type entropies of $\mathbf{1}(\widetilde{\mathbb{T}})$ and the fitting normal distribution: (a) IEEE-30, (b) IEEE-57, (c)IEEE-118, (d)IEEE- 300, (e)NYISO, (f)ERCOT, and (g)WECC. In each sub-figure the realistic bus type entropy $W^* = W_1(\mathbb{T}^*)$. is marked by a red 'star'.

Figure 2.4 and 2.5 provide the numerical results obtained from all the IEEE test cases and the available realistic power grids. With the newly defined entropy measures and the chosen sample sizes, all the empirical PDF converge to a normal distribution. In these figures the fitting normal distribution curve is shown as a thick black line and the original bus type entropy W^* is marked in each empirical PDF as a red star.

As mentioned above, given a realistic power grid topology, $\mathbb{T}^*$ and $W^* = W(\mathbb{T}^*)$ represent the original bus type assignment and the corresponding bus type entropy, respectively. What is important to us is the location of W^* versus the empirical PDF in realistic power grids, because this may help us to identify the searching scope of the best set of target bus type assignments for the random topology power grids. Table 2.2 summarizes the fitting parameters, i.e., the mean value (μ) and the standard deviation (σ), for each generated empirical PDF, and the original bus type entropy W^* , for all the IEEE test systems and the available realistic grid data.

Table 2.2: The fitting parameters of normal distribution and the original bus type entropy values

	$W_0(\mathbb{T})$ $\mu/\sigma/W_0^*$	$W_1(\mathbb{T})$ $\mu/\sigma/W_1^*$
IEEE-30	2.38/0.09/2.49	87.02/3.8/92
IEEE-57	2.31/0.058/2.44	161.4/4.63/172
IEEE-118	2.34/0.045/2.35	346.39/8.13/365.04
IEEE-300	2.57/0.026/2.53	943.21/10.58/927.5
NYISO-2935	2.74/0.007/2.70	14193/48.8/13910
ERCOT-5633	2.36/0.008/2.23	15372/56.47/14428
WECC-16994	2.72/0.0034/2.33	53813/74.33/27775

2.4.1 The Scaling Property and Guidance map

In this section we examine the relative location of the original bus type entropy W^* in the empirical PDF of randomized bus type entropy in realistic power grids. What we are looking for

is to capture a scaling property of W^* versus the network size with the help of the fitting parameters of (σ, μ). Such a possible relationship can be utilized to improve our synthetic power grid modeling with the goal of estimating the measure of W^* with respect to a related empirical PDF.

To do so, given a generated random topology power grid, we will first extract the fitting parameters (σ, μ) from the empirical normal PDF of randomized bus type assignments, then use them and the observed function of the scaling property to determine the target entropy value of a correlated bus type assignment W^*, which is consistent with that observed in realistic grids. Therefore an optimization algorithm can be implemented to search for the desired bus type assignments with respect to the target entropy value W^*.

Figure 2.4 and 2.5 indicate the location of original bus type entropy W^* within the corresponding empirical PDFs. Careful examination of these two figures reveals that the location of W^* (red star) is not stationary but there is a tangible trend for the distance between W^* and μ from higher to lower values with respect to network size N, i.e. it moves from right to the left side as the network size gets larger. A guidance map in Figure 2.6 illustrates the trend for both entropy definitions with IEEE 57-bus system being the only exception with some negligible deviation. This behavior indicates a possible existence of scaling property of the normalized distance called d, versus the network size N, which is defined as:

$$d = \frac{W^* - \mu}{\sigma} \tag{2.3}$$

The idea behind the scaling property of d versus N is to identify the behavior of realistic bus type entropy W^* in terms of position and probability, taking into account network information and topology, with

$$W^* = \mu + \sigma d_W(N) \tag{2.4}$$

Further, *d* specifies the relationship between W^* and other random bus type entropies $W(\widetilde{\mathbb{T}})$. It should be noted that the scaling property *d* is a numerical value and can be considered as a topological characteristic for power network topologies.

Table 2.3 shows the value of scaling property d for realistic power grids. The obtained results indicate that with increasing in network size, the values of $\boldsymbol{d_{W_0(\mathbb{T})}}$and $\boldsymbol{d_{W_1(\mathbb{T})}}$ have the tendency to decrease. One way to find the mathematical relationship is the curve fitting, which defines an appropriate curve to fit the observed values and uses a curve function to analyze the relationship between the variables. Suppose that from the mentioned statistical analysis, presented in Table 2.2, observations from the seven IEEE and realistic grids have been collected.

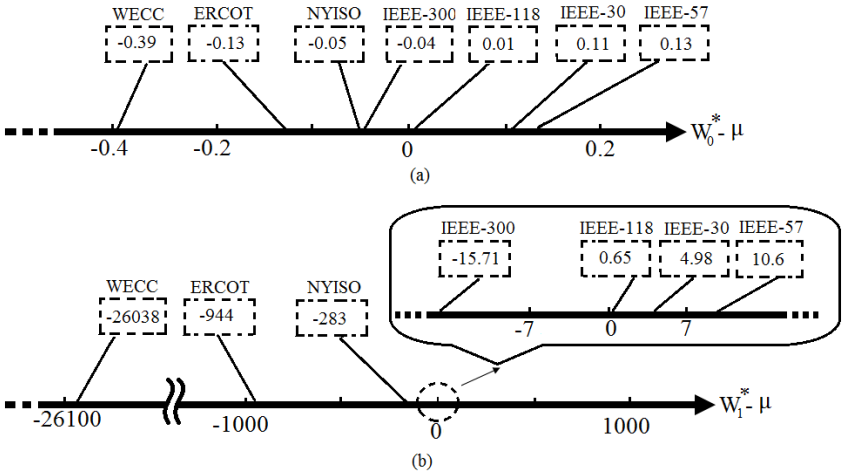


Figure 2.6: The comparison of W^*- μ for different realistic power systems based on (a) $\boldsymbol{W_0}(\mathbb{T})$ and (b) $\boldsymbol{W_1}(\mathbb{T})$

The first step towards the construction of a mathematical model is to plot these data points and estimate a function form *f(N)* to describe the general trend in the data. From Table 6 and based on the data distribution shape, a piecewise function including a linear part $f(x) = a_1x + b_1$ and

nonlinear part $f(x) = a_2 x^{b_2} + c_2$, can be considered as an appropriate model to characterize the relationship between scaling property-d and *N*. In this function $\{a_1, b_1\}$ and $\{a_2, b_2, c_2\}$ are parameters that we can adjust. In our fitting process we are trying to choose both linear and nonlinear function parameters so as to minimize the fitting error.

Table 2.3: The scaling property-d in realistic power grid systems

	N	**M**	$\boldsymbol{d_{W_0}(\mathbb{T})}$	$\boldsymbol{d_{W_1}(\mathbb{T})}$
IEEE-30	30	41	1.22	1.31
IEEE-57	57	78	2.24	2.28
IEEE-118	118	179	0.22	0.08
IEEE-300	300	409	-1.53	-1.48
NYISO-2935	2935	6567	-5.71	-5.79
ERCOT-5633	5633	7053	-16.25	-16.71
WECC-16994	16994	21539	-114.7	-350.3

Figure 2.7 shows the fitted piecewise functions for $\boldsymbol{d_{W_0}(\mathbb{T})}$ and $\boldsymbol{d_{W_1}(\mathbb{T})}$ with corresponding root-mean-square deviation (RMSE) of 1.81 and 15.31, respectively. The mathematical definition of obtained fitted curves are presented as follows:

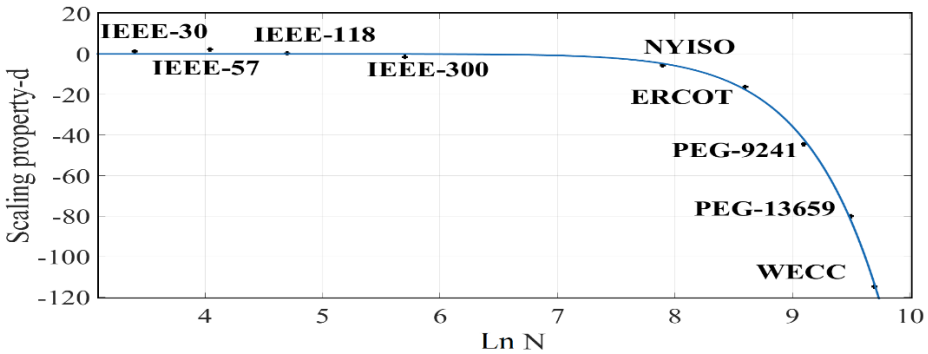
$$d_{W_0}(N) = \begin{cases} -1.721\, LnN + 8 & \text{Ln } N \leq 8 \\ -6.003 \times 10^{-14} (\text{Ln } N)^{15.48} & \text{Ln } N > 8 \end{cases} \quad (2.5)$$

$$d_{W_0}(N) = \begin{cases} -1.748\, LnN + 8.276 & \text{Ln } N \leq 8 \\ -6.053 \times 10^{-22} (\text{Ln } N)^{24.1} & \text{Ln } N > 8 \end{cases} \quad (2.6)$$

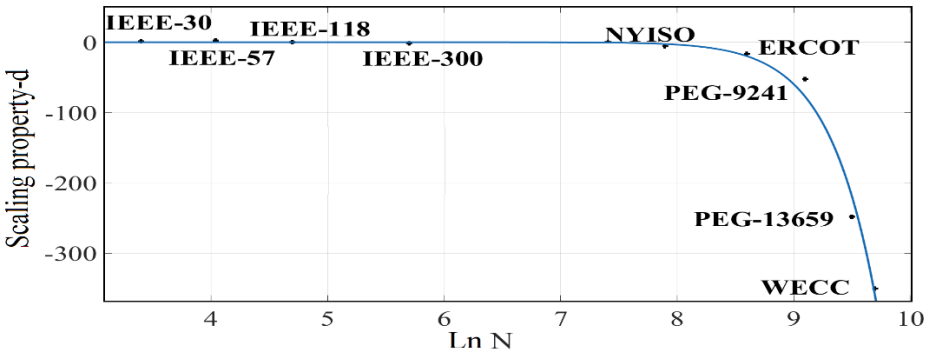
For the proposed scaling property, a linear approximation can be considered for the small size networks by ignoring the large-scale ones. However, our initial results show that as the network size increases, the values of $\boldsymbol{d_{W_0}(\mathbb{T})}$ and $\boldsymbol{d_{W_1}(\mathbb{T})}$ dramatically decrease. Accordingly, a nonlinear definition can be considered as a better approximation for the large-scale networks. Tt should be noted that in order to obtain better precision for a wide range of network sizes, Ln *N* is used instead of *N* in equations (2.5) and (2.6).

The fitted curves provide useful information to solve the bus type assignment problem. Given a random topology power grid with size N, one can easily find the value of d and further evaluate the target value W^* using (2.4). It should be noted that in equation (2.4) the distribution parameters (μ, σ) are extracted from the empirical PDF of randomized bus type assignments in the given grid topology.

Both the presented scaling functions of (2.5) and (2.6) will do a good job in the bus type assignment problem. However, we would recommend using the second one. The main reason comes from the fact that the second guidance map, with higher resolution, indirectly provides a better condition for optimization phase to find the best bus type assignment. Indeed, in optimization process we are looking for unique bus type assignment with an special bus type entropy (say W^*). This procedure leads to difficulties when we know that several different bus type assignments may have the same bus type entropy. With this in mind, from a detailed observation of proposed bus type entropy definition (2.1) and (2.2), it is clear that definition (2.6) provides a wider bandwidth for empirical PDFs (see Figure 2.7). Obviously, a wider PDF will potentially reduce the number of nominated bus type assignments for each possible bus type entropy.



(a)



(b)

Figure 2.7: The fitting curve of the scaling property of the normalized distance versus network size: (a) $\boldsymbol{d_{W_0(\mathbb{T})}}$ and (b) $\boldsymbol{d_{W_1(\mathbb{T})}}$.

The proposed synthetic grid modeling can then be improved with an optimized search algorithm as the following step-by step procedure. Given a random topology power grid generated from our proposed modeling, RT nestedSmallWorld:

Step 1 : Generate the empirical PDF curve of randomized bus type assignments $W(\widetilde{\mathbb{T}})$. with respect to the grid size data *N* and *M*, and it connecting topology. Upon the completion of step 1, the fitting parameters of μ and σ will be calculated.

Step 2 : Estimate the value of scaling property d using proposed guidance map and with respect to logrithm of network size N.

Step 3 : Find the value of W^*using(2.4) and the obtained parameters from step 1 and 2.

Step 4 : Determine the best set of bus type assignments with the target entropy value of W^* provided by step 3.

In step 4 we need to design an optimization process to search for those bus type assignments which have nearly the same bus type entropy as W^*. In the optimization process, the objective function can be defined as:

$$min_{\mathbb{T}} \ \varepsilon = |W(\mathbb{T}) - W^*| \tag{2.7}$$

The pseudo code and procedures for the proposed method can be described as in Figure 2.8. The steps in Figure 2.8 are shown in the flow chart in Figure 2.9 for completeness.

Algorithm 1: Proposed statistical methodology

Start
For a given random topology
1. Enter N, M and Link connection Matrix
2. Find the empirical PDF
3. Save the setting parameters μ and σ (from step 2)
4. Estimate the value of scaling property-d (Based on N and proposed guidance map)
5. Save the value of scaling property-d
6. Calculate W* for given topology (using (5) and with respect to d, μ and σ)
7. Optimization process
   While $|W(\tilde{\mathbb{T}}) - W^*| > \varepsilon$
   Generate random bus type assignment $\tilde{\mathbb{T}}$
   Find $W(\tilde{\mathbb{T}})$
   Calculate $|W(\tilde{\mathbb{T}}) - W^*|$
8. If $|W(\tilde{\mathbb{T}}) - W^*| < \varepsilon$ then $\tilde{\mathbb{T}}$= best bus type assignment

End

Figure 2.8: The pseudo code of the proposed method

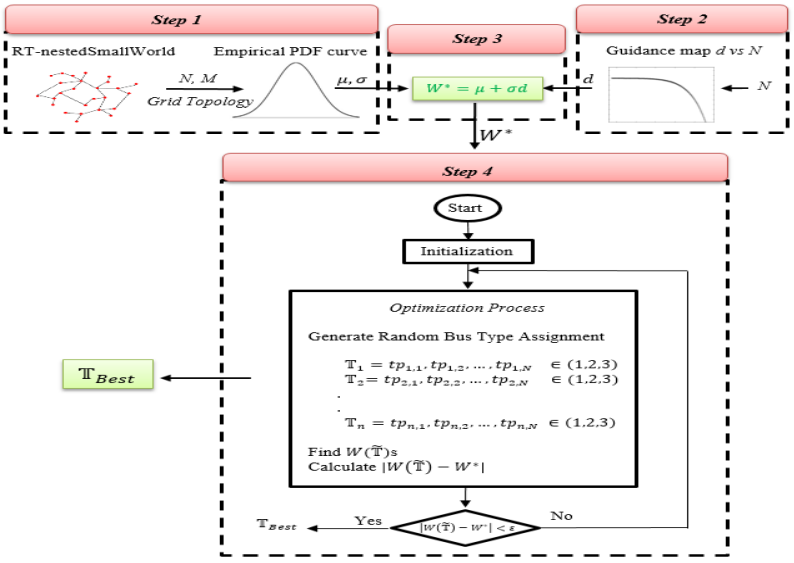


Figure 2.9: Flow diagram for the proposed methodology

2.5 Optimization Search Algorithm

Based on what discussed above, sections 2.4 and 2.5 provide a foundation for the last part of the proposed method. In optimization phase Clonal Selection Algorithm (CSA) that results in optimal or the best set of bus type assignments is presented. CSA is an efficient optimization method, inspired by the biological immunity system selection mechanism, proposed by De Castro and Van Zuben [49]. This method has successfully been applied to optimization domain and power system area in recent years. The performance of CSA can be summarized as the following procedure [49-52]:

Step 1: Randomly produce the initial population of CSA. Each individual of the population is a candidate bus type assignment for the optimization problem. Here, for a N-bus power grid we consider the bus type vector $\mathbb{T} = [\mathbb{T}_i]_{1\times N}$ as a candidate solution, with $\mathbb{T}_i = 1$ for generating buses; 2 for load buses and 3 for connection buses. The total number of candidate solutions in the population is denoted by *P*. In general, a random bus type assignment in iteration t can be written as follows:

$$\mathbb{T}_i^t = \left[\mathbb{T}_{i,1}^t, \mathbb{T}_{i,2}^t, .., \mathbb{T}_{i,k}^t, \dots, \mathbb{T}_{i,N}^t\right], \qquad i = 1 = 1,2,3, \dots, P \tag{2.8}$$

The three sets of bus types of each individual are randomly initialized with uniform distribution according a specified set of bus type ratios, $\{r_G, r_L, r_C\}$.

Step 2: Determine the value of bus type entropy $W(\mathbb{T})$, and compute the "affinity" value, with respect to each single candidate solution. In the following optimization tasks, the value of affinity corresponds to the evaluation of the objective function for the given candidate solution [49].

Step 3: Sort the random bus type assignments based on the output affinity values.

Step 4: Select the first "n" bus type assignments based on their ranking position in the sorted population and make copies of offspring candidate for producing next generation:

$$nc_i = \left[\frac{\beta N}{i}\right], \qquad \forall i = 1,2,\dots,n \tag{2.9}$$

Where [.] means the rounding function to the nearest integer toward infinity, $nc_i$is the number of offspring candidates from i^{th} candidate, and β is a constant coefficient which indicates the rate of copy. Obviously, a candidate with higher ranking will be copied more than a candidate with lower ranking. Upon the completion of step 4, the total number of possible solutions in the next-generation population will be "*NC*" as follows:

$$NC = \sum_{i=1}^{n} \left[\frac{\beta N}{i}\right] \tag{2.10}$$

Figure 2.10 graphically illustrates step 4. As can be seen, nc_1 and nc_2 present the numbers of offspring copied from the highest and second ones, respectively. The rest of selected candidates will then be multiplied sequentially according to (2.9).

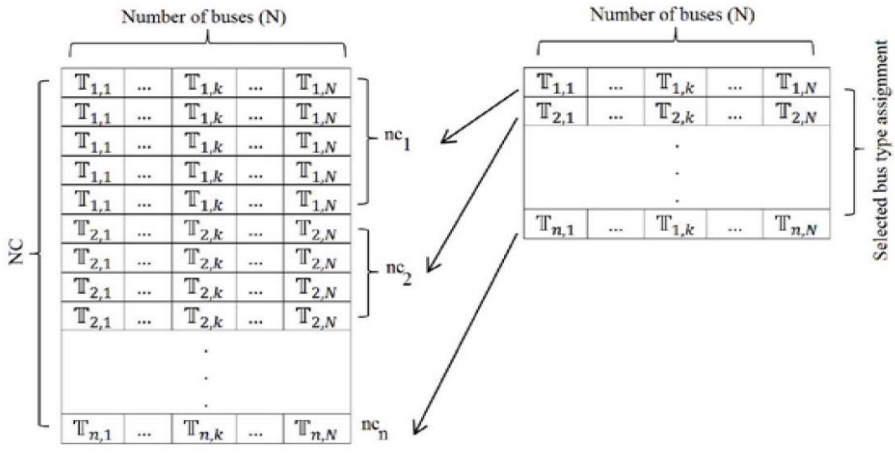


Figure 2.10: Copy operator in the CSA algorithm

Step 5: Mutate the *NC* bus type assignments in the population based on their bus type entropy values. It means that the candidate solutions with higher affinity should be mutated less than those with lower affinity.

Step 6: Determine the affinity of each mutated candidate and select the first "m" ($< P$) candidates with the highest affinity values among the NC mutated and original population. The selected m bus type assignments will enter the next generation directly.

Step 7: Generate $p = P - m$ new bus type assignment as the "antibodies" for the next generation through a random process. These randomly generated antibodies enhance search diversity of CSA, and consequently, the algorithm takes the chance to escape from the local optima.

Step 8: Return back to step 2 and repeat this cycle until the termination criteria are met. If the termination criterion satisfied, the best set of bus type assignments of the last generation is determined as the optimum solution of optimization process. Figure 2.11 demonstrates the proposed optimization procedure.

In order to show the efficiency of the proposed optimization algorithm we apply it to two power systems, i.e., the IEEE-300 bus system. To clarify; we use this system as a realistic grid sample to generate a set of correlated bus type assignment for the random topology power grids of similar network sizes using the objective function presented by (2.7).

It should be noted that the definition of objective function depends on our knowledge from the topology of power grid. The performance of the CSA for IEEE-300 bus system is shown in Figure 2.12 and 2.13. In Fig.7 the stopping criteria for iterative optimization process is considered as (Error < 0.003).

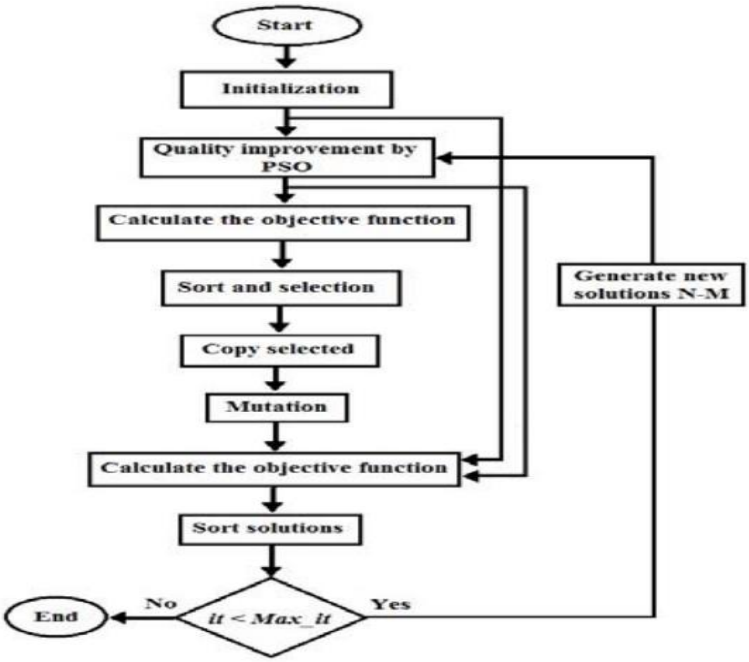
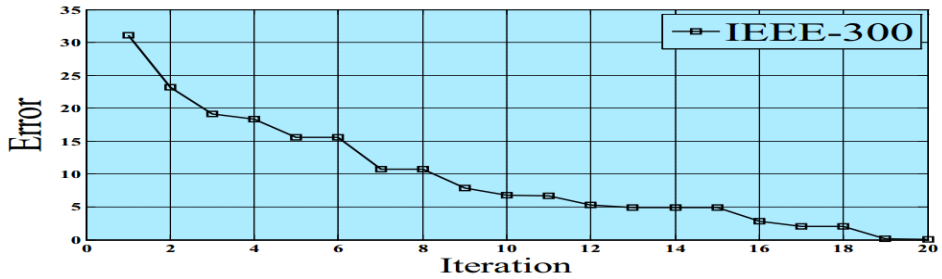
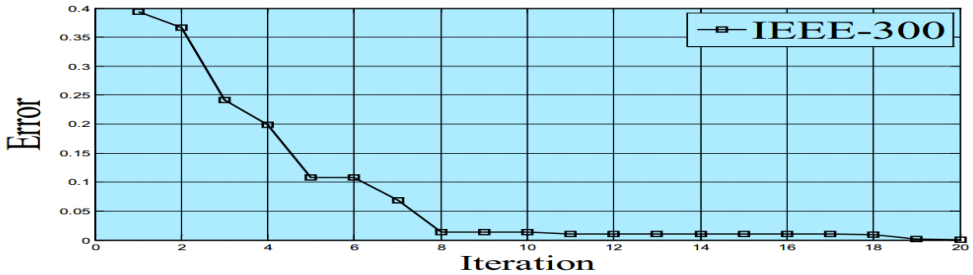


Figure 2.11: Flowchart of the clonal selection algorithm (CSA)



(a) $W_0(\mathbb{T})$



(b) $W_1(\mathbb{T})$

Figure 2.12: Convergence process of CSA algorithm in the IEEE-300 systems with the three entropy definitions $W_{0-1}(\mathbb{T})$

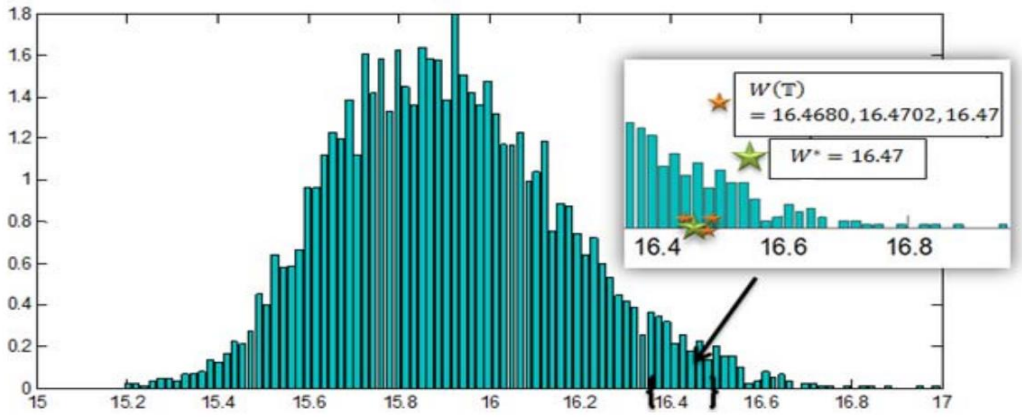


Figure 2.13: The best set of bus type assignments for a 300 bus system using the entropy definition of $W_1(\mathbb{T})$

2.7 Numerical results

In order to show the efficiency of the proposed algorithm, we performed a series of simulation and applied the proposed method to a set of random topology power grids such as RT-nestedSmallWorld test cases with 129, 147, 300 and 3000 buses. Technically, these application examples help us to show how the suggested methodology works. Furthermore, to verify the effectiveness of the proposed approach we tried to solve the bus type assignment problem for the available realistic power grids, and then checked the deviation of the searched bus type assignments from the original ones. In this way, we can evaluate the ability of proposed method to solve the bus type assignment problem.

Table 2.4 provides the topology parameters for the four random-topology case studies. It should be noted that the total number of branches and the bus type ratios are chosen according to a realistic network with most similar network-size (see Table.2.1).

Figure 2.14 and 2.15 display the empirical PDFs that result from the numerical simulations of the RT-129, 147,300 and 3000 buses systems with respect to the entropy definitions of

$W_0(\mathbb{T})$and $W_1(\mathbb{T})$. Table 2.5 shows the fitting distribution parameters (σ, μ) of the empirical PDF of randomized bus type entropy, the estimated normalized distance d evaluated based on the scaling function, and the derived target entropy values W^*, according to the two entropy definitions.

It should be noted that in a large scale power grids, different bus type assignments may have the same bus type entropy. Define the difference measure between two bus type assignment vectors $\mathbb{T}^i$ and $\mathbb{T}^j$ in a N-bus grid as:

$$\Delta\mathbb{T}^{ij} = \Delta\left(\mathbb{T}^i, \mathbb{T}^j\right) = 1 - \frac{\sum_{n=1}^{N} \delta(\mathbb{T}_n^i - \mathbb{T}_n^j)}{N} \tag{2.11}$$

where δ(·) is the Dirac delta function. And the diversity of a set of n_t bus type assignment vectors (𝕋)'s can then be defined as:

$$max_{i,j\in[1:n_t]} \quad \Delta\mathbb{T}^{ij} \tag{2.12}$$

However, our experiments on realistic power grids show that the diversity is not considerable. For the above random topology power grid cases, we have the diversity of located bus type assignments to be ≤ 10%. Therefore we accept all solutions for the bus type assignment problem. In the RTnested-smallworld-300 case, for example, the obtained results show that the output of optimization process includes 12 different bus type assignments, and diversity is 7.8%.

Table 2.4: Topology parameters of random topology case studies

Random topology Network	**N**	**M**	**Bus Type Ratio % G/L/C**
RT-129	129	189	46/46/8
RT-147	147	222	46/46/8
RT-300	300	456	23/55/22
RT-3000	3000	7119	33/44/23

We also apply the proposed algorithm to the realistic power grids. Table 2.6 provides the simulation results: the second and third columns shows the estimated bus type entropies W_0^* and W_1^* . With respect to the estimated target entropy values and using the proposed optimization algorithm we will be able to locate a set of the best bus type assignments $\mathbb{T}^{\text{best}}$ for each realistic power grid. In Table 2.6, the last column presents the maximum difference

$$\Delta\mathbb{T}^{best,*} = max_{\mathbb{T}^k \in \mathbb{T}^{best}} \ \Delta(\mathbb{T}^k, \mathbb{T}^*) \tag{2.13}$$

between the located bus type assignments and the original one for each real world power grid. Although the obtained results show a growing error with respect to the network size N, we can see a promising similarity, even for a large power grids.

Table 2.5: Scaling function results for RT-Nestedsmalworld networks

Network	d_{W_0}	μ	σ	W_0^*
RT-129	0.05	2.35	0.03	2.39
RT-147	-0.14	2.32	0.03	2.31
RT-300	-1.13	2.58	0.02	2.55
RT-3000	-4.97	2.74	0.005	2.71
Network	d_{W_1}	μ	σ	W_1^*
RT-129	0.02	390	7.53	390.15
RT-147	-0.18	447	8.09	445.75
RT-300	-1.2	1021	10.77	1008
RT-3000	-2.9	15175	40.37	15057

Table 2.6: Estimated bus type entropy for realistic power grid systems

	Estimated W_0^*	**Estimated** W_1^*	$\Delta\mathbb{T}^{best,*}$
IEEE-30	2.48	93	7.6 %
IEEE-57	2.46	170	7.2 %
IEEE-118	2.33	360	8.4 %
IEEE-300	2.55	922	10 %
NYISO-2935	2.73	13905	13.3 %
ERCOT-5633	2.24	14430	14.7 %
WECC-16994	2.31	27780	19.6 %

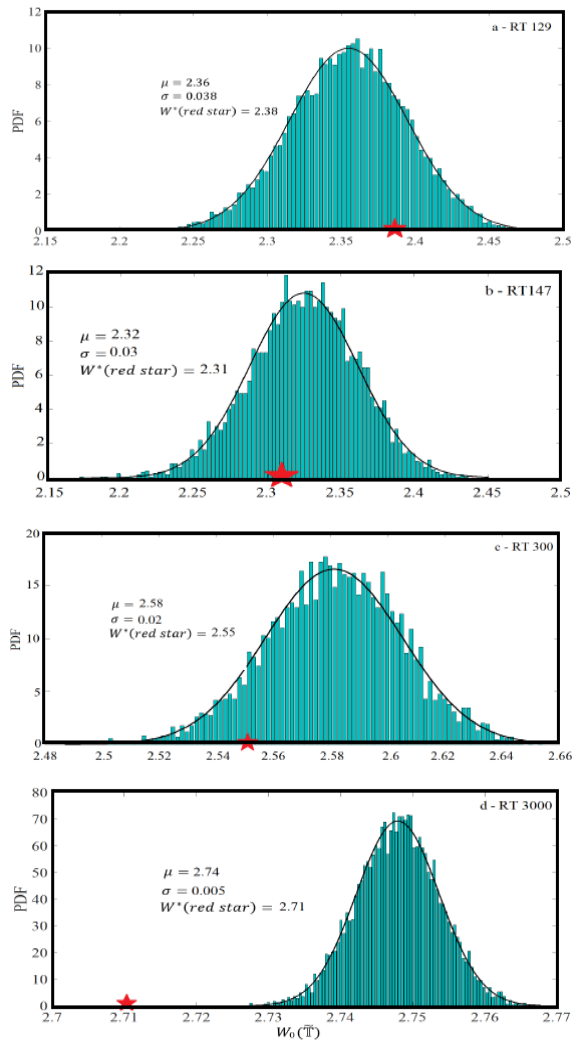


Figure 2.14: Empirical PDF and the normal distribution fitting for the randomized bus type entropy $\boldsymbol{W_0}(\widetilde{\mathbb{T}})$: (a) RT-129, (b) RT-147, (c) RT-300, (d) RT-3000

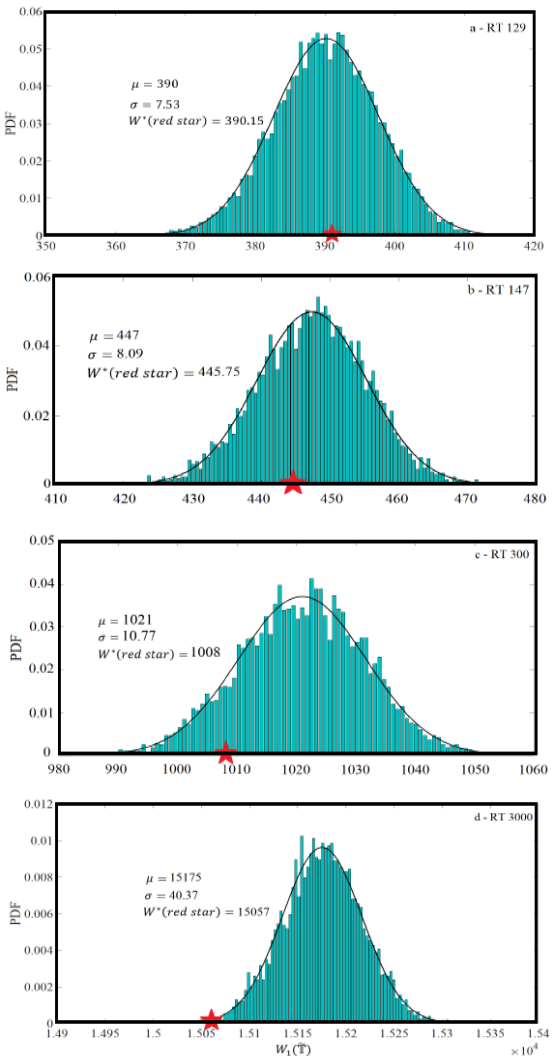


Figure 2.15: Empirical PDF and the normal distribution fitting for the randomized bus type entropy $\boldsymbol{W_1}(\widetilde{\mathbb{T}})$: (a) RT-129, (b) RT-147, (c) RT-300, (d) RT-3000

2.8 Conclusion

In this section we propose a statistical methodology to solve the bus type assignment problem in synthetic power grid models. This method includes a novel measure, called the Bus Type Entropy, the derivation of scaling property, and the optimized search algorithm. In this chapter we first define a numerical measure, called the Bus Type Entropy, to characterize correlated bus type assignment of realistic power grids. We then derive a mathematical approximation function, which

is able to indicate the appropriate entropy value that a correlated bus type assignment should assume in a specific N-bus grid network. Based on above results, the target entropy value for the best bus type assignments in a power grid topology can be directly selected given its network size. Therefore the synthetic power grid modeling has been greatly enhanced with a useful guidance map and the direct optimization procedure to search for the best bus type assignments in a specific N-bus grid topology, saving the mandatory, but most the time not attainable, requirement of the availability of a set of realistic grid data with a comparable network size, to identify the search target. Finally, we examine the performance of the proposed measure both on the random-topology grid cases and on some realistic power grids. The obtained results verify that the proposed approach can efficiently determine the best set of bus type assignments.

Chapter 3

Statistical Settings of Generation Capacity, Generation Dispatch and Load

This chapter investigates the problem of generation and load settings in a synthetic power grid modeling of high-voltage transmission network, considering both electrical parameters and topology measures. Our previous study indicated that the relative location of generation and load buses in a realistic grid are not random but correlated. And an entropy based optimization approach has been proposed to determine a set of correlated siting for generation and load buses in a synthetic grid modeling. Using the approximate scaling function of aggregate generation capacity versus network size, the exponential distribution of individual generation capacities in a grid, and the non-trivial correlation between the generation capacity and the nodal degree of a generation bus we develop an approach to generate a statistically correct random set of generation capacities and then assign them to each generation bus in a grid. Based on the statistics collected from a number of realistic power grids, we then propose a statistical algorithm to determine generation/load setting and the generation dispatch at each generation bus according to its generation capacity and the statistic of dispatch ratios.

3.1 Introduction

To accomplish the goal of developing synthetic networks for the modeling of realistic power grid systems, extensive research has been conducted in order to recognize the salient grid related properties, to collect the statistics regarding the grid topologies and electrical parameters, hence to develop some useful models [53-54]. Reference [55] provides a comprehensive study on geographically approaches in synthetic power grid modeling. This paper describes several structural statistics and uses them to present a methodology to generate synthetic line topologies with realistic parameters. In [56] the authors propose a systematic methodology to augment the synthetic network base case for energy economic studies. In this paper the cost model of generators is determined based on the fuel type and generation capacity. This model can be utilized in electricity market and power system operation analysis.

As mentioned in [53], a valid synthetic grid model needs to include at least the following critical components: (a) the electrical grid topology which is fully defined by grid admittance matrix; (b) the generation and loads settings which indicate their correlated siting and sizing; (c) the transmission constraints which include the capacities of both transmission lines and transformers and etc.

The chapter presents our recent study results on the statistics of generation capacities and settings in a synthetic grid modeling. A set of approaches has been developed to generate a statistically correct random set of generation capacities and assign them to the generation buses in a grid according to the approximate scaling function of total generation capacity versus network size, the estimated exponential distribution of individual generation capacities, the non-trivial correlation between the generation capacity and the nodal degree of a generation bus. Then we determine the generation dispatch of each generation unit according to its capacity and the dispatch

ratio statistics, which we collected and derived from a number of realistic grid test cases. The proposed approaches is readily applied to determining the load settings in a synthetic grid model and to studying the statistics of the flow distribution and to estimating the transmission constraint settings.

The main contributions of this chapter are summarized as follows:(1) a set of statistical analysis on the generation capacities and its correlation with topology metrics is presented; (2) a statistical approach is proposed to determine the generation/load capacity at any generation bus in a synthetic grid modeling which takes into account both topology and electric measures; (3) a practical algorithm is developed to calculate actual power output of generating units with respect to their estimated maximum generation capacities and statistics of dispatch ratios of realistic grid data.

3.2 The statistics of generation capacity and load

Generation/ load settings in a synthetic grid modeling means to determine both the siting and the sizing of each generation unit and load. Our initial results indicate that the relative location of generation and load buses in a realistic grid are not random but correlated. And an entropy based optimization approach has been proposed to determine a set of correlated siting for generation and load buses in a synthetic grid modeling. In this section we focus on the problem to determine the sizing of generation units, i.e., the capacity and the dispatch at each generation bus.

In this section we first examine the statistical features of generation capacities in realistic power grids in terms of aggregate generation capacity, distribution of individual capacities, and their non-trivial correlation with nodal degrees. Table 3.1 presents the evaluation results of total

generation capacity, total demand, and total backbone transmission capacity of some realistic grids, with the corresponding definitions as:

$$P_{G,Max}^{tot} = \sum_{i \in G} P_{G,Max}^{i} \tag{3.1}$$

$$P_{L}^{tot} = \sum_{i \in L} P_{L}^{i} \tag{3.2}$$

$$F_{Max}^{tot} = \sum_{l \in BKB} F_{Max}^{l} \tag{3.3}$$

Table 3.1: Total generation capacity, demand and total backbone transmission capacity in some realistic grids

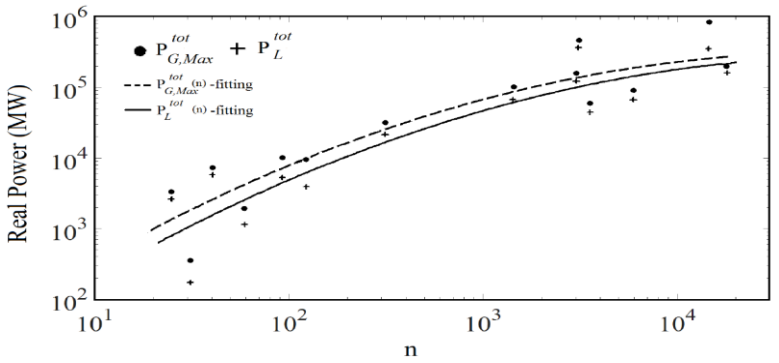
	$P_{G,\max}^{\text{tot}}$(MW)	P_{L}^{tot}(MW)	$r_{G/L}$	$F_{\max}^{\text{tot}}$(MW)
IEEE-24	3.41e3	2.85e3	1.19	1.87e4
IEEE-30	3.35e2	1.89e2	1.77	1.95e3
NE-39	7.37e3	6.25e3	1.18	3.37e4
IEEE-57	1.98e3	1.25e3	1.58	–
PEGASE-89	9.92e3	5.73e3	1.73	9.59e4
IEEE-118	9.97e3	4.24e3	2.35	–
IEEE-300	3.27e4	2.35e4	1.39	–
PEGASE-1354	1.29e5	7.31e4	1.76	8.99e5
PEGASE-2869	2.31e5	1.32e5	1.74	1.93e6
NYISO-2935	5.47e5	3.95e5	1.38	2.65e5
Polish-3375	7.11e4	4.84e4	1.47	1.37e6
ERCOT-5633	1.02e5	7.28e4	1.40	1.84e6
PEGASE-13659	9.81e5	3.81e5	2.57	2.27e6
WECC-16994	2.46e5	1.74e5	1.41	4.49e6

Figure 3.1 plots the total generation capacity and total demand in the grid systems as listed in Table 3.1 and the ratio between the two versus the network size. The scaling functions obtained from the curving fitting approach are given as follows:

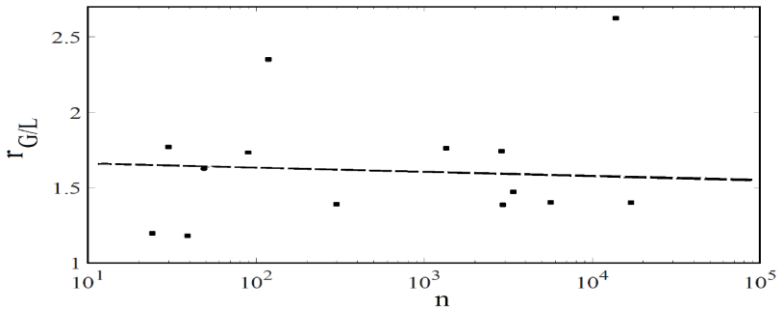
$$\log P_{G,Max}^{tot}(n) = -0.21(\log n)^2 + 2.06(\log n) + 0.66 \tag{3.4}$$

$$\log P_{L}^{tot}(n) = -0.20(\log n)^2 + 1.98(\log n) + 0.58 \tag{3.5}$$

where both $P^{tot}(n)$'s are measured in MW and the logarithm is with base 10.



(a)



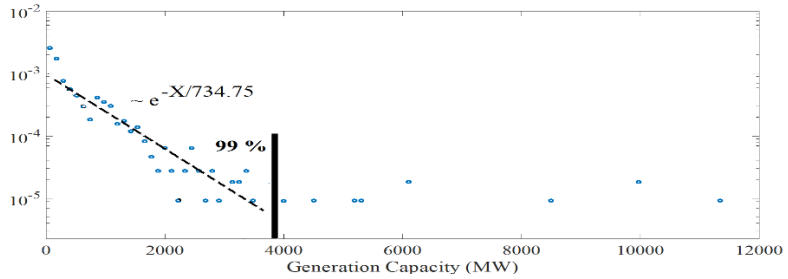
(b)

Figure 3.1: The scaling property of the total generation capacity and demand in realistic power grids:(a) the total generation marked by 'black squares' and total demand marked by '+' ; (b) the ratio of the total generation to the total demand.

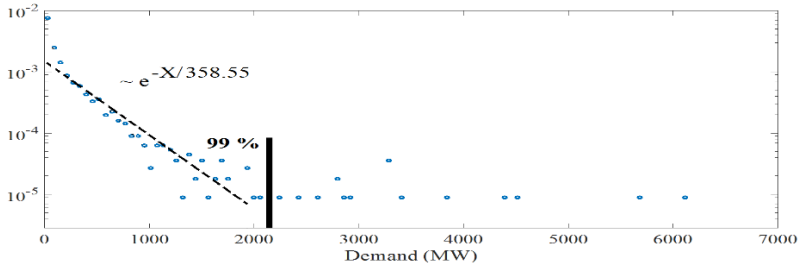
From Figure 3.1 we can see that when the network size is small (i.e. n < 300), the total generation capacity and demand in a grid tend to grow as a power function, i.e., $P^{tot} \propto n^{2.0}$. However, as the network size becomes larger, the scaling curves begin to bend down and grow slower than that power function. However the ratio of the two, $r_{G/L} = P^{tot}_{G,Max}/P^{tot}_{L}$, tends to slowly drop down from above 1.50 and draw closer to 1.00 as the network size increases.

Our initial study on the statistical distribution of generation capacity and demand within a power grid based on some realistic grid data such as the PEGASE, the WECC, and the NYISO systems, shows that the generation capacity and load settings approximately follow an exponential

distribution, as shown in Figures 3.2 -3.4. It is interesting to note that in the PEGASE 13659-bus system, about 99.9% of the generators have capacities following an exponential distribution except 0.1% with very large capacities falling out of the normal range. In the WECC 16994-bus system and the NYISO 2935-bus systems, approximately 99% of the generation capacities (and the loads as well) follow an exponential distribution while only 1% with extremely large capacities (or demands) falling out of the expected normal range. We will continue the study on this aspect to determine the cause of observed distribution exceptions, which may either come from an inherent heavy-tailed distribution or only result from boundary equalization in a network reduction modeling. Our study also shows that there exists correlations between the total number of branches connecting a bus (that is, its node degree) and the generation or load attached to the bus.

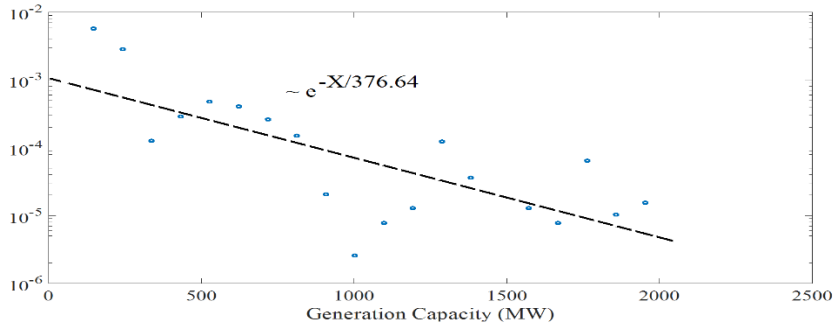


(a)

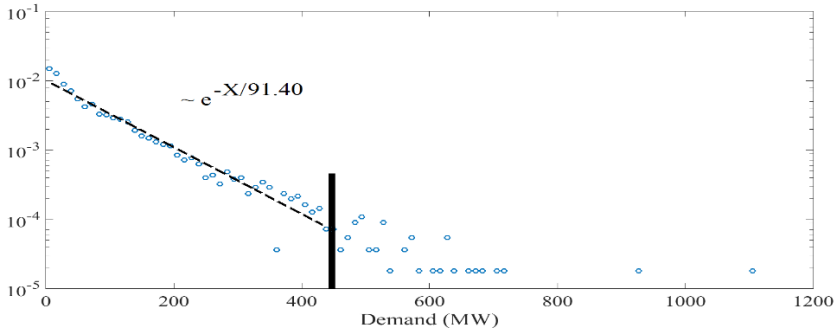


(b)

Figure 3.2: The empirical PDF of (a) generator capacities and (b) demands in power grids in the NYISO-2935 bus system

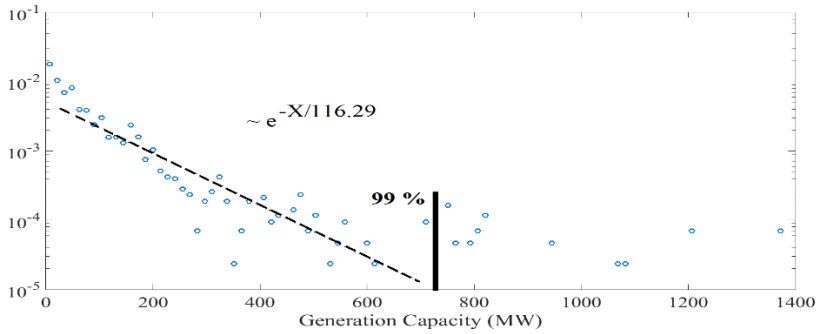


(a)

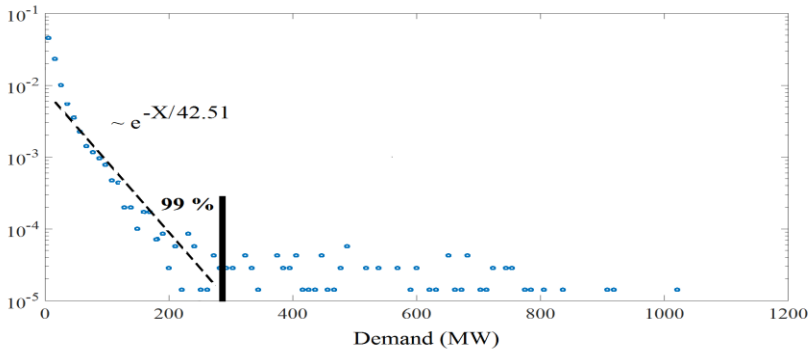


(b)

Figure 3.3: The empirical PDF of (a) generator capacities and (b) demands in the PEGASE-13659 bus system



(a)



(b)

Figure 3.4: The empirical PDF of (a) generator capacities and (b) demands in the WECC-16994 bus system

3.2.1 Statistical-based algorithm to assign generation capacities to generation buses

After studying the scaling property and distribution of generation capacities in a grid, it would be critical to examine the correlation between the generation capacities and other topology metrics. To simply the following statistical analysis and the algorithm development for generation capacity generation and assignment in a synthetic grid modeling, we define two normalized variables as

$$\overline{P_{g_n}^{Max}} = P_{g_n}^{Max} / \max_i P_{g_i}^{Max}, \tag{3.6}$$

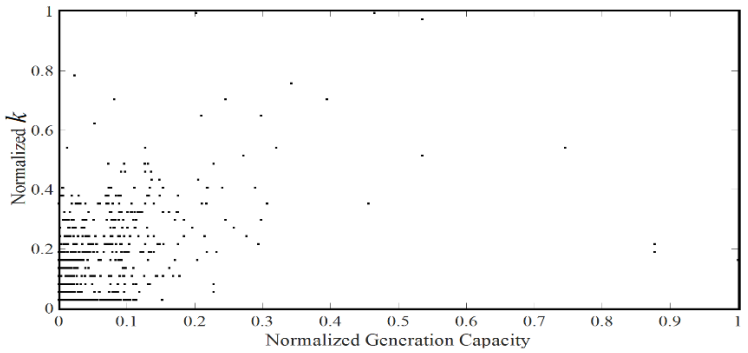
$$\overline{k_n} = k_n / \max_i k_i. \tag{3.7}$$

So that both variables will assume values limited within [0, 1]. The statistics collected from the date of a number of realistic grids indicate that there exists a considerable correlation between the nodal degree of a generation bus and its capacity with a Pearson coefficient of $\rho\left(\overline{\mathrm{P_{g_n}^{Max}}}, \overline{\mathrm{k_n}}\right) \in [0.25, 0.5]$. Figure 3.5.a shows the scatter plots of normalized generation capacity versus the normalized node degree of some sample grids as the NYISO-2935 and WECC-16994 systems, which exhibit similar distribution patterns. That is, most data points are densely located within the region of $\overline{P_{g_n}^{Max}} \in [0, 0.2]$ and $\overline{k_n} \in [0, 0.5]$, while very few located in the region of $\overline{P_{g_n}^{Max}} \geq 0.6$.

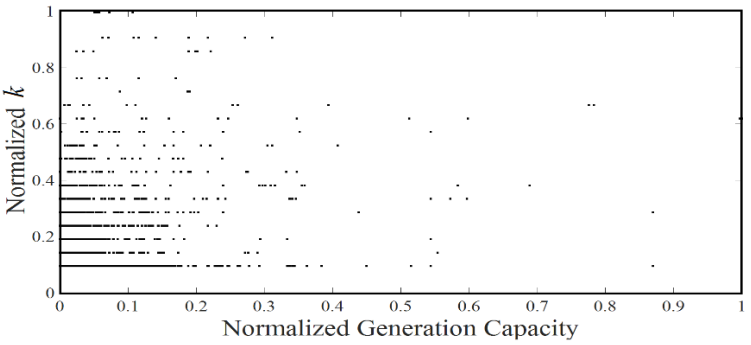
When two variable, say $\overline{P_{g_n}^{Max}}$ and $\overline{k_n}$ are considered, then we may put them together to get a pair of numbers, that is, a point $(\overline{P_{g_n}^{Max}}, \overline{k_n})$ in the two-dimensional space. These two-dimensional variables are considered mainly by their density function $f(\overline{P_{g_n}^{Max}}, \overline{k_n})$, which integrated on a set A gives the probability of the event that the value of $(\overline{P_{g_n}^{Max}}, \overline{k_n})$ is in the set A:

$$\Pr(A) = \Pr\left(\left(\overline{P_{g_n}^{Max}}, \overline{k_n}\right) \in \mathrm{A}\right) \tag{3.8}$$

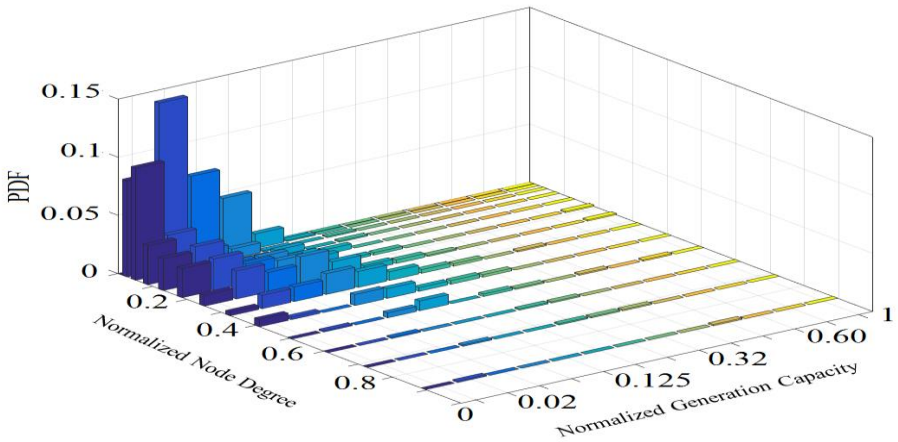
Figure 3.5.c illustrates the 2-D empirical probability density function (PDF) of normalized node degree versus normalized generation capacity in realistic power grids. Based on the obtained empirical PDF, a two- dimensional probability distribution table shown in Table 3.2 can be formulated to enable the algorithm development in next section to assign the generated capacity values to each generation bus in a grid according to its normalized nodal degree.



(a)



(b)



(c)

Figure 3.5: Scatter plots of normalized node degree versus normalized generation capacities in (a) NYISO and (b) WECC systems, and (c) 2-D empirical PDF of normalized node degree versus normalized generation capacity in realistic power grid

In the rest of this section we will introduce an approach to generate a statistically correct random set of generation capacities and assign them to the generation buses in a grid according to the approximate scaling function of total generation capacity versus network size, the estimated exponential distribution of individual generation capacities, and the non-trivial correlation between the generation capacity and the nodal degree of a generation bus.

Table 3.2: Probability analysis of normalized node degree and normalized generation capacity in realistic power grid

$\overline{P_{g_n}^{Max}}$ \ $\overline{k_n}$	0.00–0.01	0.01–0.03	0.03–0.06	0.06–0.1	0.1–0.15	0.15–0.21	0.21–0.28	0.28–0.36	0.36–0.45	0.45–0.55	0.55–0.66	0.66–0.78	0.78–1.00	Marginal Prob
1.00–0.78	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.002
0.78–0.66	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.004
0.66–0.55	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.003
0.55–0.45	0.000	0.001	0.000	0.004	0.008	0.000	0.002	0.001	0.000	0.001	0.000	0.001	0.000	0.018
0.45–0.36	0.006	0.002	0.000	0.009	0.008	0.003	0.003	0.002	0.000	0.000	0.000	0.000	0.000	0.034
0.36–0.28	0.003	0.011	0.012	0.017	0.013	0.007	0.003	0.002	0.000	0.001	0.000	0.000	0.000	0.072
0.28–0.21	0.009	0.024	0.016	0.024	0.013	0.004	0.003	0.001	0.000	0.000	0.000	0.000	0.001	0.097
0.21–0.15	0.025	0.027	0.016	0.013	0.009	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.001	0.097
0.15–0.1	0.027	0.031	0.010	0.010	0.005	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.088
0.1–0.06	0.033	0.017	0.003	0.003	0.005	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.063
0.06–0.03	0.090	0.030	0.01	0.008	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.151
0.03–0.01	0.082	0.140	0.070	0.04	0.010	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.360
0.01–0.00	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Marginal Prob	0.283	0.291	0.147	0.141	0.077	0.022	0.017	0.008	0.001	0.003	0.000	0.001	0.002	1.000

Given a random topology power grid with N buses among which N_g buses have generation units, we may determine the aggregate generation capacity $P_g^{tot}(N)$using equation (3.4) and generate a statistically correct random set of N_g generation capacities which follows an exponential distribution of generation capacities with 1% of generated capacities switched to super large values. Then some scaling adjustment may be necessary to remain the same aggregate generation capacity given by $P_g^{tot}(N)$. Next an algorithm will be developed to assign the generation capacities

to each generation bus with respect to the statistical pattern presented in Table 3.2. Below is a step-by-step algorithm procedure description:

Step 1: Estimate the total generation capacity $P_g^{tot}(N)$ using (3.4).

Step 2: Generate a statistically correct random set of generation capacities $[{P_g}_n^{Max}]_{1\times N_g}$. It should be noted that 99 % of generated capacities follow the exponential distribution and remaining one percent is guaranteed to take supper large values (2~3 times greater than all generation capacities which follow the exponential distribution).

Step 3: Do the scaling of generated capacities if $\sum_{n=1}^{N_g} {P_g}_n^{Max} > 1.05P^{tot}$ to make the aggregate generation capacity remain the range specified by $P^{tot}(N)$. And the scaling function is given as:

$$[{P_g}_n^{Max}]'_{1\times N_g} = [{P_g}_n^{Max}]_{1\times N_g} \times \frac{P^{tot}}{\sum_{n=1}^{N_g} {P_g}_n^{Max}} \quad (3.9)$$

where $[{P_g}_n^{Max}]'_{1\times N_g}$ is the updated generation capacities.

Step 5: Normalize both generation capacities and node degrees and categorize them evenly into 100 square regions with specific range of $\overline{{P_g}_n^{Max}}$ and $\overline{k_n}$.

Step 4: Calculate the node degree for all generation buses 1 to $\boldsymbol{N_g}$ based on the topology information of given random power grid.

Step 6: Check the similarity with Table 3.2 and reorder the mismatched segments.

Step 7: Assign the generated capacities to nominated generation buses with respect to their node degrees.

Step 8: Convert the normalized values to the actual values.

The flowchart of the proposed algorithm is depicted in Figure 3.6.

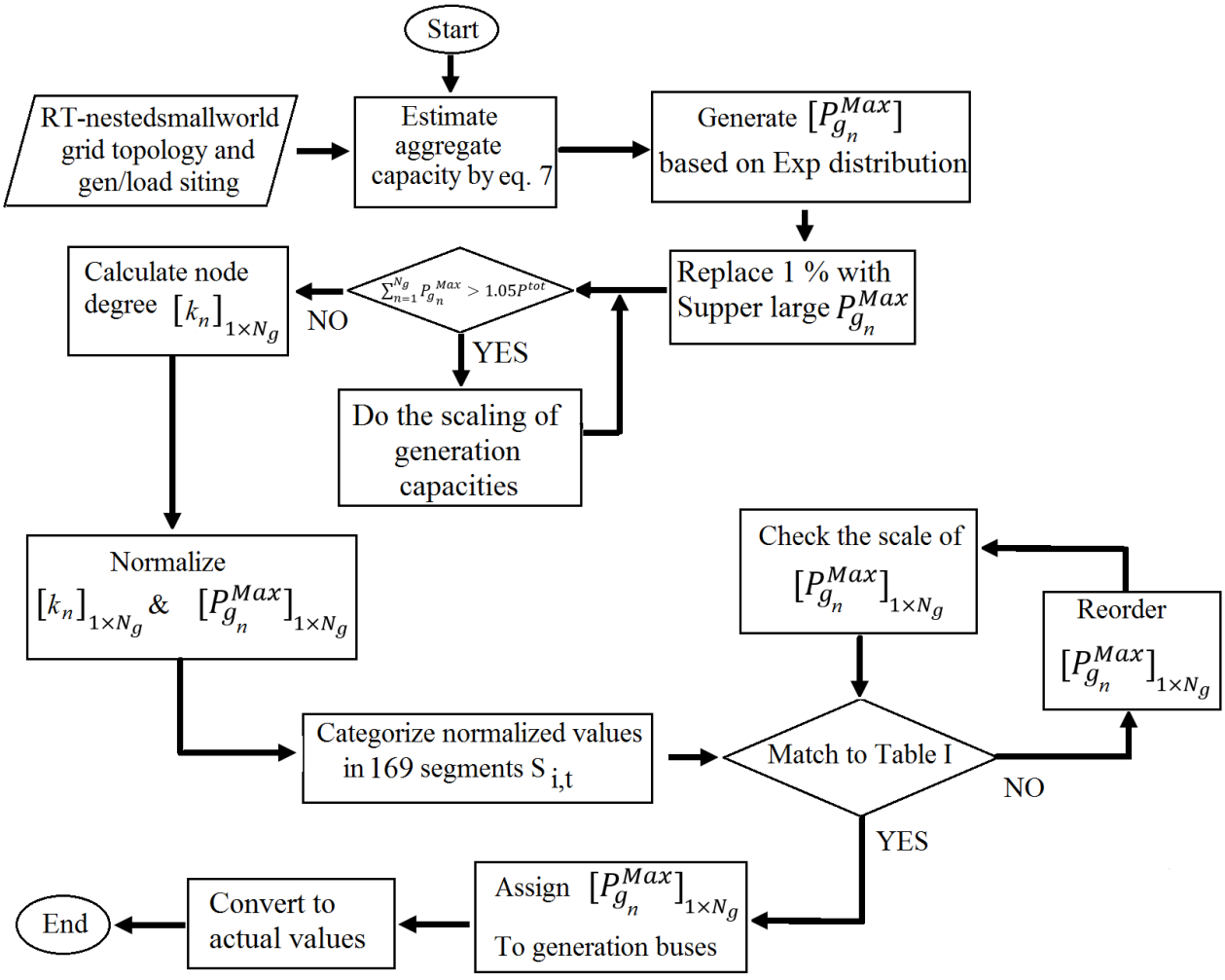


Figure 3.6: Flowchart of the proposed algorithm to assign random generation capacities to generating units

3.3 Statistical-based algorithm to assign loads to load buses

This section introduces a statistical –based approach to generate a set of static loads and assign them to the load buses. In this section we first investigate the statistical features of static loads in

realistic power grids in terms of total demand, distribution of individual loads, and their non-trivial correlation with nodal degrees.

In section 3.2 we reported the results of our analysis on scaling function of total demand in a grid versus its network size. Figure 3.1.a illustrates the scaling of aggregate demand as a function of network size

$$\log P_L^{tot}(N) = -0.20(\log N)^2 + 1.98(\log N) + 0.58 \quad (3.10)$$

where $P_L^{tot}(N) = \sum_{n=1}^{N_L} P_{L_n}$ denotes the total generation capacity and N_L is the total number of load buses. The obtained results show that in realistic power networks the total demand tends to grow as a power function. It is important to point out that in a given grid topology although the total demand can be fully determined by the presented scaling function, failure to maintain a balance between total load and resources causes frequency to vary from its target value. Thus, it is crucial to consider both scaling function and aggregate sources to achieve a reasonable value for total demand.

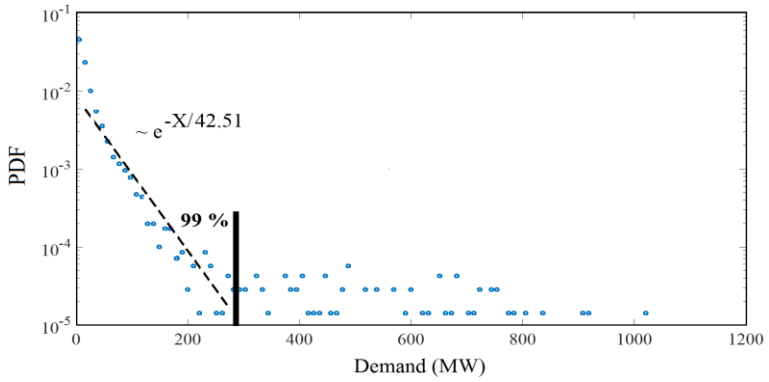


Figure 3.7: Empirical PDF of loads in WECC-16994 bus system

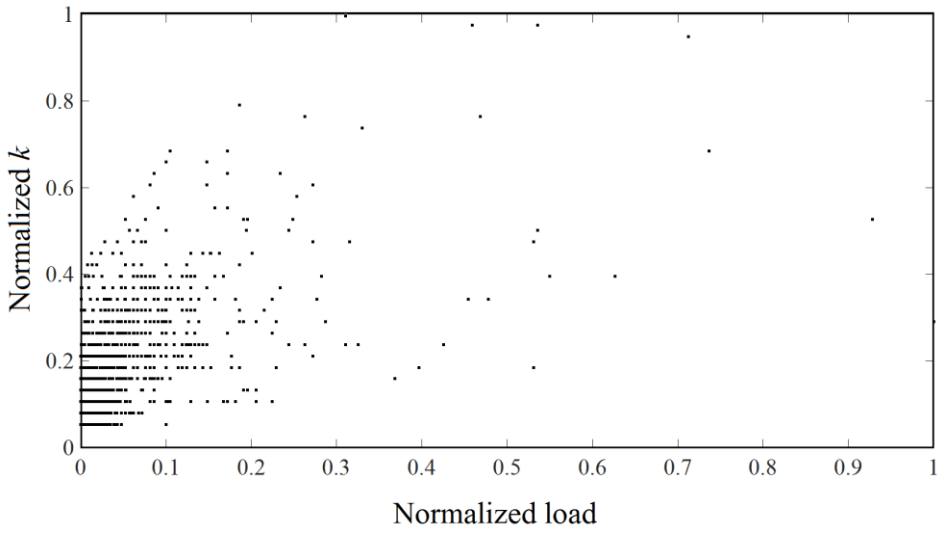


Figure 3.8: Scatter plots of normalized node degree versus normalized loads

Our initial experiments on the statistical distribution of loads within realistic power grids show that, like generation capacities, about 99% of the loads follow an exponential distribution with about 1% having extremely large demands falling out of the normal range defined by the expected exponential distribution. Figure 3.7 show the statistical distribution of loads in the WECC-16994 bus system. The fitting curve is depicted as a dashed line for the distribution function of P_L.The straight line in the log plot implies that about 99% of load capacity in realistic power grids tend to drop down as an exponential function with mean value of $\beta = 42.51$.

Given a realistic power grid with N buses among which $N_L$buses have loads, we may examine the correlation between the total number of branches connecting a bus (that is, its node degree) and the total load attached to the bus, like what we did in section 3.2.1 for the generation capacities. To simply the following statistical analysis, we consider the normalized node degree presented in (3.7) and normalized load as

$$\overline{P_{L_n}} = P_{L_n} / \max_i P_{L_i} \tag{3.11}$$

So that the normalized loads will assume values limited within [0, 1]. Our statistical results show that in realistic power grids the Pearson's coefficient of correlation varies in the range of 0.3-0.6. Figure 3.8 displays the scatter plot of normalized load capacities and normalized node degree which can be further used to generate the 2-D empirical PDF of some sample grids like WECC-16994 buses system.

By averaging the statistics of available realistic grid data, we may extract an empirical 2-dimensional probabilistic density function (PDF) for the normalized load values and nodal degree $\left(\overline{P_{L_n}}, \overline{k_n}\right)$. Based on the 2-D empirical PDF over the obtained uneven grid division (see Figure 3.9) a two-dimensional probability distribution table shown in Table 3.3 can be formulated to enable an algorithm to assign the generated load values to each load bus in a grid according to its normalized nodal degree.

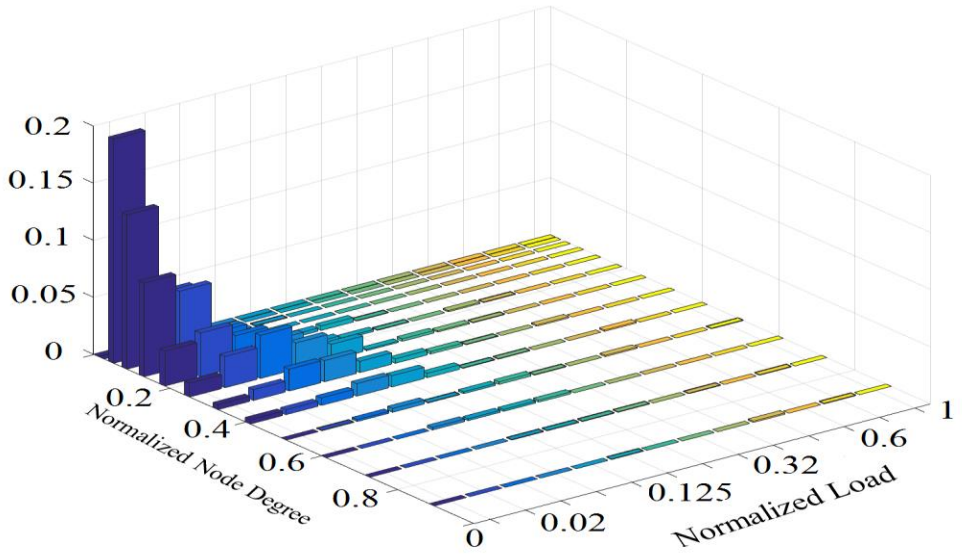


Figure 3.9: The 2-D empirical PDF of load versus normalized node degree in WECC-16994 bus system.

Table 3.3: Probability analysis of normalized node degree and normalized load in WECC-16994 bus system

$\overline{P_L}$ \ $\overline{k_n}$	0.00–0.01	0.01–0.03	0.03–0.06	0.06–0.1	0.1–0.15	0.15–0.21	0.21–0.28	0.28–0.36	0.36–0.45	0.45–0.55	0.55–0.66	0.66–0.78	0.78–1.00	Marginal Prob
0.78–1.00	0	0	0	0	0	0.001	0	0	0	0.001	0	0.001	0	0.002
0.66–0.78	0	0	0	0	0.001	0.001	0.001	0.001	0	0.001	0	0.001	0	0.003
0.55–0.66	0	0	0	0.002	0.002	0.002	0.002	0	0	0	0	0	0	0.008
0.45–0.55	0	0.001	0.001	0.003	0.001	0.001	0.002	0.001	0	0.001	0	0	0.001	0.012
0.36–0.45	0.003	0.004	0.006	0.010	0.010	0.003	0.001	0.001	0	0	0.001	0	0	0.041
0.28–0.36	0.004	0.009	0.019	0.018	0.009	0.003	0.002	0.001	0	0.001	0	0	0	0.069
0.21–0.28	0.012	0.027	0.037	0.022	0.013	0.001	0.002	0.001	0.001	0	0	0	0	0.118
0.15–0.21	0.030	0.038	0.027	0.015	0.003	0.001	0.001	0	0.001	0.001	0	0	0	0.119
0.1–0.15	0.082	0.066	0.022	0.004	0.002	0.003	0.001	0	0	0	0	0	0	0.183
0.06–0.1	0.135	0.058	0.010	0.002	0	0	0	0	0	0	0	0	0	0.205
0.03–0.06	0.196	0.033	0.005	0	0.001	0	0	0	0	0	0	0	0	0.235
0.01–0.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000
0.00–0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0.000
Marginal Prob	0.464	0.238	0.130	0.076	0.042	0.018	0.012	0.005	0.002	0.005	0.001	0.001	0.005	1.000

The approach to creating a statistically correct random set of load capacities and assigning them to the load buses begins by determining the aggregate load capacity $P_L^{tot}(N)$ using equation (3.10) and generating a statistically set of N_L load capacities which follows an exponential distribution of generation capacities with 1% of generated loads switched to super large values. In order to accurately formulate a synthetic power grid we need to assign load capacities to the load buses in a way consistent with that of a realistic grid. Therefore, the proposed approach will be developed to assign the load capacities to each load bus with respect to the statistical pattern presented in Table 3.3. Expect for the statistical pattern, the procedure is exactly like that of generation capacities assignment in section 3.2.1 (See Figure 3.6)

3.4 The statistics of generation dispatch in realistic power grids

In realistic power grids generation units are committed to serve the time-varying demand of customer loads according to their costs, operational limits, network constraints, and environmental constraints and other factors. However, in the development of synthetic grid modeling, we lack most of electrical and economical information as mentioned above. Therefore, we seek a solution to determine the generation dispatch in a grid based on the statistical analysis so that it will be consistent with the statistics observed in realistic power grids.

The main idea is that by studying the possible correlation between generation capacities $P_{g_n}^{Max}$ and power dispatch $P_{g_n}(t_0)$ in realistic power grids an appropriate method can be constructed to estimate the power output of generating units. Figure 3.10 shows the scatter plot of actual generation versus generation capacities in a typical realistic power grid. From Figure 3.10, it is evident that there exists strong correlation between the two as $\rho\{P_{g_n}^{Max}, P_{g_n}(t_0)\} = 0.9563$ in this case and greater than 70% in other realistic grids. Two additional observations could be made regarding this figure: (1) in a typical realistic power grid about 0 ~ 20 % of generators are out of service, and interestingly only a small number of them belong to the super large power units (red dots in Figure 3.10. It should be noted here that electricity market operation requirements of capacity reserve, load levels of the system, or the annual overhaul schedule might be the reasons causing this phenomenon; (2) about 50 % of committed power units are operated very closed to their maximum generation capacities and the power dispatch for the rest of generating units vary between minimum and maximum generation capacity.

Figure 3.11 shows the scatter plot of generation capacities versus dispatch factor $\alpha_n = P_{g_n}(t)/\, P_{g_n}^{Max}, n = 1, \ldots, N_g$, in the WECC-16994 bus system. Figure 3.11 implies that in a typical power grid small and mid-size power units tend to have a wider range of dispatch factor compared with those unit of larger size. That means, as a unit size becomes larger, its dispatch factor tends to grow larger too and draw closer to 1.00. Figure 3.12 shows the statistical distribution of uncommitted generation capacities. The distribution is fitted very well by the exponential PDF with mean value of $\beta = 90.9$, i.e. about 99% of statistics satisfy to exponential distribution. Figure 3.13 plots the empirical 2-D distribution function of $(\overline{P_{g_n}^{Max}}, \alpha_n)$ extracted from WECC-16994 system, in a similar way as what has been done in Section 3.2. It should be noted here that the main motive of using 2-D empirical PDF is to depict the actual relationship between the generation dispatch and the total generation capacity attached to the bus.

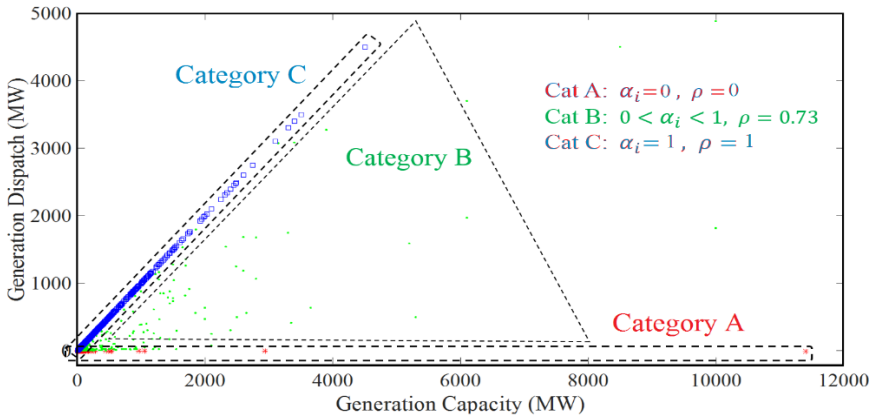


Figure 3.10: Scatter plot for generation dispatch versus generation capacity in realistic power grid

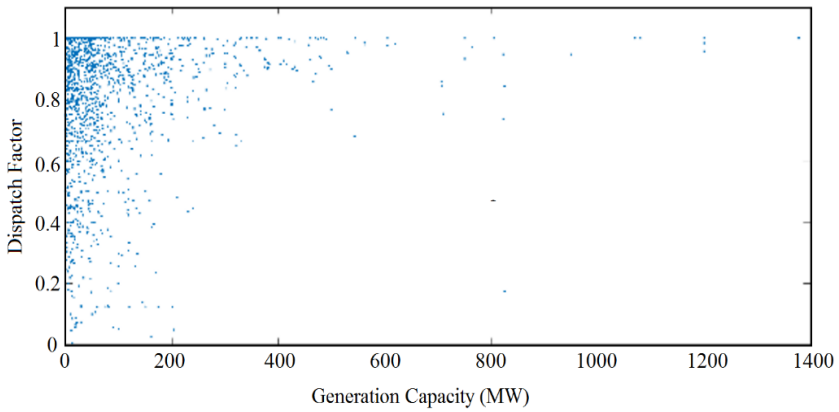


Figure 3.11: Scatter plot for dispatch factor α versus maximum generation capacity in WECC-16994 bus system

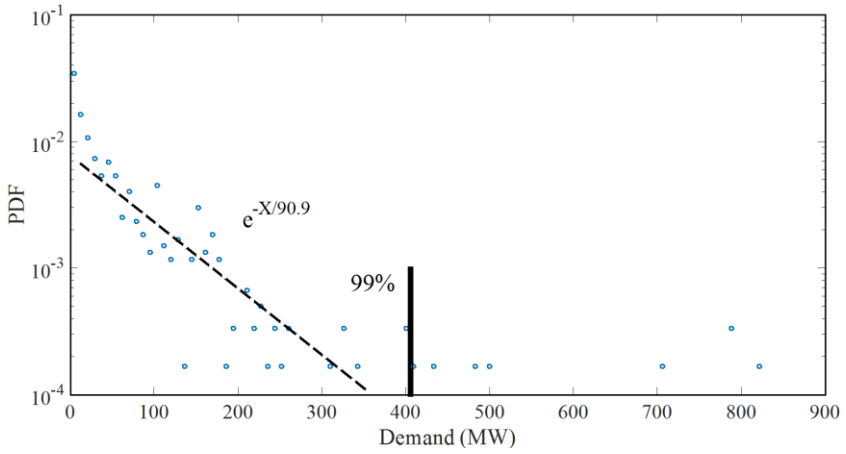


Figure 3.12: Empirical PDF of uncommitted generation capacities in realistic power grid

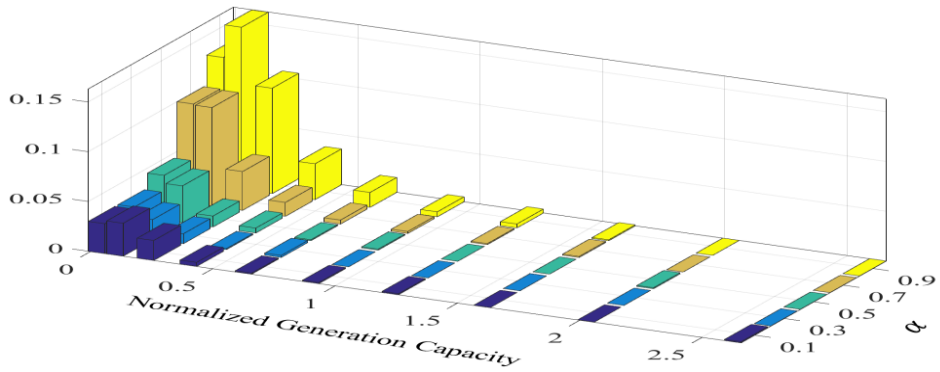


Figure 3.13: The 2-D empirical PDF of generation dispatch versus normalized generation capacity in WECC-16994 bus system.

Table 3.4: Probability analysis of generation dispatch and normalized generation capacity in WECC-16994 bus system

		$\overline{k_n}$										Marg Prob
		0 / 0.05	0.05 / 0.15	0.15 / 0.3	0.3 / 0.5	0.5 / 0.75	0.75 / 1.05	1.05 / 1.4	1.4 / 1.8	1.8 / 2.25	2.25 / 3	
α	0.8 / 1	0.130	0.163	0.106	0.036	0.015	0.005	0.004	0.001	0	0	**0.455**
	0.6 / 0.8	0.10	0.099	0.039	0.014	0.004	0.002	0.001	0.001	0	0	**0.260**
	0.4 / 0.6	0.045	0.037	0.011	0.005	0.001	0.001	0	0	0.001	0	**0.101**
	0.2 / 0.4	0.030	0.020	0.010	0.001	0.002	0	0	0	0	0	**0.072**
	0 / 0.2	0.030	0.033	0.020	0.004	0.001	0.001	0.001	0	0	0.001	**0.091**
Marg Prob		**0.335**	**0.352**	**0.186**	**0.06**	**0.022**	**0.009**	**0.006**	**0.002**	**0.001**	**0.001**	1

The above mentioned results presented in Table 3.4 can provide us useful guidelines to determine the generation dispatch of all installed generators in a synthetic power grid. The obtained pattern will help us to propose a statistics-based algorithm to improve the generation setting in existing synthetic power grid modeling, namely, to assign the best set of generation dispatches which are consistent with that of realistic grids. Briefly, this method is composed of three steps as follows:

Step 1: *Uncommitted units*; select a set of generation capacities form $[P_{g_n}^{Max}]_{1\times N_g}$ and consider them as uncommitted units. In this step, (0~20) % of generating units are considered as uncommitted units with $\alpha = 0$ which follow the exponential distribution in Figure 3.12 and remaining units are guaranteed to take dispatch factor value between (0 ~ 1]. It should be noted that the selected capacities should follow the empirical PDF presented by Figure 3.9. Hence, the capacity of nominated units must be close enough to the random values generated by exponential distribution.

Step 2: *Committed units*; this step provides a procedure to determine the generation dispatch at each generation bus according to an empirical 2-D distribution function of ($P_{g_n}^{Max}, \alpha_n$) extracted

from realistic grids. Table 3.4 summarizes the obtained results from statistical analysis of WECC-16994 systems. This table can provide us useful guidelines to determine the generation dispatch of committed generators (category B) in a synthetic power grid. First we need to select 40~50 % of remaining units as committed but not full-load units. In this procedure, the nominated units are selected according to the empirical PDF of generation capacities of category B, in a similar way as what has been done in step 1. Then an algorithm will be developed to assign the best generation dispatch to each generation bus with respect to the statistical pattern presented in Table 3.4.

Step 3: *Full-load units*; in this step all remaining units are considered as full-load sources with dispatch factor $\alpha = 1$.

The flowchart of the proposed algorithm is depicted in Figure 3.14.

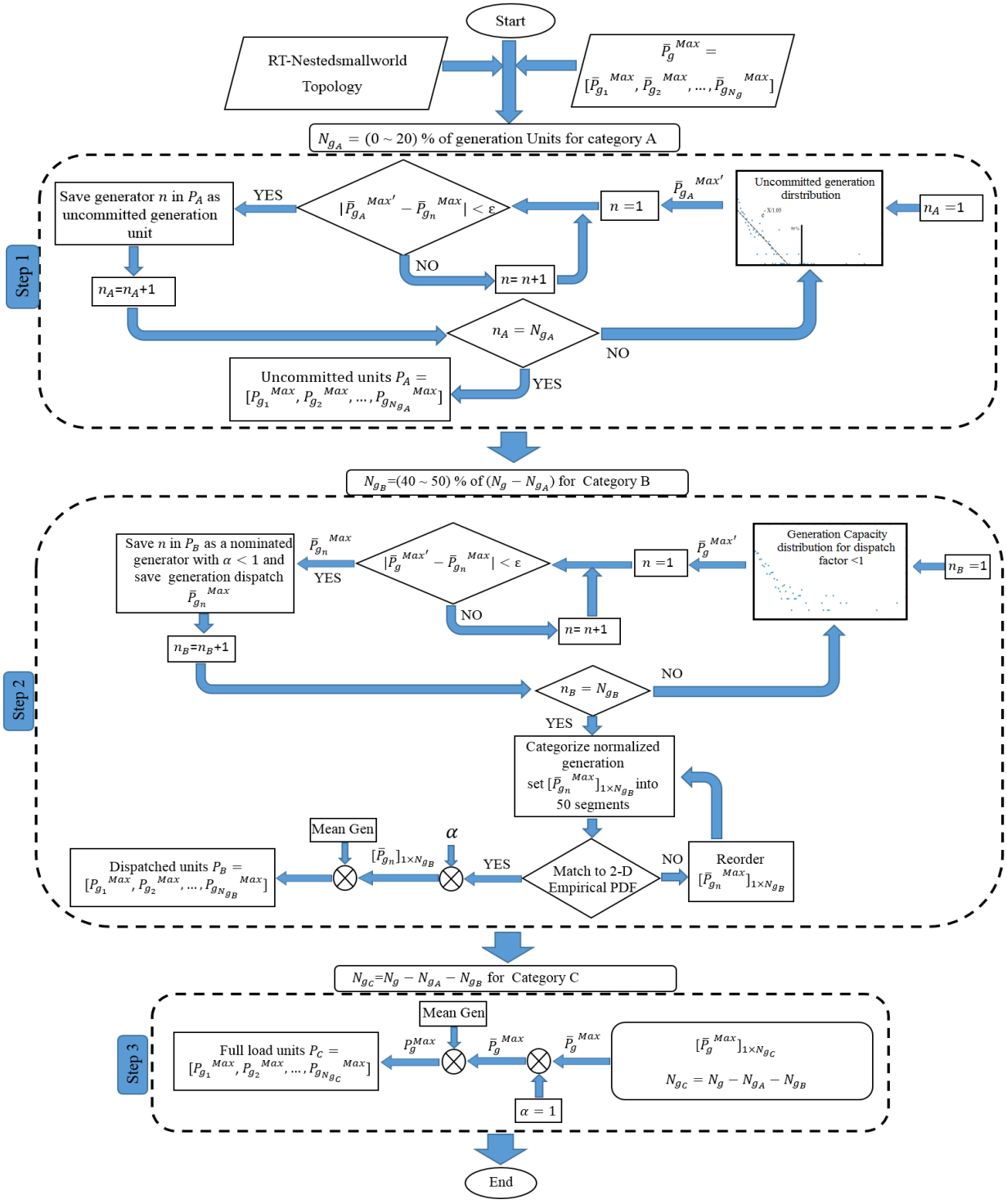


Figure 3.14: Flowchart of the proposed algorithm to assign random generation dispatch to generating units

3.5 Conclusion

In this chapter we examine the statistical features of generation capacities in realistic power grids in terms of aggregate generation capacity, distribution of individual capacities, and their non-trivial correlation with nodal degrees. Our study on the statistical distribution of generation capacities shows that more than 99% of the generation units follow an exponential distribution with about 1% having extremely large capacities falling out of the normal range defined by the expected exponential distribution. Based on the obtained results presented in this paper there exists non-trivial correlations between the total number of branches connecting a bus (say node degree k) and the total generation attached to the bus. This paper also indicates that there exists strong correlation in between the power dispatch and the generation capacities.

Based on the above results, we develop an algorithm to generate a statistically correct random set of generation capacities and then assign them to each generation bus in a grid. We then propose a statistical approach to determine the generation dispatch at each generation bus according to its generation capacity and the statistic of dispatch ratios.

Chapter 4

Transmission line capacity assignment

In this chapter we mainly focus on the statistical analysis of transmission line capacities. This section addresses the research question of whether there exists a relationship between the transmission line capacities and topological characteristics of power grids. We utilize the realistic power grid data sets to extract a model for the transmission line capacity in the synthetic high voltage grids. Motivated by the need for accurate line capacity assignment and in order to improve the synthetic power grid modeling, we performed a set of statistical experiments on some available realistic networks such as, NYISO_2935, ERCOT-5633 and WECC- 16994 bus systems, and we found that (a) the capacity of transmission lines follows a well-known distribution and it can be fully defined by mathematical definitions (b) the line capacities assignment in a given synthetic power grid model should be improved using obtained statistical results which is found consistent with what is observed in realistic power grids. To our best knowledge, this is a first attempt of statistical analysis of transmission line capacities and its application in generating synthetic power grids.

4.1 Introduction

To avoid exceeding a transmission line limits, most lines have established a maximum power flow level that cannot be exceeded at any time. This static thermal rating is based on the worst-case scenario with respect to environmental conditions: a hot day with full sun and a very low wind that is typically assumed to be around 2ftps. If the line's maximum power flow is exceeded under these conditions, the line could sag to the point where it would come into contact with trees, trucks, boats, or other nearby hazards [57-60]. However, power levels that would cause unacceptable sag in a transmission line under these conditions may not cause unacceptable sag on a cooler, cloudy, windy day, when the rising temperature caused by current through the line is offset by the colder operating environment and the cooling effect of the wind. Because the environmental conditions are almost never as unfavorable as those used to calculate the static thermal rating of a line, the line's true maximum capacity is underestimated the vast majority of the time.

This chapter mainly focuses on statistical analysis of transmission line capacities in terms of both topological and electrical parameters. In this chapter we examine transmission line capacities based on both network topology metrics and some newly proposed electrical indexes. The obtained results show that the issue of transfer capacity assignment not only emerges as an electrical optimization concern, but some topological metrics must be considered to find the best line capacity assignment that is consistent with what is manifested in real-world grid. These results then will be used to develop a new methodology to appropriately characterize the line capacity assignment and improve the synthetic power grid modeling.

4.2 Line capacity distribution in real-world power grids

In power electric systems transferring power from one bus to another is distributed over transmission lines based on Kirchhoff's current and voltage laws. These laws are derived from topology of the network and are unique. It should be noted that every transmission line has a limit on the amount of power it can transfer. This limit can be imposed by several phenomena including voltage and stability limits. Technically, the transfer limits must be estimated to encompass both normal and overloaded operations. Therefore, given the connecting topology of a power grid, the transfer limits are traditionally considered 20 % upper than the normal operation. In this section we will use the statistical analysis to assess whether the distribution of transmission line capacities follows a well-known mathematical distribution. The importance of this issue stems from the fact that in a synthetic power grid modeling, we need a realistic pattern to generate the reasonable values for transmission line capacities. Although, the operational concerns should be considered to determine the reasonable values for the line capacities, the statistical analysis needs to account for the constraints of transmission capacity setting.

Using ERCOT-5633 and WECC-16994 buses systems, as two realistic power grid examples, we investigate the distribution of realistic line capacities. Due to the lack of sufficient sample data, we have to combine the transmission line data from the two systems. It should be noted that the distribution of a random variable, say transmission line capacity, is a function that describes how likely we can obtain the different possible values of the random variables. Figure 4.1 depicts the distribution of normal, long-term emergency and short-term emergency rating. Technically,

normal rating refers to maximum loading that the line can carry continuously. Long-Term Emergency (LTE) rating is defined as the maximum loading, which may be carried by a line up to four hours. And Short-Term Emergency (STE) rating refers the maximum loading for a line up to fifteen minutes. The statistics of the generated distribution functions show that in the realistic power grids, the capacity of transmission lines follows the Generalized Extreme Value Distribution (see equation 4.1 and 4.2).

From Figure 4.1 the probability that capacity rating of transmission lines falls into [0 MW, 1000MW] is much higher than other intervals. Further, this figure provides a pattern to generate reasonable capacity rating in synthetic power grid methods. Table 4.1 shows the best fitting distribution functions with the corresponding estimated parameters. Based on the obtained results, it is clear that the distribution of capacity lines can be fully defined by a mathematical distribution, and the fitted curve, presented by (4.1) and (4.2), provides a pattern to generate reasonable line capacity values for a given power grid topology.

Generalized Extreme value Distribution:

$$G(x|\mu,\sigma,\xi) = \frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)} \tag{4.1}$$

Where μ, σ and ξ are fitting parameters, and:

$$t(x) = \begin{cases} \left(1 + \left(\frac{x-\mu}{\sigma}\right)\xi\right)^{-1/\xi} & if \xi \neq 0 \\ e^{-(x-\mu)/\sigma} & if\ \xi = 0 \end{cases} \tag{4.2}$$

Although the presented distributions can generate reasonable values for the capacity of transmission lines, it is really challenging to find the best line capacity assignment. Indeed, we need to evaluate the statistical behavior of realistic capacities and the relationships between the

line capacities and topological characteristics of power grids. In order to address this issue, first we need to describe different types of links in real-world networks.

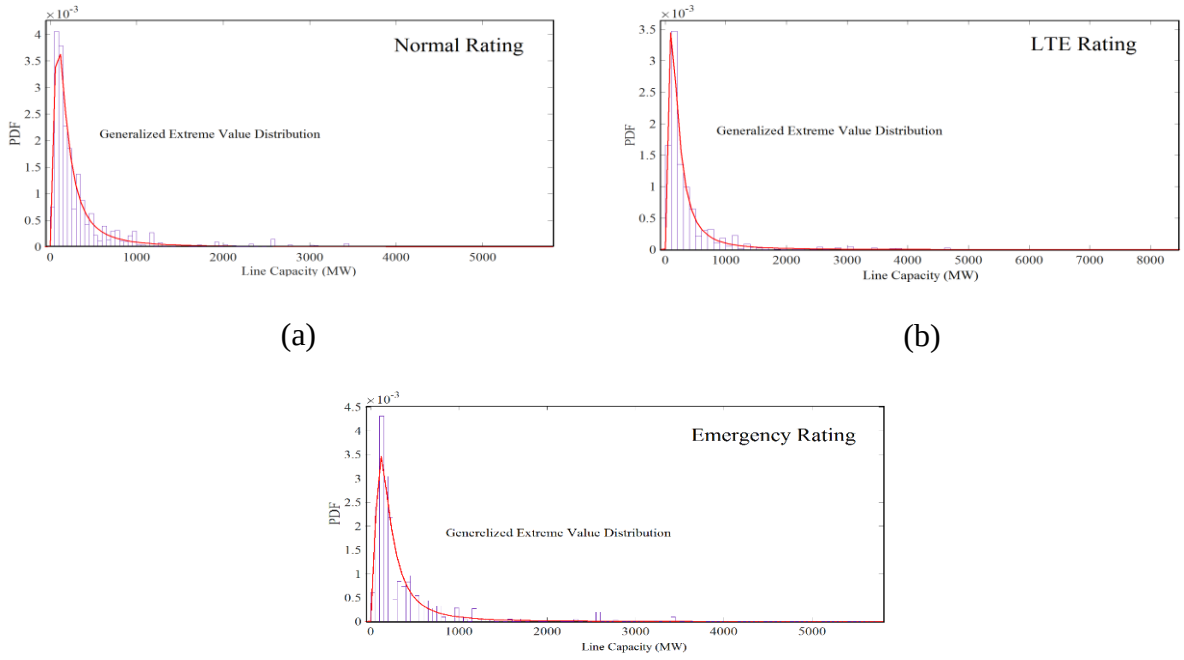


Figure 4.1: Line capacity distribution for (a) Normal, (b) long-term emergency and (c) short-term emergency rating.

Table 4.1: Distribution fitting for transmission line capacities

		Parameter Estimated	
	Distribution Fitting	**ERCOT**	**WECC**
Normal Rating	$G(x\|\mu,\sigma,\xi)$	μ = 230.77	μ = 136.92
		σ =130.21	σ =110.95
		ξ =0.40	ξ =0.64
LTE Rating	$G(x\|\mu,\sigma,\xi)$	μ = 240.80	μ = 153.92
		σ =136.60	σ =125.19
		ξ =0.41	ξ =0.64
STE Rating	$G(x\|\mu,\sigma,\xi)$	μ =247.14	μ =160.52
		σ =142.1	σ =121.68
		ξ =0.42	ξ =0.58

4.3 Transmission line capacity related with topology features

In this section we continue our experiments on realistic power grids to evaluate the statistical behavior of their line capacities. First we investigate the relationship between the link types and different values of transmission line capacities. Then we define and examine a new index, called *Neighboring Capacity Ratio (NCR)*, to characterize the dependency of line capacity on topological metrics. As mentioned in previous section, in a typical power grid, all the buses can be categorized into three groups; Generation buses (G) which connect generators, Load buses (L) which connect loads and Connection buses (C) which form the transmission network. Considering different bus types G/L/C, all the branches in a given network can be grouped into six link type categories; GG, LL, CC, GL, GC and LC. That is, the link type of a branch is determined by the bus types of its end buses. Table 4.2 and 4.3 provide the topological analysis of transmission line capacities in realistic power grids. Due to the wide range of line capacities it is necessary to categorize the line capacities into smaller intervals, so that we can study each category separately. In this way we will be able to model their topological behavior more precisely.

Table 4.2: The ratio of link types in WECC power grid

Line Capacity- MW	**GG**	**LL**	**CC**	**GL**	**GC**	**LC**
(0-400)	100 %	100 %	78 %	95 %	91 %	92 %
(400-800)	0	0	15 %	3 %	8 %	5 %
(800-1200)	0	0	4 %	1 %	1 %	2 %
(1200-2000)	0	0	3 %	1 %	0	1 %
Total number of links	10	2	8229	593	183	5359

Table 4.3: The ratio of link types in WECC power grid

Line Capacity -MW	GG	LL	CC	GL	GC	LC
(0-400)	100 %	100 %	65 %	83 %	73 %	80 %
(400-800)	0	0	20 %	9 %	15 %	17 %
(800-1200)	0	0	10 %	8 %	12 %	2 %
(1200-2000)	0	0	5 %	0	0	1%
Total number of links	2	4	2539	86	49	1775

The node degree of bus *i* in a grid equals the total number of branches it connects and can be obtained from the *i*th diagonal entry of the Laplacian Matrix, i.e., $k_i = L(i,i)$. From node degree and topology of the grid it is possible to have a unique measure for the connectivity of each link in the grid, called *Average Link Degree (ALD),* that is:

$$\bar{k}_m = \frac{k_{(m,i)} + k_{(m,j)}}{2} \qquad for\ m = 1\ to\ M \tag{4.3}$$

where $k_{m,i}$ and $k_{m,j}$ represent the node degree for two connected buses, and *M* is the total number of links.

The important thing to note here is that since the definition (4.3) includes the node degree for both connected buses, we need to represent two link types, different from what we presented at the beginning of this section. The first type, called *Internal Link*, is considered a connection link which has some neighbors from both sides, and the second type, called *Boundary Link*, is directly connected to the load/generation buses. Internal and boundary links can be fully recognized with respect to the grid topology (see Figure 4.2)

Figure 4.3 displays the distribution function of ALD for internal links that results from the numerical simulation of ERCOT and WECC systems. For the sake of comparison, plots for various line capacity intervals are shown in a single figure. Figure 4.3 shows an explicit trend for ALD in

which as we move toward intervals with higher capacities, we obtain higher fitting parameters (μ, σ). Another even more interesting discovery is about the type of four distributions. For line capacities between 0 to 400 MW (interval 1) the ALD follows the Generalized Extreme Value distribution. However, other three categories follow the Burr distribution $\beta(x|c,k) = ck\,(x^{c-1}/(1+x^c)^{k+1})$ where c and k are fitting parameters. From the obtained statistics we can see that given a realistic power grid the ALDs have heavy-tailed distributions.

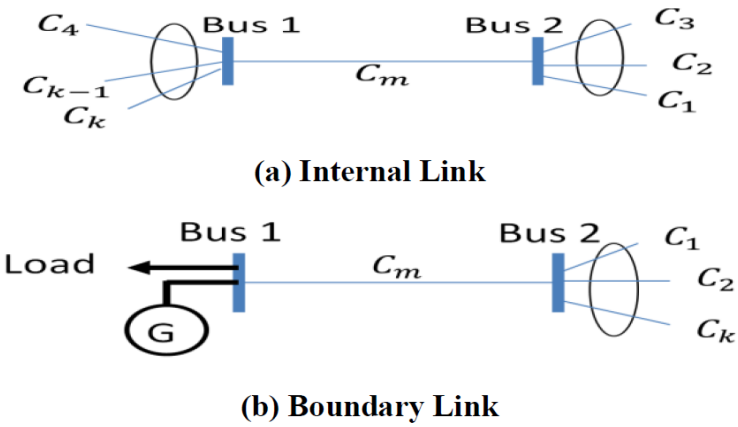


Figure 4.2: (a) Internal link including several neighbors from both sides and (b) Boundary link directly connected to generation/load buses.

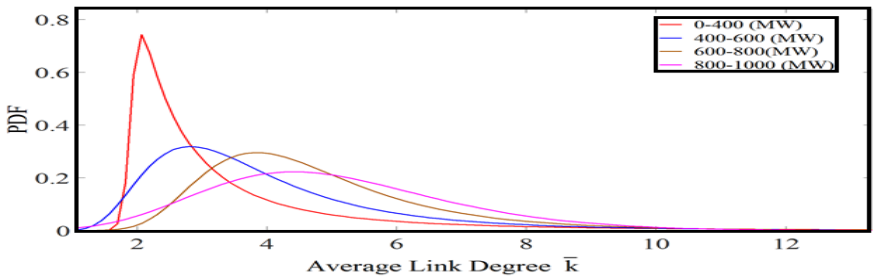


Figure 4.3: Average link degree distribution for the internal links resulted from ERCOT and WECC systems.

What presented above shows the distribution of ALD for internal links. However, one can easily consider another strategy to represent the statistical behavior of boundary links. To address this issue, we propose a new index, τ, called Neighboring Capacity Ratio (NCR) . For a given transmission line m, index τ is defined as the ratio of the average neighbor line capacities $\overline{C}$ to the line capacity C_m. The proposed index is presented as follows:

$$\tau_m = \frac{\overline{C}_m}{C_m} \tag{4.4}$$

$$\overline{C}_m = \frac{C_{1,m}+C_{2,m}+\cdots+C_{k,m}}{k} \tag{4.5}$$

where $c_{k,m}$ is the capacity of *k th* neighbor line (see Figure 4.2.b). Figure 4.4 plots the distribution function of τ for ERCOT and WECC power grids. The numerical results show that for boundary links, not only the proposed index follows Burr distribution, but also it is limited around $\tau = 2$ and no greater than $\tau \cong 5$. What should be noted here is the proposed index τ can be applied to the internal links as well. Figure 4.5 illustrates the distribution of τ for the internal links. As we can see for both ERCOT and WECC power grids, the fitted curves follow the Bur distribution (like what we see for boundary links). Figure 4.5 also indicates that the deviation of τ is roughly greater than $\tau = 2$ and it falls into a very restricted range. Table 13 shows the best fitting distribution parameters, and the mean and standard deviation values for both ERCOT and WECC systems.

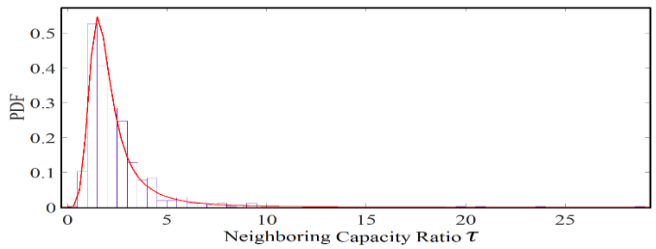
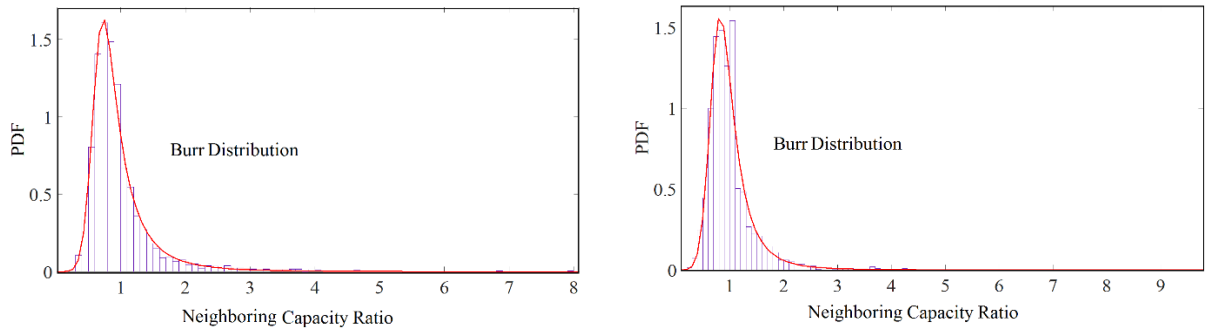


Figure 4.4: Neighboring capacity ratio distribution for the boundary links resulted from ERCOT and WECC systems.



(a)WECC (b) ERCOT

Figure 4.5: Neighboring capacity ratio distribution for the internal links resulted from (a) WECC and (b) ERCOT systems.

Table 4.4: Distribution fitting for neighboring capacity ratio

			Burr Distribution	
	Mean	**Std**	**c**	**k**
WECC	0.99	0.58	6.8	0.43
ERCOT	1.011	0.48	6.4	0.42

Examining and interpreting structural and electrical properties of power systems enable the development of an appropriate synthetic modeling that could be utilized to generate power grid test cases with accurate grid topology and electric parameters. Our statistical results show that the capacity of transmission lines follows a well-known distribution and it can be fully defined by mathematical definition. The obtained distribution provides a pattern to generate reasonable capacity line values for a given synthetic power grid. Also, with respect to the proposed measure, *Neighboring Capacity Ratio (NCR)*, our experiments reveals some useful and interesting relations presented between the transmission capacity of a specific line and that of neighboring lines in the grid. Generally speaking this ratio follows a Burr distribution with the mean value around 1.0. Our statistical analysis on real-world power grid provides potential insights to propose a practical

method to find the best transmission line capacity assignment in random topology power grid modeling.

4.4 Statistical-based algorithm to assign maximum capacity to transmission line

The last section of this chapter introduces a statistical –based approach to generate a set of transmission line capacities and assign them to the transmission lines. To accomplish this goal we first investigate the scaling function of total transmission line capacity in a grid versus its network size. Then we study the possible relationship between transmission line capacity F_l^{max} and the real-time flow of power F_l through the transmission lines during the normal operation of realistic power grids.

Figure 4.6 depicts the total transmission capacity of power grids, as defined in (3.3), versus the network size for each sample grid in Table 3.1, where the solid line represents an approximate fitting curve of the observed scaling property. Note that for the purpose of simplicity, the logarithm in this section is with base 10. From Figure 4.6 we can see that when the network size is small, the total transmission line capacity in a grid tend to has small value. However, as the network size becomes larger, the total backbone transmission capacity tends to grow as a power function of network size, $F_{Max}^{tot}(n) \propto n^{0.9059}$.

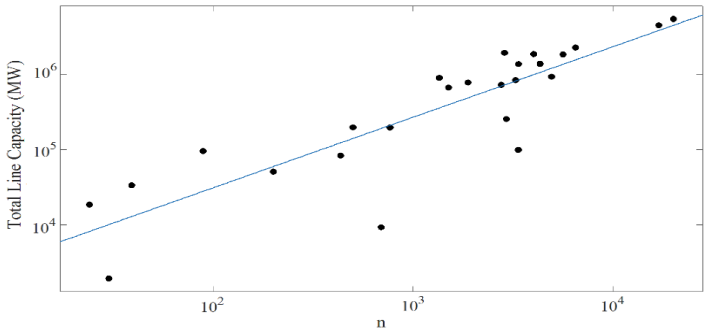


Figure 4.6: The scaling property of the total backbone transmission capacity in realistic power grids.

Figure 4.7 shows the empirical PDF marginal capacity $\beta_l = F_l^{max} - F_l/F_l^{max}, l = 1, \dots, M$, in the WECC-16994 bus system. This figure implies that in a typical power grid majority of transmission lines tend to have a big marginal capacity. Figure 4.7 also illustrates the distribution of marginal capacity β. As we can see in realistic power grids, the fitted curve follows the generalized extreme value distribution

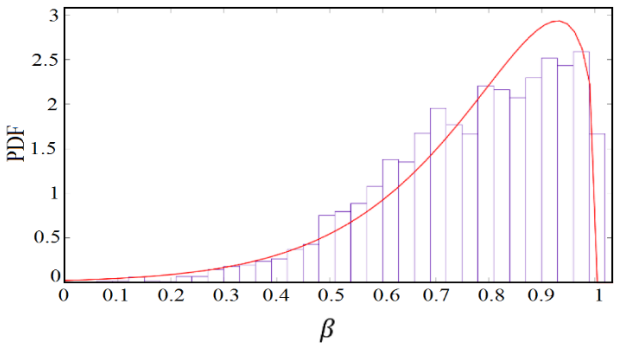


Figure 4.7: The Empirical PDF of capacity margin in WECC-16994 bus system

The above mentioned results can provide us useful guidelines to determine the transmission line capacities in a synthetic power grid with respect to their short-term power flows. The obtained empirical distribution will help us to propose a statistics-based algorithm to improve the transmission line setting in existing synthetic power grid modeling, namely, to assign the best set

of transmission line capacities which are consistent with that of realistic grids. Briefly, this method is composed of two steps as follows:

The statistics collected from the date of a number of realistic grids also indicate that there exists a considerable correlation between the capacity margin of a transmission line and its short-term power flow with a Pearson coefficient of -0.5745. Figure 4.8 displays the scatter plot of normalized short-term power flow and capacity margin which can be further used to generate the 2-D empirical PDF of some sample grids like WECC-16994 buses system.

By averaging the statistics of available realistic grid data, we may extract an empirical 2-dimensional probabilistic density function (PDF) for the normalized short-term power flow and capacity margin (see Figure 4.9). Based on the 2-D empirical PDF over the obtained uneven grid division a two-dimensional probability distribution table shown in Table 4.5 can be formulated to enable an algorithm to assign the generated capacity margin to each transmission line in a grid according to its normalized short-term power flow.

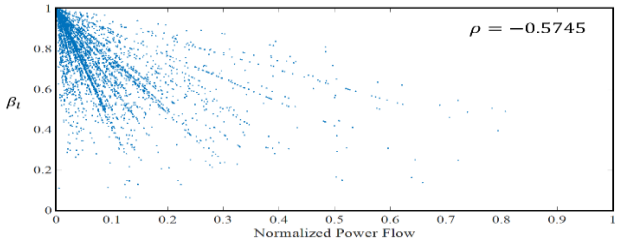


Figure 4.8 Scatter plot for capacity margin versus normalized short-term power flow in WECC-16994 bus system

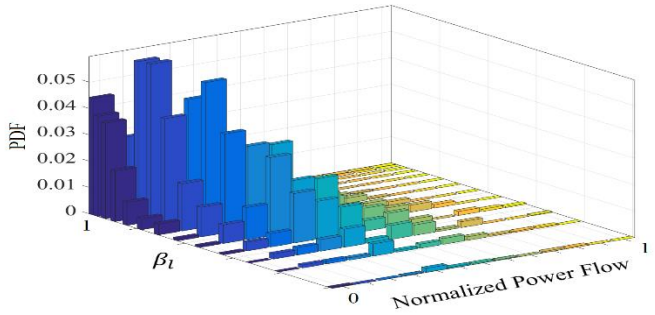


Figure 4.9: The 2-D empirical PDF of capacity margin versus normalized short-term power flow in WECC-16994 bus system

Table 4.5: Probability analysis of capacity margin and normalized short-term power flow in WECC-16994 bus system

		$\bar{F}_l$													Marginal Prob
		0.00 / 0.01	0.01 / 0.03	0.03 / 0.06	0.06 / 0.1	0.1 / 0.15	0.15 / 0.21	0.21 / 0.28	0.28 / 0.36	0.36 / 0.45	0.45 / 0.55	0.55 / 0.66	0.66 / 0.78	0.78 / 1.00	
β_l	1.00 / 0.99	0.044	0.001	0	0	0	0	0	0	0	0	0	0	0	0.441
	0.99 / 0.97	0.038	0.011	0.001	0	0	0	0	0	0	0	0	0	0	0.050
	0.97 / 0.94	0.036	0.029	0.008	0	0	0	0	0	0	0	0	0	0	0.073
	0.94 / 0.90	0.019	0.059	0.011	0.006	0.001	0	0	0	0	0	0	0	0	0.099
	0.90 / 0.85	0.008	0.059	0.027	0.008	0.003	0.002	0.001	0	0	0	0	0	0	0.112
	0.85 / 0.79	0.004	0.040	0.046	0.020	0.006	0.002	0.004	0	0	0	0	0	0	0.126
	0.79 / 0.72	0.004	0.017	0.054	0.024	0.012	0.004	0.001	0.003	0.001	0	0	0	0	0.126
	0.72 / 0.64	0	0.011	0.037	0.031	0.030	0.011	0.008	0.002	0.002	0.001	0	0	0	0.137
	0.64 / 0.55	0.001	0.007	0.012	0.029	0.019	0.018	0.005	0.004	0.003	0.002	0.001	0	0	0.104
	0.55 / 0.45	0.001	0.003	0.005	0.018	0.014	0.010	0.003	0.005	0.001	0	0.001	0	0.001	0.066
	0.45 / 0.34	0	0.001	0.002	0.004	0.006	0.001	0.005	0.003	0	0.002	0	0	0	0.024
	0.34 / 0.22	0	0.001	0.001	0	0.004	0	0.001	0.002	0.001	0	0	0	0	0.010
	0.22 / 0.00	0	0	0	0.	0.001	0.001	0.001	0	0	0	0	0	0	0.003
Marginal Prob		0.160	0.240	0.200	0.140	0.090	0.050	0.020	0.020	0.010	0.005	0.002	0	0.001	1.000

Step 1: In this step a number *m* of marginal capacities are generated from the specific distribution presented by figure 4.7 and an algorithm will be developed to assign the best capacity margin to each transmission line with respect to the statistical pattern presented in Table 4.5. And then the maximum transmission capacity of each line can be calculated by equation (4.6) with respect to the assigned marginal capacity and the real-time power flow resulted from DC power flow calculation [61-62].

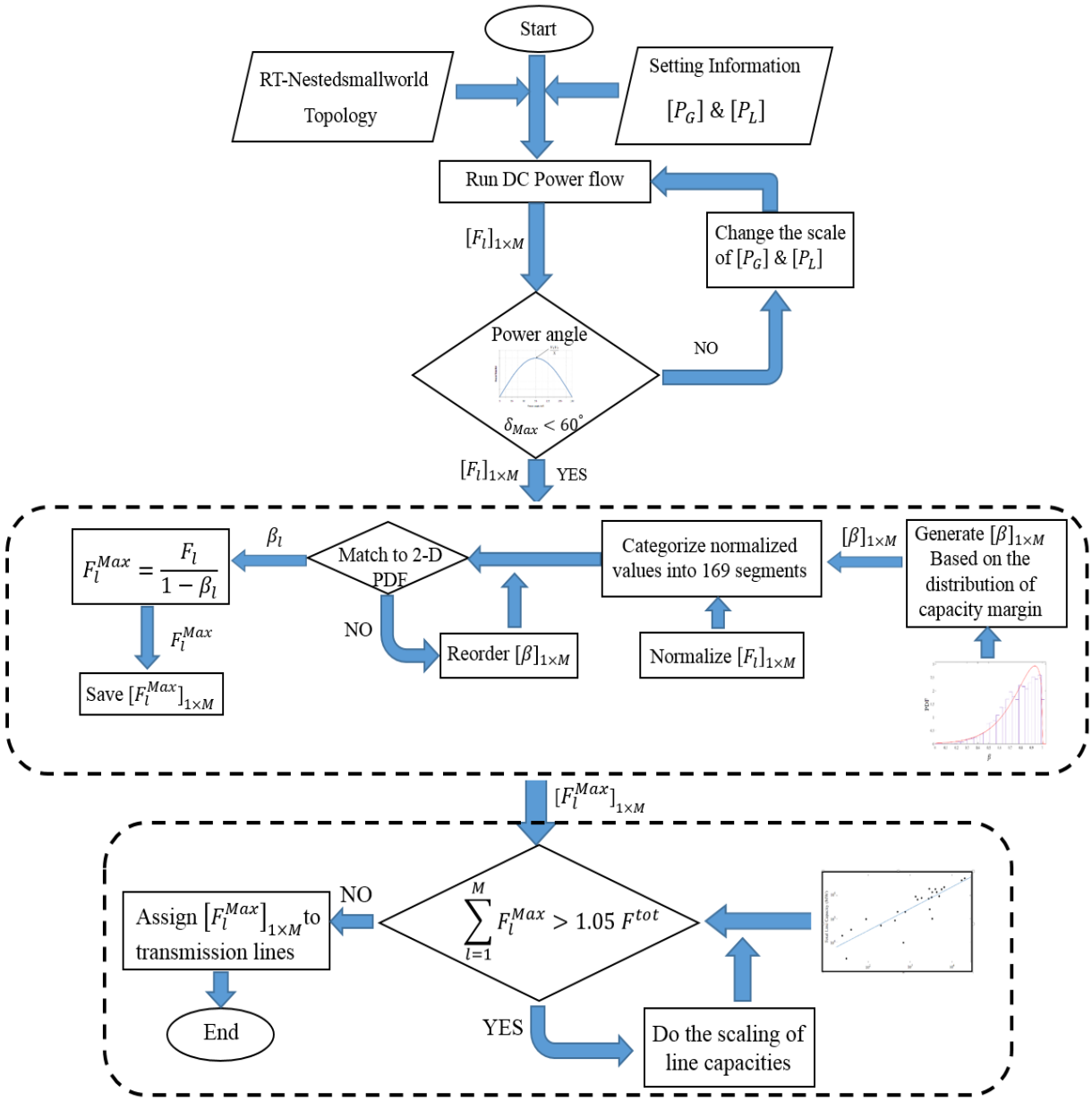


Figure 4.9: Flowchart of the proposed algorithm for transmission line capacity assignment in synthetic power grid modeling

$$F_l^{Max} = \frac{F_l}{1-\beta_l} \tag{4.6}$$

The real-time active power transfer across the line can is proportional to $\sin\delta$ where δ is called the power angle, which is the phase difference between the voltages on sending and receiving

buses. What should be noted here is in order to have a secure system the maximum power angle should be less than 60 degrees. Hence, we have to do the scaling of generation dispatches and loads if $\delta^{max} > 60°$, to make all power angles remain the secure range.

Step 2: Do the scaling of transmission capacities if $\sum_{l=1}^{M} F_l^{max} > 1.05F^{tot}$ to make the aggregate transmission capacity remain the range specified by $F^{tot}(n)$. And the scaling function is given as:

$$[F_l^{max}]'_{1\times M} = [P_{g_n}^{Max}]_{1\times M} \times \frac{F^{tot}}{\sum_{l=1}^{M} F_l^{max}} \quad (4.7)$$

The flowchart of the proposed algorithm is depicted in Figure 4.9.

Conclusion

In this chapter, we investigate the relationship between transmission line capacity and some network topology metric and a new proposed electrical index called "neighboring capacity ratio". It is found that the capacity of transmission lines follows a well-known distribution and it can be fully defined by mathematical definition. The obtained results show that the issue of transfer capacity assignment not only emerges as an electrical optimization concern, but some topological metrics must be considered to find the best line capacity assignment that is consistent with what is manifested in real-world grid. These results are used to develop a new methodology to appropriately characterize the line capacity assignment and improve the synthetic power grid modeling.

Chapter 5

Synthetic grid modeling integration

In synthetic power grid research area a number of models have been proposed based on the observed statistical characteristics of the grids. All these models provide useful perspectives of power grid characteristics. However, the topology of the generated power grids fails to accurately or fully represent a realistic power system. We note that power grid networks are much more than a graph topology. In order to facilitate numerical simulations for grid controls and operations, one also needs to include realistic electrical parameter settings such as transmission line capacities, the generation/ load settings. This thesis seek an effective way to improve the accuracy of existing synthetic power grid models, by providing five statistical-based algorithms. The proposed algorithms will allow us to recognize the best set of bus type assignment and the most accurate setting for generation, load and transmission line capacities.

Figure 5.1 shows the flowchart of the proposed five algorithm presented in this thesis that could be utilized to generate synthetic power grid test cases with accurate grid topology and electric parameters

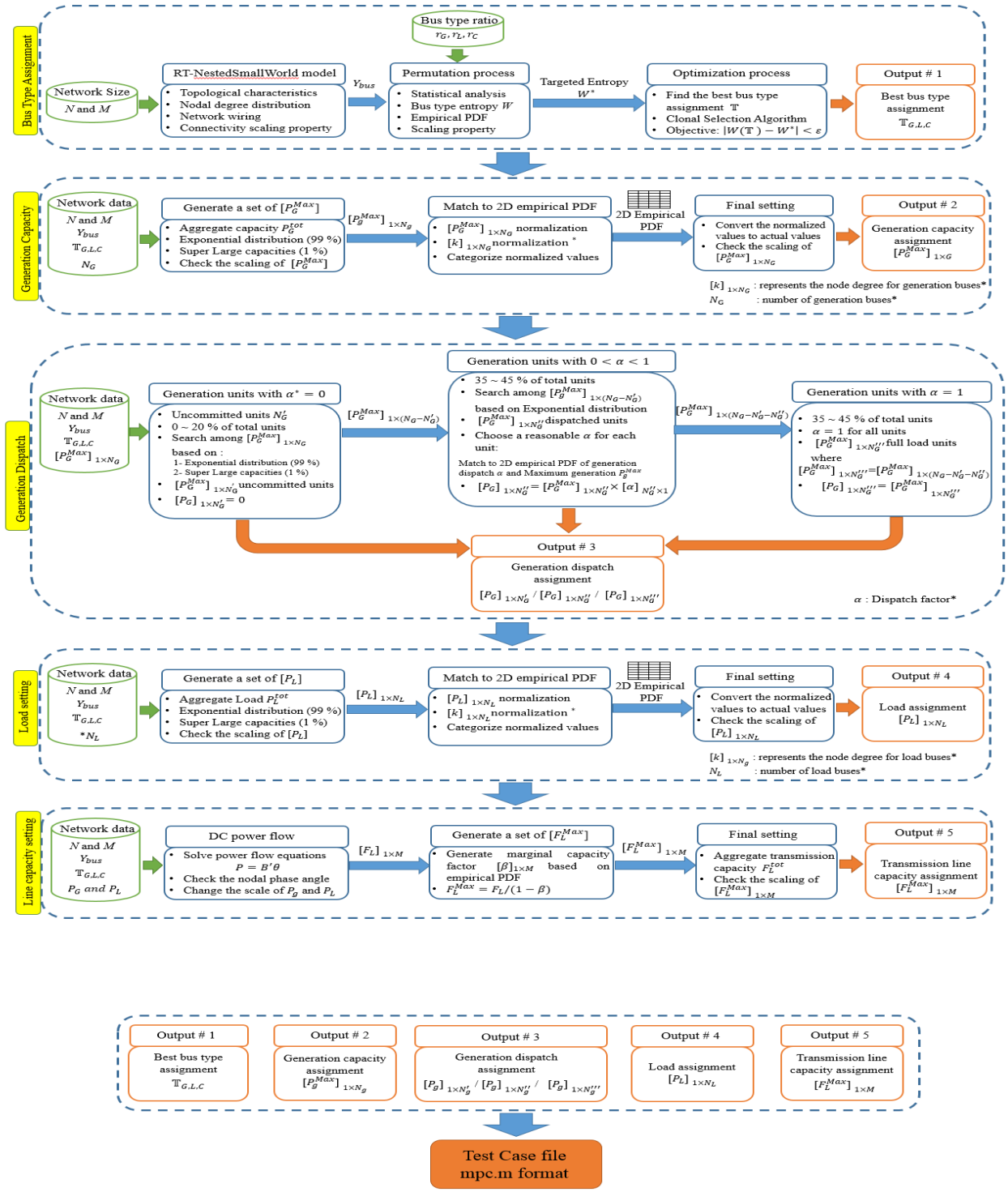


Figure 5.1: Summary of the proposed five algorithm to generate synthetic power grid test cases with accurate grid topology and electric parameters

Chapter 6

Thesis Summary

Several emerging issues, including the resiliency of electric power delivery during extreme weather events, expanding use of distributed generation, the rapid growth of renewable generation and the economic benefits of improved grid efficiency and flexibility, are challenging the way electricity is delivered from suppliers to consumers. This grid of the future requires advances in transmission and distribution system management with algorithms to control and optimize how power is transmitted and distributed on the grid. However, the development of these systems has been hindered because the research community lacks high-fidelity, public, large-scale power system models that realistically represent current and evolving grid characteristics. Due to security and privacy concerns, much of the real data needed to test and validate new tools and techniques is restricted. To help drive additional innovation in the electric power industry, there is a need for grid models that mimic the characteristics of the actual grid, but do not disclose sensitive information.

These models can be used to produce a sufficiently large number of power grid test cases with scalable network size featuring the same kind of small-world topology. However, in the existing models the approaches to address some electrical and topological settings such as (1) bus type

assignment, (2) generation and load settings, and (3) transmission line capacity assignments, are not sufficient enough to apply to realistic simulations.

To address this challenges, the first part of this thesis proposes a statistical methodology to solve the bus type assignment problem. In this part we have defined a novel numerical measure, *the Bus Type Entropy*, to characterize the bus type assignment of realistic power grids. Our proposed measure incorporates both bus type ratios and the link type ratios. Therefore it utilizes the correlated characteristics of realistic grids bus type assignments. This new measure will allow us to recognize the specific set of bus type assignments, either directly extracted from the realistic grids or formulated but consistent with a realistic grid, in the spectrum of random ones generated from permutation. Therefore some useful guidance will be provided to design an optimal algorithm to improve existing synthetic power grid modeling. We examine the performance of the proposed methodology on a set of realistic and synthetic power grids based on the empirical probability density function analysis. Numerical results obtained from all case studies verify the effectiveness of the proposed method to find the best set of bus type assignment.

In the second part of this thesis we examine the statistical features of generation capacities and loads in realistic power grids in terms of aggregate generation\load capacity, distribution of individual capacities, and their non-trivial correlation with nodal degrees. Our study on the statistical distribution of generation\load capacities shows that more than 99% of the generation units follow an exponential distribution with about 1% having extremely large capacities falling out of the normal range defined by the expected exponential distribution. Based on the obtained results presented in this part we came to the conclusion that there exists non-trivial correlations between the total number of branches connecting a bus (say node degree k) and the total generation attached to the bus. Based on the above results, we develop an algorithm to generate a statistically

correct random set of generation capacities and load, and then assign them to each generation and load bus in a grid. We then propose a statistical approach to determine the generation dispatch at each generation bus according to its generation capacity and the statistic of dispatch ratios.

The last part of this thesis mainly focuses on the statistical analysis of transmission line capacities. This part addresses the research question of whether there exists a relationship between the transmission line capacities and topological characteristics of power grids. We utilize the realistic power grid data sets to extract a model for the transmission line capacity in the synthetic high voltage grids. Motivated by the need for accurate line capacity assignment and in order to improve the synthetic power grid modeling, we performed a set of statistical experiments on realistic power grids and we found that (a) the capacity of transmission lines follows a well-known distribution and it can be fully defined by mathematical definitions (b) the line capacities assignment in a given synthetic power grid model should be improved using obtained statistical results which is found consistent with what is observed in realistic power grids. The obtained results and discoveries are used to develop a new methodology to appropriately characterize the line capacity assignment and improve the synthetic power grid modeling.

Bibliography

[1] Wood and B. Wollenberg, Power system generation, operation and control. New York: Wiley, 1984.

[2] "Power systems test case archive" [Online]. Available: http://www.ee.washington.edu/research/pstca/.

[3] H. Oh, "A New network reduction methodology for power system planning studies," IEEE Trans. on Power Systems, vol. 25, pp. 677-684, 2010.

[4] D. Shi, "Power system network reduction for engineering and economic analysis,"Ph.D. dissertation, Arizona State University, 2012.

[5] M. Rahnamay-Naeini, Z. Wang, N. Ghani, A. Mammoli and M. M. Hayat,"Stochastic analysis of cascading-failure dynamics in power grids," IEEE Trans. on Power Systems, vol. 29, pp. 1767-1779, 2014.

[6] A. Papaemmanouil and G. Andersson, "On the reduction of large power system models for power market simulations," PSCC'17, Stockholm, Sweden, pp. 1308-1313, 2011.

[7] M.Reza, "Stability analysis of transmission system with high penetration of distributed generation" PhD thesis, 2006.

[8] Z. Wang, A. Scaglione, and R. J. Thomas, "A Markov-Transition Model for Cascading Failures in Power Grids," 2012 45th Hawaii International Conference on System Sciences, Maui, HI, pp. 2115-2124, 2012.

[9] O.P. Malik, "Evaluation of power systems into smarter network", Journal of control, Automation and Power system", vol.24, pp 139-147, 2013.

[10] J. Grainger, W. Grainger, “Power System Analysis”, McGraw-Hill Science/Engineering/Math, 1994.

[11] S. M. Fatemi, S. Abedi, G. B. Gharehpetian, S. H. Hosseinian and M. Abedi, "Introducing a Novel DC Power Flow Method With Reactive Power Considerations," in IEEE Transactions on Power Systems, vol. 30, pp. 3012-3023, 2015.

[12] S.H, Low“Convex relaxation of optimal power flow” IREP Symposium Bulk Power System Dynamics and Control, pp.1-6, 2013.

[13] Z. Wang, R. J. Thomas, "On Bus Type Assignments in Random Topology Power Grid Models," 48th Hawaii International Conference on System Sciences, Kauai, HI, pp. 2671-2679, 2015.

[14] D. J. Watts, S. H. Strogatz, “Collective dynamics of ’Small-World’ networks,” Nature, vol. 393, pp.393-440, 1998.

[15] Z. Wang, R. J. Thomas, A. Scaglione, “Generating random topology power grids,” Proc. 41st Annual Hawaii International Conference on System Sciences (HICSS-41), Big Island, , vol. 41, 2008.

[16] L. Fu, W. Huang, S. Xiao; Y. Li; S. Guo, “Vulnerability Assessment for Power Grid Based on Small-world Topological Model,” Power and Energy Engineering Conference (APPEEC), Asia-Pacific , pp.1-4, 2010.

[17] Z. Wang, A. Scaglione, and R. J. Thomas, “Generating Statistically Correct Random Topologies for Testing Smart Grid Communication and Control Networks”, IEEE Transactions on Smart Grid, vol. 1, pp.28- 39, 2010.

[18] S. Soltan and G. Zussman,"Generation of synthetic spatially embedded power grid networks", arXiv:1508.04447, 2015.

[19] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, "A methodology for the creation of geographically realistic synthetic power flow models", 2016 IEEE Power and Energy Conference at Illinois (PECI), pp. 16, 2016.

[20] G. A. Pagani, M. Aiello,"The power grid as a complex network: a survey", Physica A: Statistical Mechanics and its Applications, vol.392, pp. 2688-2700, 2013.

[21] Z. Wang, S.H. Elyas, and R. J. Thomas, "A novel measure to characterize bus type assignments of realistic power grids", PowerTech 2015, Eindhoven, Netherlands, 2015.

[22] P. Hines, S. Blumsack, E. Cotilla Sanchez, and C. Barrows, "The Topological and Electrical Structure of Power Grids," Proc. 2010 43rd Hawaii International Conference on System Sciences (HICSS-43), 2010.

[23] E. Cotilla Sanchez, P. Hines, C. Barrows, and S. Blumsack, "Comparing the Topological and Electrical Structure of the North American Electric Power Infrastructure," IEEE Systems Journal, vol. 6, pp.616-626, 2012.

[24] R. Albert and A. Barab_asi, "Statistical mechanics of complex networks," Reviews of Modern Physics, vol. 74, pp.47-97. 2002.

[25] P. Erdos, and A. Renyi, "On random graphs. I.," Publicationes Mathematicae, vol. 6, pp.290-297, 1959.

[26] W. J. Reed, and M. Jorgensen, “The double pareto-lognormal distribution: a new parametric model for size distributions,” Communications in statistics-Theory and methods, vol 33, 2004.

[27] K. Rahimi and B. Chowdhury, "A hybrid approach to improve the resiliency of the power distribution system," 2014 North American Power Symposium (NAPS), Pullman, WA, pp. 1-6, 2014.

[28] A. Parchure, S. J. Tyler, M. A. Peskin, K. Rahimi, R. P. Broadwater and M. Dilek, "Investigating PV Generation Induced Voltage Volatility for Customers Sharing a Distribution Service Transformer," in IEEE Transactions on Industry Applications, vol. 53, pp. 71-79, 2017.

[29] K. Rahimi, S. Mohajeryami and A. Majzoobi, "Effects of photovoltaic systems on power quality," 2016 North American Power Symposium (NAPS), Denver, CO, pp. 1-6, 2016.

[30] D. Devaraj and B. Yegnanarayana, "Genetic-algorithm-based optimal power flow for security enhancement," in IEE Proceedings - Generation, Transmission and Distribution, vol. 152, pp. 899-905, 2005.

[31] M. Todorovski and D. Rajicic, "An initialization procedure in solving optimal power flow by genetic algorithm," in IEEE Transactions on Power Systems, vol. 21, pp. 480-487, 2006.

[32] J. Condren, T. W. Gedra and P. Damrongkulkamjorn, "Optimal power flow with expected security costs," in IEEE Transactions on Power Systems, vol. 21, pp. 541-547, May 2006.

[33] M. Rosas-Casals, S. Valverde, and R. V. Sol, “Topological Vulnerability of the European Power Grid under Errors and Attacks”, Int. J. Bifurcation Chaos 17, pp.2465-2475, 2007.

[34] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, “Critical points and transitions in an electric power transmission model for cascading failure blackouts,” Chaos, vol. 12, pp.985-994, 2002.

[35] M. Parashar, J. S. Thorp and C. E. Seyler, "Continuum modeling of electromechanical dynamics in large-scale power systems," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 51, pp. 1848-1858, 2004.

[36] E.N. Gilbert,” Random graphs” The Annals of Mathematical Statistics, Vol.30, pp. 1141 – 1144. 1959.

[37] A. Barabási and R. Albert, “Emergence of Scaling in Random Networks” Science 286, pp. 509 – 512, 1999.

[38] C. Herrmann et al,” Connectivity distribution of spatial networks” Physical Review E 68, pp. 1 – 6, 2003.

[39] B.M. Waxman, IEEE Journal on Selected Areas in Communications 6.9, pp.1617 – 1622, 1988.

[40] M. E. J. Newman. “Random graphs as models of networks” John Wiley, New York, NY, 2002.

[41] Z. Wang, A. Scaglione, and R. J. Thomas, “On modeling random topology power grids for testing decentralized network control strategies,” presented at the 1st IFAC Workshop Estimation and Control Netw. Syst. (NecSys’09), Venice, Italy, 2009.

[42] Z. Wang, A. Scaglione, and R. J. Thomas, “The node degree distribution in power grid and its topology robustness under random and selective node removals,” presented at the 1st

IEEE Int. Workshop Smart Grid Commun. (ICC'10SGComm), Cape Town, South Africa, 2010.

[43] "NYISO" [Online]. Available: http://www.nyiso.com/public/index.jsp.

[44] "ERCOT" [Online]. Available: http://www.ercot.com/

[45] "WECC" [Online]. Available: https://www.wecc.biz/Pages/home.aspx

[46] D. E. Whitney , D. Alderson, "Are technological and social networks really different," Proc. 6th International Conference on Complex Systems (ICCS06), Boston, MA, 2006.

[47] M. Newman, "The structure and function of complex networks," SIAM Review, vol. 45, pp.167-256, 2003.

[48] J. L. Rodgers, and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," The American Statistician, volume 42(1), pp.59-66. 1988.

[49] L. Castro, F. Zuben, "Learning and optimization using the clonal selection principle", IEEE Transactions Evolutionary Computation, vol.6, pp:239-251, 2002.

[50] L. S. Batista, D. B. Oliveira, F. G. Guimaraes, E. J. Silva and J. A. Ramirez, "Dynamic Multiobjective Clonal Selection Algorithm for Engineering Design," in IEEE Transactions on Magnetics, vol. 46, no. 8, pp. 3033-3036, 2010.

[51] S. B. Nejad, S. H. Elyas, A. Khamseh, I. N. Moghaddam and M. Karrari, "Hybrid CLONAL selection algorithm with PSO for valve-point Economic load Dispatch," 2012 16th IEEE Mediterranean Electrotechnical Conference, pp. 1147-1150, 2012.

[52] M. Jafari, A. Mohammad Shahri and S. H. Elyas, "Optimal tuning of Brain Emotional Learning Based Intelligent Controller using Clonal Selection Algorithm," ICCKE 2013, Mashhad, pp. 30-34, 2013.

[53] S.H Elyas, Z. Wang, "Improved synthetic power grid modeling with correlated bus type assignments" IEEE Transactions on power system, accepted 2016, available: http://ieeexplore.ieee.org/stamp/stamp.jsp? arnumber=7763878.

[54] E. Schweitzer, A. Scaglione, R. Thomas, T. Overbye, "Analysis of the coupling between power system topology and operating condition for synthetic test case validation" CIGRE 2016 grid of future symposium, Philadelphia, PA, 2016.

[55] A. Birchfield, T. Xu, K. Gegner, K. Shetye, T. Ovrbye, "Grid structural characteristics as validation criteria for synthetic network" IEEE Transactions on power systems, 2016.

[56] T. Xu, A. Birchfield, K. Gegner, K. Shetye, T. Overbye, "Application of large-scale synthetic power system models for energy economic studies," HICSS- 50, Waikoloa, HI, 2017.

[57] H. E. Green, "The Numerical Solution of Some Important Transmission-Line Problems," in IEEE Transactions on Microwave Theory and Techniques, vol. 13, pp. 676-692, 1965.

[58] D. Shirmohammadi, P. R. Gribik, E. T. K. Law, J. H. Malinowski and R. E. O'Donnell, "Evaluation of transmission network capacity use for wheeling transactions," in *IEEE Transactions on Power Systems*, vol. 4, no. 4, pp. 1405-1413, 1989.

[59] J. C. Villumsen, G. Brønmo and A. B. Philpott, "Line capacity expansion and transmission switching in power systems with large-scale wind power," in *IEEE Transactions on Power Systems*, vol. 28, pp. 731-739, 2013.

[60] Yaqing Liu, M. Zitnik and R. Thottappillil, "An improved transmission-line model of grounding system," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 43, pp. 348-355, 2001.

[61] B. Stott, J. Jardim and O. Alsac, "DC Power Flow Revisited," in *IEEE Transactions on Power Systems*, vol. 24, pp. 1290-1300, 2009.

[62] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans and W. L. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," *The 8th IEE International Conference on AC and DC Power Transmission*, London, UK, pp. 58-62, 2006.

Appendix A

List of Publications

[1] S. H. Elyas, Z. Wang, "Improved Synthetic Power Grid Modeling with Correlated Bus Type Assignments", IEEE Transaction on Power Systems, 2016. (available in : http://ieeexplore.ieee.org/abstract/document/7763878/) (Top journal in the area of power system research).

[2] S. H. Elyas, Z. Wang, "Statistical analysis of transmission line capacities in electric power grids" ISGT, 2016.

[3] Z. Wang, S. H. Elyas, R. J. Thomas, "A novel measure to characterize bus type assignments of realistic power grids", Power Tech, 2015.

[4] S. H. Elyas, Z. Wang, "A multi- objective optimization algorithm for bus type assignments in random topology power grid model" HICSS 49, 2016.

[5] Z. Wang, S. H. Elyas, R. J. Thomas, "Generating synthetic electric power system data with accurate electric topology and parameters" UPEC, 2016.

[6] S. H. Elyas, Z. Wang, R. J. Thomas, "On the Statistical Settings of Generation Capacities and Dispatch in a Synthetic Grid Modeling" IREP, 2017

[7] Z. Wang, S. H. Elyas, " On the Scaling Property of Power Grids", 50th Hawaii International Conference on System Sciences (HICSS), 2017.