2015 48th Hawaii International Conference on System Sciences

On Bus Type Assignments in Random Topology Power Grid Models

Zhifang Wang
Electrical and Computer Engineering
Virginia Commonwealth University
Richmond, VA, USA
zfwang@vcu.edu

Robert J. Thomas
Electrical and Computer Engineering
Cornell University
Ithaca, NY, USA
rjt1@cornell.edu

***Abstract*—In order to demonstrate and test new concepts and methods for the future grids, power engineers and researchers need appropriate randomly generated grid network topologies for Monte Carlo experiments. If the random 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. Our previous work [1] proposed a random topology power grid model, called *RT-nested-smallworld*, based on the findings from a comprehensive study of the topology and electrical properties of a number of realistic grids. The proposed model can be utilized to generate a large number of power grid test cases with scalable network size featuring the same small-world topology and electrical characteristics found from realistic power grids.**

On the other hand, we know that dynamics of a grid not only depend on its electrical topology but also on the generation and load settings; and the latter closely relates with an accurate bus type assignment of the grid. Generally speaking, the buses in a power grid test case can be divided into three categories: the generation buses (G), the load buses (L), and the connection buses (C). In [1] our proposed model simply adopts random assignment of bus types in a resulting grid topology, according to the three bus types' ratios.

In this paper we examined the correlation between the three bus types of G/L/C and some network topology metrics such as node degree distribution and clustering coefficient. We also investigated the impacts of different bus type assignments on the grid vulnerability to cascading failures using IEEE 300 bus system as an example. We found that (a) the node degree distribution and clustering characteristic are different for different type of buses (G/L/C) in a realistic grid; (b) the changes in bus type assignment in a grid may cause big differences in system dynamics; and (c) the random assignment of bus types in a random topology power grid model should be improved by using a more accurate assignment which is consistent with that of realistic grids.

***Keywords*-Electric Power grid, random topology, graph network.**

I. Introduction

Electric power grid is one of the largest interconnected machines on the earth, whose dynamics include vast complexities resulting from the large number of interconnected grid components [2]. The transients occur in disparate time frames and arise from the nonlinear and time-varying nature of transmission flow and disturbance propagation. What is more, as smart grid innovation advances, higher penetration of renewable generation, electric vehicles, and demand side participation will surely introduce even more complexity to the grid dynamic model. It is a common practice for power engineers and researchers to trim a substantial amount of the complexity from the complete model and instead utilize various sets of approximations and assumptions to develop new concepts, methods, analysis, and control schemes. And in order to in some sense demonstrate and prove the new concept and method numerical simulations and experiments on some standard power grid test case has become an indispensable critical step.

In the past power engineers and researchers mainly depended on a small number of historical test systems such as the IEEE 30 or 300 bus systems [3]-[6]. Even if they could utilize some real-world and/or synthesized grid data for test [7]-[11], the verification and evaluation results, obtained from the limited number of test cases, could not be deemed as adequate or convincing in the sense that the same concept or method, if extended to other instance of grid networks, may not yield the same or comparable performance. We believe that future power engineers and researchers need appropriate randomly generated grid network topologies for Monte Carlo experiments to demonstrate and test new concepts and methods. If the random 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 such as the IEEE-30, 57, 118, or 300-bus systems [3] or some other existing grid models [12]-[15].

A number of researchers also noticed similar needs to generate scalable-size power grid test cases. Different power grid models were proposed based on observed statistical characteristics in the literature. For example, [12] proposed a Tree-topology power grid model to study power grid robustness and to detect critical points and transitions in the flow transmission to cause cascading failure blackouts. In [14] the authors used Ring-structured power grid topologies to study the pattern and speed of contingency or disturbance propagation. [15] first proposed statistically modeling the power grid as a small-world network in their work on random graphs. [16] gave a statistical model for power networks in an effort to grasp what class of communication network topologies need to match the underlying power networks,

1530-1605/15 $31.00 © 2015 IEEE
DOI 10.1109/HICSS.2015.322

so that the provision of the network control problem can be done efficiently. [17] used a small-world graph model to study the intrinsic spreading mechanism of the chain failure in a large-scale grid. All these models mentioned above provide useful perspectives to 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 correctly or fully reflect that of a realistic power system, especially its distinct sparse connectivity and scaling property versus the grid size. Besides, power grid network is more than just a topology: in order to facilitate numerical simulations of grid controls and operations, it also needs to include realistic electrical parameter settings such as line impedances, and generation and load settings.

In our previous work [1], we finished a comprehensive study on the topological and electrical characteristics of electric power grids including the nodal degree distribution, the clustering coefficient, the line impedance distribution, the graph spectral density, and the scaling property of the algebraic connectivity of the network, etc. Based on the observed characteristics, a random topology power grid models, called *RT-nested-smallworld*, has been proposed to generate a large number of power grid test cases with scalable network size featuring the same electrical and topology with realistic power grid features. The only shortcoming of this model lies in its random bus type assignments purely according to the different bus type ratios. 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 connecting generators,
- L the load buses supporting custom demands, and
- C the connection buses forming the transmission network.

In a typical power grid, the generation bus consists of 10$\sim$40% of all the grid buses, the load bus is 40$\sim$60%, and the connection bus 10$\sim$20%.

In this paper we examined the correlation between the three bus types of G/L/C and network topology metrics such as the node degree distribution and the clustering coefficients. We also investigated the impacts of different bus type assignments on the grid dynamics such as its vulnerability to cascading failures using IEEE 300 bus system as an example. We found that (a) the node degree distribution and clustering characteristic differ for the three types of buses G/L/C in a realistic grid; (b) changes in bus type assignment in a grid may cause big differences in system dynamics such as its vulnerability to cascading failures; and (c) the random assignment of bus types in a random topology power grid model should be improved and replaced with a more accurate assignment which is consistent with what is observed in realistic power grids.

The rest of the paper is organized as follows: Section II presents the system model for our analysis of a power grid network; Section III and IV examine the node degree distribution and the clustering characteristics respectively for different types of buses in a power grid; Section V utilizes IEEE 300 bus system as an example to study how the bus assignment affects the resulting grid's electrical characteristics in terms to cascading failure vulnerability; and Section VI concludes the paper.

II. System Model

The power network dynamics are controlled by its network admittance matrix Y and by power generation distribution and load settings. The generation and load settings, given the bus type assignments in the grid, can take relatively independent probabilistic models, and have been discussed in the paper [16]. The network admittance matrix

$$Y = A^T \Lambda^{-1}(l_l) A \tag{1}$$

can be formulated with two components, i.e. the line impedance vector z_l and the line-node incidence matrix A. For a grid network with n nodes and m transmission lines, its line-node incidence matrix $A := (A_{l,k})_{m\times n}$, arbitrarily oriented, is defined as: $A_{l,i} = 1$; $A_{l,j} = -1$, if the l^{th} link is from node i to node j and $A_{l,k} = 0, k \neq i, j$. And the Laplacian matrix L can be obtained as $L = A^T A$ with

$$L(i,j) = \begin{cases} -1, & \text{if there exists link } i-j, \quad \text{for } j \neq i \\ k, & \text{with } k = -\sum_{j\neq i} L(i,j), \quad \text{for } j = i \\ 0, & \text{otherewise,} \end{cases} \tag{2}$$

with $i, j = 1, 2, \cdots, n$. As indicated in [1] , the Laplacian of a grid network fully defines its topology and all the topology metrics can be derived from it.

[26] introduced a stochastic cascading failure model which incorporates the statistics of the generation and loads in a grid therefore to derive the statistics of the line flow process. For the tractability of the problem, the *DC power flow* approximation was utilized to characterize a power grid network, which is a standard approach widely used in optimizing flow dispatch and for assessing line overloads [27]. Consider a power grid transmission network with n nodes interconnected by m transmission lines, the network flow equation can be written as follows:

$$\begin{aligned} P(t) &= B'(t)\theta(t), \\ F(t) &= diag\,(y_l(t))\,A\theta(t) \end{aligned} \tag{3}$$

where $P(t)$ represents the vector of injected real power, $\theta(t)$ the phase angles, and $F(t)$ the flows on the lines. The matrix $B'(t)$ is defined as $B'(t) = A^T diag\,(y_l(t))\,A$, where $y_l(t) = s_l(t)/x_l$ with x_l the line reactance and $s_l(t)$ the line state; $s_l(t) = 0$ if line l is tripped, and $s_l(t) = 1$ otherwise; $diag\,(y_l(t))$ represents a diagonal matrix with entries of $\{y_l(t), \quad l = 1, 2, \cdots, m\}$. The vector of line states, $\mathbf{s}(t) = [s_1(t), s_2(t), \cdots, s_m(t)]^T$ with $s_l(t) \in \{0,\ 1\}$, is defined as the network state.

The operating condition of the grid, represented by the real power injection $P(t) = [G(t)^T \quad -L(t)^T]^T$, where $G(t)$ is the generation and $L(t)$ is the load portion, assumes to be a conditionally multivariate Gaussian random process, given the network state of $\mathbf{s}(t)$. The probability density function of $\{P(t)|\mathbf{s}(t)\}$ is, therefore, fully specified giving its conditional mean and covariance, denoted as follows:

$$\mu_P(t) = \begin{bmatrix} \mu_g(t) \\ -\mu_l(t) \end{bmatrix}, \; C_P(t,\tau) = \begin{bmatrix} \Sigma_{gg}(t,\tau) & \Sigma_{gl}(t,\tau) \\ \Sigma_{lg}(t,\tau) & \Sigma_{ll}(t,\tau) \end{bmatrix}. \tag{4}$$

The time dependence is due to the fact that the random process is intrinsically non-stationary, e.g., the generation settings in the power grid are adjusted periodically to balance the loads during the day. The covariance reflects the uncertainty in the load/generation settings coming from multiple sources such as the forecasting deviation, the measurement errors, and the volatility caused by demand response and renewable generation.

We can then compute the statistics of line flow process as

$$\begin{aligned} \mu_F(t) &= \sqrt{y_t}(\widetilde{A}_t^T)^\dagger \mu_P(t) \\ C_F(t,\tau) &= \sqrt{y_t}(\widetilde{A}_t^T)^\dagger C_P(t,\tau)(\widetilde{A}_t)^\dagger \sqrt{y_t}, \end{aligned} \tag{5}$$

with $\widetilde{A}_t = \sqrt{y_t}A = USV^T$ and $\sqrt{y_t} = diag\{\sqrt{y_l(t)}\}$.

III. Topological Characteristics of G/L/C Buses in a Power Grid

The statistics of power grid topology have been studied by many researchers, such as [15], [20], [23]. The topology metrics studied include some basic ones, such as network size n, the total number of links m, average nodal degree $\langle k\rangle$, average shortest path length in hops $\langle l\rangle$, and more complex ones, such as the ratio of nodes with larger nodal degrees than $\overline{k}$, $r\{k > \overline{k}\}$, and the Pearson coefficient [21] and the clustering coefficient [15]. [1] also examined the scaling property of network connectivity and recognized the special graph spectral density of power grids. As indicated in [1], all the topology metrics mentioned above can be derived from the grid Laplacian L. In this section we will focus on the average node degree, the clustering coefficient, and the node degree distribution since these metrics closely relates with the bus type assignment in a grid.

A. Average Node Degrees for $G/L/C$ Grid Buses

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

$$\langle k\rangle = \frac{1}{N}\sum_i L(i,i). \tag{6}$$

While the average node degree of a special type of buses can be defined as

$$\langle k\rangle_t = \frac{1}{N_t}\sum_i L_t(i,i), \qquad t = G,\, L, \text{ or } C, \tag{7}$$

Table I
Ratio of Bus Types in Real-World Power Networks

	(n,m)	$r_{G/L/C}(\%)$
IEEE-30	(30,41)	20/60/20
IEEE-57	(57,78)	12/62/26
IEEE-118	(118,179)	46/46/08
IEEE-300	(300, 409)	23/55/22
NYISO	(2935,6567)	33/44/23

where N_t represents the total number of t-type buses in a grid. The linear dependence between nodal degree and the node type can be evaluated using a correlation coefficient defined as

$$\rho(t,k_t) = \frac{\sum_{i=1}^N (t_i - E(t))(k_i - E(k))}{\sqrt{\sum_{i=1}^N (t_i - E(t))^2}\sqrt{\sum_{i=1}^N (k_i - E(k))^2}}. \tag{8}$$

Table I gives the network size and the percentage ratios of three types of buses in the IEEE model and the NYISO systems[1], from which one can see that the generation buses (G) may consist of 10-40% of total grid; the load buses (L) 40-60%; while the connection buses (C) 10-20%. Therefore, in a typical grid network, the generation buses has the widest ratio range; the load buses has the highest percentage; and the connection buses the lowest with the most narrow range of ratios.

Table II shows the average node degrees in the IEEE model grids, the WECC[2], and the NYISO systems based on selected metrics (6) and (7) and the correlation coefficient between node degree and bus types $\rho(t,k_t)$. From Table II one can clearly see the sparse connectivity of power grid networks since the average nodal degree does not scale as the network size increases. Instead it falls into a very restricted range. On the other hand, if examining the average node degree of the three type of buses in a grid separately, we see that they can be different from each other. In the IEEE 30, and 300 bus systems, the connection buses have the highest nodal degree. In the IEEE 57, and 118 bus systems, the generation buses are most densely connected. However, in the NYISO system, the load buses have the most connected. The last column of the table, $\rho(t,k_t)$, shows that there exist non-trivial correlation between the node degrees and the bus types in a typical power grid. A positive correlation coefficient, i.e., $\rho(t,k_t) > 0$ implies that the connection bus tends to have higher average nodal degree; while a negative

[1]For the reader's reference, the IEEE 30, 57, and 118 bus systems represent different parts of the American Electric Power System in the Midwestern US; the IEEE 300 bus system is synthesized from the New England power system. The WECC is the electrical power grid of the western United States and NYISO represents New York state bulk electricity grid.

[2]The average node degrees for each type of buses in the WECC system are not available here due to the lack of the bus type data.

Table II
CORRELATION BETWEEN NODE DEGREE AND BUS TYPES IN REAL-WORLD POWER NETWORKS

	$\langle k\rangle$	$\langle k\rangle_G$	$\langle k\rangle_L$	$\langle k\rangle_C$	$\rho(t,k_t)$
IEEE-30	2.73	2.00	2.61	3.83	0.4147
IEEE-57	2.74	3.86	2.54	2.67	-0.2343
IEEE-118	3.03	3.56	2.44	3.40	-0.2087
IEEE-300	2.73	1.96	2.88	3.15	0.2621
NYISO	4.47	4.57	5.01	3.33	-0.1030
WECC	2.67	-	-	-	-

Table III
BUS TYPES AND CLUSTERING COEFFICIENTS OF REAL-WORLD POWER NETWORKS AND RANDOM GRAPH NETWORKS

	$C(R)$	C_{all}	C_G	C_L	C_C	$\rho(t,C_t)$
IEEE-30	0.0943	0.2348	0.1944	0.2537	0.2183	0.0210
IEEE-57	0.0489	0.1222	0.1524	0.1352	0.0778	-0.1064
IEEE-118	0.0260	0.1651	0.1607	0.1969	0.0167	-0.0538
IEEE-300	0.0091	0.0856	0.1227	0.0895	0.0364	-0.1428
NYISO-2935	0.0015	0.2134	0.2693	0.2489	0.0688	-0.2382
WECC-4941	0.0005	0.0801	-	-	-	-

correlation coefficient $\rho(t,k_t) < 0$ that the generation or the load bus tends to be more densely connected.

B. Clustering Coefficient for $G/L/C$ Grid Buses

The *clustering coefficient* is an important characterizing measure to distinguish the *small world* topology of a power grid network. It assesses the degree to which the nodes in a topology tend to cluster together. A *small world* network usually has a clustering coefficient significantly higher than that of a random graph network, given the same or comparable network size and total number of edges. The random graph network mentioned here refers to the network model defined by Erd*ö*s-R*é*nyi (1959) , with n labeled nodes connected by m edges which are chosen with uniform randomness from the $n(n-1)/2$ possible edges [19].

The clustering coefficient of a grid is defined by Watts and Strogatz as the average of the clustering coefficient for each node [15]:

$$C = \frac{1}{N}\sum_{i=1}^{N} C_i. \tag{9}$$

with C_i being the node clustering coefficient as $C_i = \lambda_G(i)/\tau_G(i)$, where $\lambda_G(i)$ is the number of edges between the neighbors of node i and $\tau_G(i)$ the total number of edges that could possibly exist among the neighbors of node i. For undirected graphs, obviously $\tau_G(i) = \frac{1}{2}k_i(k_i-1)$ given k_i as the node degree. As pointed out in [18], the clustering coefficient for a random graph network theoretically equals the probability of randomly selecting links from all possible links. That is, $C(R) = \frac{m}{n(n-1)/2} = \frac{\langle k\rangle}{n-1}$.

Hence we can extend above definition to each type of $G/L/C$ buses to obtain the averaged clustering coefficient of the corresponding category. That is, given a bus type, the corresponding C_t is:

$$C_t = \frac{1}{N_t}\sum_{i=1}^{N_t} C_{t_i}, \qquad t = G,\ L,\ \text{or}\ C. \tag{10}$$

Table III shows the clustering coefficients of the IEEE power systems, the NYISO, the WECC systems[3], and the random graph Networks with same network size and same total number of links. The former is denoted as C_{all} and $C_{G/L/C}$, and the latter as $C(R)$. The relatively larger clustering coefficient of power networks than that of a random graph network, i.e., $C_{all} \gg C(R)$ was used in [15] as an indicator that power grids tend to assume a *small-world* topology.

On the other hand, if examining the clustering coefficients of the three types of grid buses separately, we see that they can be very different from each other. The last column of the table, $\rho(t,C_t)$, shows that there exist non-trivial correlation between the bus types and the clustering coefficient in a typical power grid. A positive correlation coefficient, i.e., $\rho(t,C_t) > 0$ implies that the connection bus tends to have higher average clustering coefficient; while a negative correlation coefficient $\rho(t,C_t) < 0$ that the generation or the load bus tends to be more densely clustered. It is worth noting that except the IEEE 30 bus system, all the other IEEE model systems and the NYISO system have $C_G \gg C_C$ and $C_L \gg C_C$, indicating that the generation and load buses in a typical power grid, tend to cluster together with the rest of the network much more tightly than the connection buses.

C. Nodal Degree Distribution for $G/L/C$ Grid Buses

In [1] we examined the empirical distribution of nodal degrees $\underline{k} = \text{diag}(L)$ in the available real-world power grids. The histogram probability mass function (PMF) of node degree is obtained as

$$p(k) = \frac{\sum_{i=1}^{N} \mathbf{1}_{k_i=k}}{N} \tag{11}$$

Fig. 1 shows the histogram PMF in log-scale for the nodal degrees of the NYISO system. The PMF curve approximates a straight line in the semi-logarithm plot (i.e., shown as $\log(p(k))$ vs. k), implying an exponential tail analogous to the Geometric distribution. However, it is also noticed that there exist a range of small degrees, that is, when $k \leq 3$, that the empirical PMF curve clearly deviates from that of a Geometric distribution. Although many researchers chose to use an exponential (or Geometric) distribution for power grid node degrees [22], we have provided evidence that the nodal degrees in a power grid can be better fitted

[3]The clustering coefficients for each type of buses in the WECC system are not available here due to the lack of the bus type data.

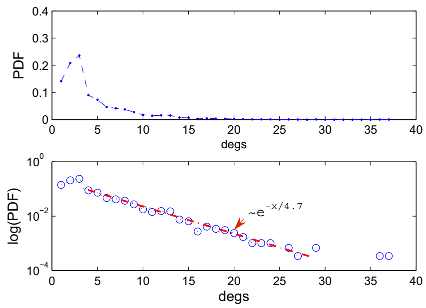


Figure 1. Empirical PDF of Nodal Degrees in Real-world Power Grids (NYISO)

by using a mixture distribution which comes from the sum of a truncated Geometric random variable and an irregular Discrete random variable. We also proposed a method to estimate the distribution parameters by analyzing the poles and zeros of the average probability generation function (PGF).

The PGF of a random variable X is defined as $G_X(z) = \sum_k \text{Pr}_{(x=k)} z^k$. Given a sample data set of X with the size of N, its PGF can also be estimated from the mean of z^X because $E(z^X) = \frac{1}{N}\sum_k n_{(x=k)} z^k \approx \sum_k \text{Pr}_{(x=k)} z^k$. where $n_{(x=k)}$ denotes the total number of the data items equaling to k. Due to $\lim_{N\to\infty} \frac{n_{(x=k)}}{N} = \text{Pr}_{(x=k)}$, we can have $E(z^X) \approx G_X(z)$ with a large enough data size.

If a random variable can be expressed as a sum of two independent random variables, its probability mass function (PMF) is then the convolution of the PMFs of the components variables, and its probability generation function (PGF) is the product of that of the component variables.

The examination of PGFs concluded that the node degree distribution in power grids can be very well approximated by a sum of two independent random variables, that is,

$$\mathcal{K} = \mathcal{G} + \mathcal{D}, \tag{12}$$

where $\mathcal{G}$ is a truncated Geometric with the threshold of k_{max}

$$\begin{aligned} \text{Pr}_{(\mathcal{G}=k)} &= \frac{(1-p)^k p}{\sum_{i=0}^{k_{max}} (1-p)^i p} \\ &= \frac{(1-p)^k p}{1-(1-p)^{k_{max}+1}}, \quad k = 0, 1, 2, \cdots, k_{max} \end{aligned} \tag{13}$$

with the PGF as

$$\begin{aligned} G_{\mathcal{G}}(z) &= \frac{\sum_{k=0}^{k_{max}} (1-p)^k p z^k}{1-(1-p)^{k_{max}+1}} \\ &= \frac{p\left(1-((1-p)z)^{k_{max}+1}\right)}{(1-(1-p)^{k_{max}+1})(1-(1-p)z)} \end{aligned} \tag{14}$$

And $\mathcal{D}$ is an irregular Discrete $\{p_1, p_2, \cdots, p_{k_t}\}$,

$$\text{Pr}_{(\mathcal{D}=k)} = p_k, \quad k = 1, 2, \cdots, k_t \tag{15}$$

with the PGF as

$$G_{\mathcal{D}}(z) = p_1 z + p_2 z^2 + p_3 z^3 + \cdots + p_{k_t} z^{k_t} \tag{16}$$

Table IV
ESTIMATE COEFFICIENTS OF THE TRUNCATED GEOMETRIC AND THE IRREGULAR DISCRETE FOR THE NODE DEGREES IN THE NYISO SYSTEM

node groups	max($\underline{k}$)	p	k_{max}	k_t	$\{p_1, p_2, \cdots, p_{k_t}\}$
All	37	0.2269	34	3	0.4875, 0.2700, 0.2425
Gen	37	0.1863	36	1	1.000
Load	29	0.2423	26	3	0.0455, 0.4675, 0.4870
Conn	21	0.4006	18	3	0.0393, 0.4442, 0.5165

Therefore the PMF of $\mathcal{K}$ is

$$\text{Pr}_{(\mathcal{K}=k)} = \text{Pr}_{(\mathcal{G}=k)} \otimes \text{Pr}_{(\mathcal{D}=k)} \tag{17}$$

And the PGF of $\mathcal{K}$ can be written as

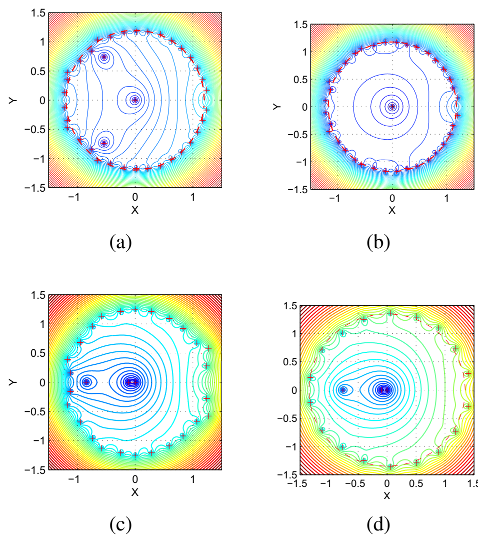
$$G_{\mathcal{K}}(z) = \frac{p\left(1-((1-p)z)^{k_{max}+1}\right)\sum_{i=1}^{k_t} p_i z^i}{(1-(1-p)^{k_{max}+1})(1-(1-p)z)} \tag{18}$$

The equation (18) indicates that the PGF $G_{\mathcal{K}}(z)$ has k_{max} zeros evenly distributed around a circle of radius of $\frac{1}{1-p}$ which are introduced by the truncation of the Geometric $\mathcal{G}$ (because the zero at $\frac{1}{1-p}$ has been neutralized by the denominator $(1-(1-p)z)$ and has k_t zeros introduced by the irregular Discrete $\mathcal{D}$ with $\{p_1, p_2, \cdots, p_{k_t}\}$.

Fig. 2 below show the contour plots of PGF of node degrees for different types of buses in the NYISO system. From the contour plots one can easily locate the zeros in PGF, and further determine the coefficients of corresponding distribution functions. The estimated coefficients for all the buses and each type of buses in the NYISO systems are listed in the Table IV. Fig. 3 compares the probability mass function (PMF) with estimate coefficients and the empirical PMF of the NYISO system and shows that the former matches the latter with quite good approximation; and the G/L/C types of buses can be characterized with different PMFs for their node degrees respectively.

Some interesting discoveries include: (a) Clearly each plot in Fig. 2 contains evenly distributed zeros around a circle, which indicate a truncated Geometric component; (b) Besides the zeros around the circle, most contour plots also have a small number of off-circle zeros, which come from an embedding irregular Discrete component; (c) The contour plot for each group of nodes has zeros with similar pattern but different positions. This implies that each group of node degrees has similar distribution functions but with different coefficients. Therefore it is necessary and reasonable to characterize the node degrees distribution according to the bus types. Otherwise if the node degrees aggregate into one single group, just as in Fig. 2(a), some important characteristics of a specific type of node degrees would be concealed (e.g., comparing (a) and (b)-(d) in Fig. 2).

From the topology analysis for G/L/C grid buses based on the average node degree, the clustering coefficient, and



(a) (b) (c) (d)

Figure 2. The Contour Plot of $E(z^X)$ of Node Degrees for Different Groups of Buses in the NYISO system: (a) All buses; (b) Gen buses; (c) Load buses; (d) Connection buses; the zeros are marked by red '+'s.

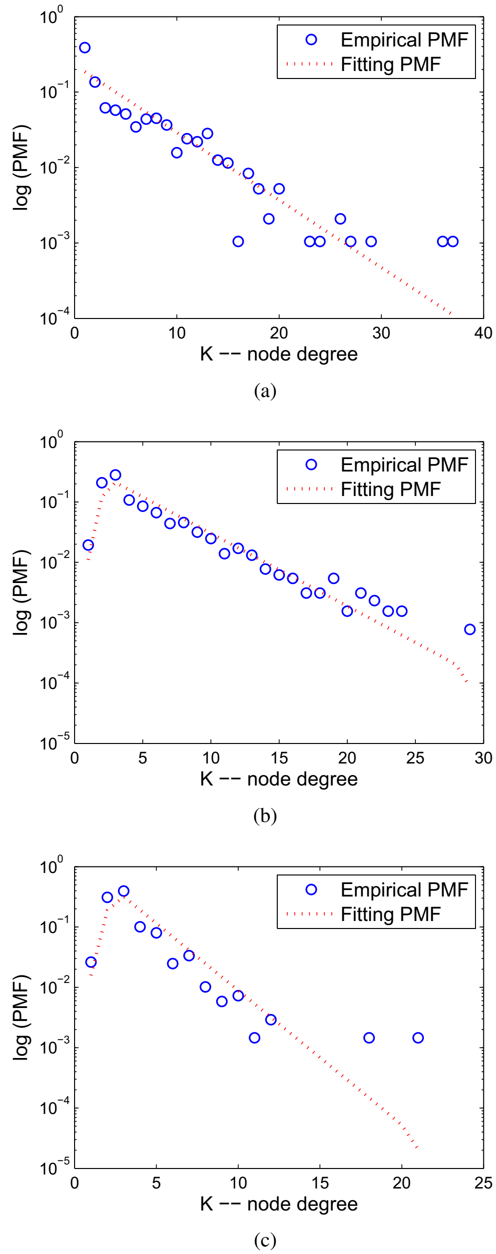
the node degree distribution, we can see that these metrics closely depend on and change with the bus type assignment in a grid. For different type of G/L/C buses there will be different topology characteristics. In other words, the bus type assignment in a typical power grid is not simply uniform. Therefor the random topology power grid model should avoid to use random bus type assignment in its topology. Otherwise the generated test cases may contains unrealistic settings that affect the grid dynamics and lead to misleading results.

IV. Impacts of Bus Type Assignments on Grid Vulnerability

In this section we will use the stochastic cascading failure model introduced in [26] to examine how the bus type assignments may affect the grid's vulnerability to cascading failures.

With the derived flow statistics from (5), one can then approximate the dynamics of the line state therefore the grid state accordingly using a Markovian transition model. A line is considered as *overloaded* if the power flow through it exceeds the line limit determined by its thermal capacity or static/dynamic stability conditions, i.e., which is called the line's *overload threshold*, where $F_l^{\max}$. Therefore the normalized overload distance for a line flow can be written as $a_l = \left(F_l^{\max} - \mu_{F_l(t)}\right)/\sigma_{F_l(t)}$ with $\sigma_{F_l} = \sqrt{C_{F_{ll}}(t,0)}$. And its overload probability can be approximated as $\rho_l(t) \approx \mathcal{Q}(a_l)$.

The persistent overload condition may cause a line to trip shortly, consequently, a transition in the state $\mathrm{s}(t)$. Fig.4 illustrates a line flow process for which two kinds of sojourn intervals can be defined: the *overload intervals* (when the



(a) (b) (c)

Figure 3. Comparing the Empirical and Fitting PMF of Node Degrees in NYISO: (a) Gen buses; (b) Load buses; (c) Connection buses

flow magnitude stays above its threshold: $|F_l(t)| \geq F_l^{\max}$), and the *normal-load intervals* (when $|F_l(t)| \leq F_l^{\max}$), which are associated with an overload and normal-load line-tripping rate respectively, denoted as λ_l^* and λ_l^0. Obviously $\lambda_l^* \ll \lambda_l^0$. Assuming that $F_l(t)$ is Gaussian and differentiable, one can compute the average level crossing rate (i.e. the expected number of crossings of the threshold $F_l^{\max}$ in either direction) [28] as $\gamma_l = \frac{W}{\pi} e^{-a_l^2/2}$, where $W = \sqrt{-R''_{F_l}(0)/R_{F_l}(0)}$ is the equivalent bandwidth.

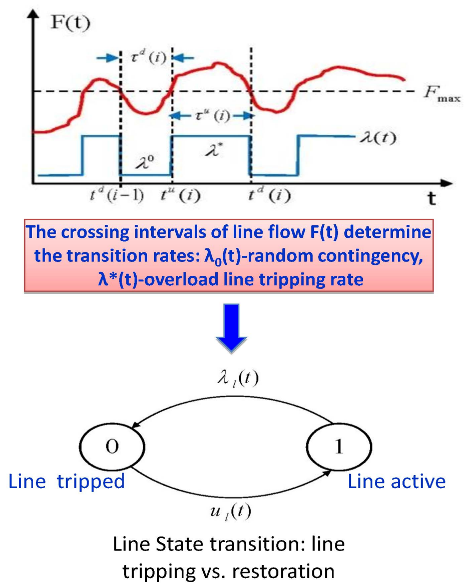


Figure 4. The flow process and the line state transition

The probabilistic distribution of crossing intervals and line states can then be derived by using the Rice's results on Gaussian random process [29] and compute the expected safety time for each line to stay connected given current network topology and operating state. More details of the derivation can be found in [26]. Given the overload probability ρ_l of the l-th line, the level crossing density γ_l at $F_l^{\max}$, and the line tripping rates λ_l^* for an overload line and λ_l^0 caused by random contingencies, the expected number of crossings after which the l^{th} line finally gets tripped is $\bar{\kappa}_l = [(1+\beta_l) - (\beta_l - \alpha_l)\rho_l] / (1-\alpha_l\beta_l)$, and the expected life-time of the line is:

$$\mathcal{T}_l = (\bar{\kappa}_l - 1)/\gamma_l + E\{\Delta t_l\}, \tag{19}$$

where $E\{\Delta t_l\}$ is the mean duration of the *last* interval $E\{\Delta t_l\} = \left[\Delta T_l^* + \Delta T_l^0\right] / (1 - \alpha_l\beta_l)$, with $\alpha_l = E\left\{p\left\{s_l\left(t_l^d(i)\right) = 1\right\}\right\}$, $\beta_l = E\left\{p\left\{s_l\left(t_l^u(i)\right) = 1\right\}\right\}$, $\Delta T_l^* = (1-\alpha_l)\left[\beta_l + (1-\beta_l)\rho_l\right] / \lambda_l^*$ and $\Delta T_l^0 = (1-\beta_l)\left[1 - (1-\alpha_l)\rho_l\right] / \lambda_l^0$. Therefore, given the statistical settings of a grid, one can then estimate of the expected life-time of all the lines and evaluate the grid vulnerability to the arrival of possible cascading failures.

Here we will use the IEEE 300 bus system as an example to investigated the impacts of different bus type assignments on the grid vulnerability to cascading failures. Keeping unchanged the ratios of G/L/C grids buses and the corresponding statistics of the generation and load settings, we apply different random bus type assignments to the original grid topology. Then we examine and compare the expected life-time $\mathcal{T}_l$ of each line in the new grid system and the original IEEE 300 bus system. The simulation parameters are given as follows. The initial operating equilibrium and conditions are taken and derived from the power flow solution of the system data from [3]. Taking the mean of $P(0)$ as $\mu_P(0) = [G(0)^T, -L(0)^T]^T$, we set the standard deviation of the loads as $\sigma_L = 0.07|L(0)|$, but ignore the variance in G. For simplicity, we assume that the loads and generation are statistically independent of one other. The line overload thresholds are set as $F^{\max} = 1.20|F(0)|$. Here we take $F(0)$ as the original flow distribution under normal operating conditions and assume that the line capacity allows a 20% load increase. The overload and normal-load line tripping rates are set as $\lambda^* = 1.92 \cdot 10^{-2}$Hz and $\lambda^0 = 7.70 \cdot 10^{-11}$Hz respectively. Analysis on the load record from realistic power grids [30] has shown that the load process can be approximated as a low-pass Gaussian process with an equivalent bandwidth of $W \approx 10^{-5}$Hz. Since the flow process in a grid can be seen as a linear projection from the load process, we can apply the equivalent bandwidth W to the flow processes in the grid. Fig. 5 shows the comparison results of the first 60 most critical lines with shortest expected life-time. The blue solid line plots the results of the original grid of IEEE 300 bus system. The green and purple dashed lines represent the $\mathcal{T}_l$ results obtained from two test cases of IEEE 300 bus system with random bus type assignments. The red solid line with "x" marks is the $\mathcal{T}_l$ averaged from the results of 10 random assignment cases. From the figure we can see that given the same topology and generation and load statistical settings, the test cases with random bus type assignments tend to have larger expected life time than that of the realistic grid settings. In some cases the increase of the expected life-time can be up to 150% of the realistic $\mathcal{T}_l$. In other words, if we utilize random bus type assignment in our random topology power grid model, i.e. *RT-nested-smallworld*, 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.

V. Conclusion

In this paper we examined the correlation between the three bus types of G/L/C and some network topology metrics such as the average node degree, the clustering coefficient and the node degree distribution. We also investigated the impacts of different bus type assignments on the grid vulnerability to cascading failures using IEEE 300 bus system as an example. It is found that the node degree distribution and clustering characteristic are different for different type of buses (G/L/C) in a realistic grid; 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 our random topology power grid model, i.e. *RT-nested-*

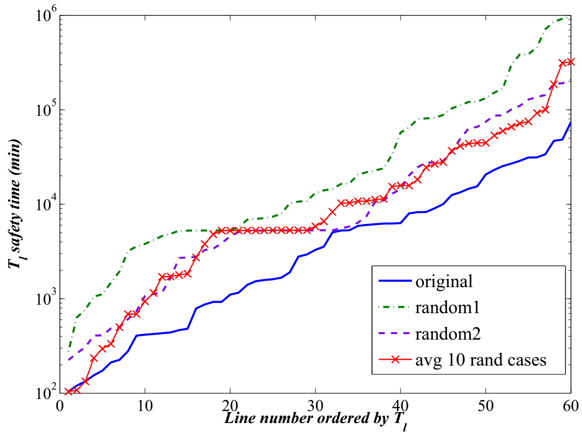


Figure 5. Safety Time of Grid Lines of IEEE 300 Bus System: Original vs Random Bus Type Assignments

smallworld, 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 a random topology power grid model should be improved by using a more accurate assignment which is consistent with that of realistic grids.

Acknowledgment

The authors would like to thank Dr. Cris Moore from Santa Fe Institute and the graduate student Rameez Khimani from Virginia Commonwealth University for the initial discussions on the design of bus type assignment algorithms for random topology power grid models.

References

[1] 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*, volume 1(1):28-39, 2010.

[2] F. L. Alvarado, "Computational Complexity in Power Systems", *IEEE Trans on PAS*, vol. PAS-95, no. 4, Jul/Aug 1976.

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

[4] P. Maghouli, S.H. Hosseini, M.O. Buygi, M. Shahidehpour, "A Multi-Objective Framework for Transmission Expansion Planning in Deregulated Environments", IEEE Transactions on Power Systems, vol.24(2):1051-1061, May 2009.

[5] J. Choi, T.D. Mount, R.J. Thomas, "Transmission Expansion Planning Using Contingency Criteria," IEEE Transactions on Power Systems, vol.22, no.4, pp.2249-2261, Nov. 2007.

[6] G. Bei, "Generalized Integer Linear Programming Formulation for Optimal PMU Placement," Power Systems, IEEE Transactions on , vol.23, no.3, pp.1099-1104, Aug. 2008.

[7] Z. Wang, A. Scaglione and R.J. Thomas, "Compressing Electrical Power Grids", 1st IEEE International Conference on Smart Grid Communications (SmartGridComm), Gaithersburg, Maryland, October 4-6, 2010.

[8] S. Galli, A. Scaglione, and Z. Wang, "Power Line Communications and the Smart Grid", 1st IEEE International Conference on Smart Grid Communications, Gaithersburg, Maryland, Oct 4-6 2010.

[9] H. Wang, J.S. Thorp , "Optimal locations for protection system enhancement: a simulation of cascading outages," Power Delivery, IEEE Transactions on , vol.16, no.4, pp.528-533, Oct 2001.

[10] J. Matevosyan, L. Soder, "Minimization of imbalance cost trading wind power on the short-term power market," Power Systems, IEEE Transactions on , vol.21, no.3, pp.1396-1404, Aug. 2006.

[11] A.G. Bakirtzis, P.N. Biskas, "A decentralized solution to the DC-OPF of interconnected power systems," Power Systems, IEEE Transactions on , vol.18, no.3, pp.1007-1013, Aug. 2003.

[12] 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.

[13] 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*, volume 12(4), 2002, pp.985-994.

[14] M. Parashar, J. S. Thorp, "Continuum modeling of electromechanical dynamics in large-scale power systems," *IEEE Trans. on Circuits and Systems*, volume 51(9), 2004, pp.1848-1858.

[15] D. J. Watts, S. H. Strogatz, "Collective dynamics of 'Small-World' networks," *Nature*, volume 393, 1998, pp.393-440.

[16] Z. Wang, R. J. Thomas, A. Scaglione, "Generating random topology power grids," *Proc. 41st Annual Hawaii International Conference on System Sciences (HICSS-41)*, volume 41, p. 183, Big Island, Jan 2008.

[17] 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), 2010 Asia-Pacific , pp.1-4, 28-31 March 2010.

[18] R. Albert and A. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, volume 74(1), 2002, pp.47-97.

[19] P. Erdös, and A. Rényi, "On random graphs. I.," *Publicationes Mathematicae*, volume 6, 1959, pp.290-297.

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

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

[22] M. Rosas-Casals, S. Valverde, and R. Solé, "Topological vulnerability of the European power grid under erros and attacks," *International Journal of Bifurcations and Chaos*, volume 17(7), 2007, pp.2465-2475.

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

[24] E. P. Wigner,"Characteristic Vectors of Bordered Matrices with Infinite Dimensions," *The Annals of Math*, volume 62, 1955, pp.548-564.

[25] E. P. Wigner, "On the Distribution of the Roots of Certain Symmetric Matrices," *The Annals of Math*, volume 67, 1958, pp.325-328.

[26] Z. Wang, A. Scaglione, and R. J. Thomas, "A markov-transition model for cascading failures in power grids," in *45th Hawaii International Conference on System Sciences*, Maui, Hawaii, Jan. 2012, pp. 2115–2124.

[27] A. Wood and B. Wollenberg, *Power System Generation, Operation and Control*. New York: Wiley, 1984.

[28] A. Papoulis, *Probability,Random Variables, and Stochastic Processes, chapter 16*. McGraw-Hill International Editions, 1991.

[29] S. O. Rice, "Distribution of the duration of fades in radio transmission:gaussian noise model," *Bell Syst. Tech. J.*, vol. 37, p. 581-635, May 1958.

[30] National Grid, http://www.nationalgrid.com/uk/Electricity/Data/Demand+Data.