Building Synthetic Power Transmission Networks of Many Voltage Levels, Spanning Multiple Areas

Adam B. Birchfield, Ti Xu, Komal S. Shetye, and Thomas J. Overbye
Department of Electrical and Computer Engineering, Texas A&M University
abirchfield@tamu.edu, txu@tamu.edu, shetye@tamu.edu, overbye@tamu.edu

Abstract

Synthetic power grids, that is, test cases designed to match realistic structural and statistical characteristics of actual grids, are useful for research, development, and demonstration of innovations, since the cases are fictitious and thus free from data confidentiality issues. Building on previous work, this paper addresses a couple of related problems in the transmission network synthesis process. These issues appear as created cases become larger and involve multiple areas and overlapping nominal voltage levels. A fast, scalable hierarchical clustering is designed to assign voltage levels to substations considering the needs of the system, the specific constraints of the area, and smooth interconnections between neighboring areas with different voltage levels. A line topology generation framework is considered that is appropriate for many networks of different voltage levels, constructed together for a useful, realistic grid. These methods are demonstrated in a new 2000 bus test case, validated and publicly released.

1. Introduction

Public test cases representing large electric power systems at a high level of fidelity and quality are few to non-existent, despite the potential value such cases would have to the power systems research community. Since the security of an actual power grid prohibits free dissemination of its detailed models, researchers are left to cases they can obtain through non-disclosure agreements, which hinder publication, or literature test cases that differ widely from the actual grid in size, structure, and complexity.

The concept of synthetic power grids refers to a systematic way of building fully public test cases for the research community. These cases' size, structure, and features are anchored in a robust statistical and structural analysis of the actual grid. Synthetic grid models are situated on a real geographic footprint, with ties to existing public energy data; however, the transmission networks themselves are entirely fictitious, with no direct correspondence whatsoever to any actual grid. The grids look and feel real, and are solidly effective for a variety of research studies; but since they are synthetic they are able to be freely published.

Initial work on synthetic power grids with this framework [1]-[2] builds fully ac-solvable realistic power grid models of size 150 and 2000 buses that meet specified criteria. Earlier foundational work includes the topology analysis of [3]-[15]. These are only samples of the large body of literature that has examined the graph structure of power grids in the last few years. In [11]-[15], network generation was done, with only the topology considered, and no ac circuit solution. In [16], a large test case is developed, however, it is only dc-solvable and approximates the actual grid network, rather than being fully synthetic. Since the publication of [1]-[2], additions to the base synthesis algorithm include generator cost curves for economic and optimal power flow (OPF) analysis [17], initial transient stability models for dynamic simulations [18], more detailed transmission line parameters [19]-[20], and a more robust set of validation criteria, including two additional cases with size 200 and 500 buses [21]. In [22]-[24], further developments in alternative methods for network topology generation are given, not including solvable circuits. The increased attention and interest in synthetic test cases, and their potential to drive innovation for a smarter grid, warrants research to make these grid models better, more complex, and larger.

This paper addresses a topic only briefly introduced in [2], which is network synthesis considering multiple nominal voltage levels and multiple geographic areas. Whereas for grid models with a size below 1000 buses it is acceptably realistic to have only two voltage levels (such as 115 kV and 345 kV) that fully cover a single area, for larger cases realistic grids will have more voltage levels, some or

HICSS

Fig. 1. The eight areas of Texas used for the system (from [2]). The areas are: Far West (blue), West (red), North (yellow), North Central (pink), South Central (indigo), South (green), East (cyan), and Coast (orange).

all of which will only span part of the entire system. Multiple areas and voltage levels, which are essential to building cases the size of real continental interconnects, present several challenges in that voltage levels can only be connected to each other through transformers when they coincide at substations, and thus the overall branch-bus topology must be considered in conjunction with the topology of the substation voltage networks individually.

Two related problems are formulated and solved in this paper, which fall in the category of system planning. Given a set of substations geographically placed across a system, with each area having its designated set of voltage levels, the first problem is to assign these voltage levels to specific substations appropriate to real statistics and system needs. Second, the algorithms to build the network topology must be adjusted to ensure that the designated characteristics are met, with multiple voltage level networks interacting. While the various algorithms presented by this paper do not consider all the complex factors that are used in actual power system planning, the resulting substation bus assignments and network properties are shown to match important features of the real grid, and the resulting cases can be validated for realism.

By way of example, a new public test case is presented by this paper. Similar to the case of [2], the geographic footprint of Texas is used to build a 2000 bus case, named ACTIVSg2000. This case is entirely new and includes more voltage levels as well as an updated algorithm, building upon [2]. Section 4 discusses this case in detail, and it is used as an example throughout the paper.

2. Choosing which substations should be upgraded to higher voltage levels

In [2], two voltage levels were used, 115 kV and 345 kV. Each substation was given a 115 kV bus, and 15% were selected at random to be given a 345 kV bus. The random selection was not uniform, but larger generators and loads were given a higher probability. This served as an excellent simplified approximation. To improve realism and prepare for future larger and more complex cases, the analysis of this paper rebuilds the case with four voltage levels, assigned to areas in Fig. 1 according to Table I. Though these areas were used before, in this paper they are integral to the formation of the network.

The values in Table I are designed to create four overlapping voltage networks that cover the whole case with sufficient coverage for each area and sufficient diversity across the case. Many areas have

TABLE I
VOLTAGE LEVEL SPECIFICATIONS USED IN ACTIVSG2000

Area	Percent of substations containing			
	500 kV	230 kV	161 kV	115 kV
Far West		22%		100%
West	8%	18%		100%
North	10%		100%	
North Central	12%		100%	
South Central	8%	18%		100%
South	8%	18%		100%
East	13%	22%		100%
Coast	13%	22%		100%

three voltage levels, while some have only two. The 500 kV network ties most of the system together, but the underlying networks vary considerably. The percentages given are based on the principle, observed on actual grids, that nearly all substations will have a connection to an area's lowest voltage level, while about 10-20% will have a connection to the higher level. If there is a third level, 5-15% of substations will be contained.

With substations selected based on public population and energy data, similar to the method of [2], the set of substations are geographically assigned to the eight areas. Then, voltage level assignments must be made, satisfying the designed percentages from Table I and also the needs of the network.

2.1. Upgrades for connections within areas

Since the lowest voltage level covers the most substations, all of them are initialized to that level;
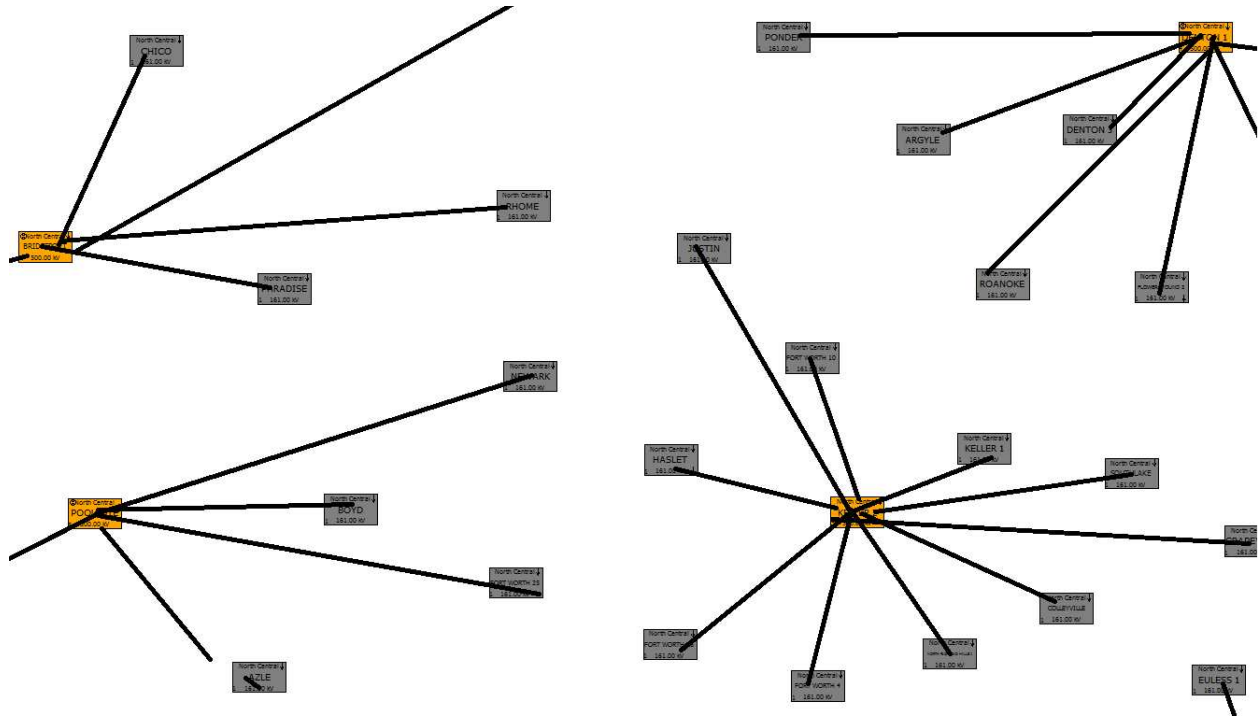
Fig. 2. Example of the parent-child hierarchy in the 161-500 kV North Central area near Fort Worth. The boxes represent substations. The 500 kV substations are selected to minimize the product of distance with the total load and generation of each child 161 kV substation. Note that this does not show the transmission lines, only the hierarchical structure used to assign voltage levels.

then, some are upgraded to higher levels. If desired, some of the higher-voltage substations could subsequently have the lower voltage levels removed.

In deciding which voltage levels to add to which substations, two considerations are made. The first is that there must be a sufficient high-voltage network available to meet the demand of the substation and its nearest neighbors' load. Second, the voltage levels must line up across area borders so that the areas can interconnect at all levels. Subsection 2.2 will address the second concern; this section deals with the former.

Computational complexity benefits from the fact that this first concern only considers the needs of a single area. No matter how large a system grows, the time to complete this step will only grow linearly with the number of areas, with the size of each area being the primary determining factor.

The approach taken is to cluster the substations into groups, where one substation in the group is considered the "parent" substation and the others are "children." Ultimately, the parent substations will be upgraded to the higher voltage level, though it will not necessarily need to connect to its children in the transmission network. The parent-child hierarchy is illustrated in Fig. 2. For networks with three voltage levels, two iterations of the clustering are done, first to cluster the lower voltage substations into medium-voltage substations, then to cluster the medium-

voltage substations into the highest-voltage substations, with two levels of resulting hierarchy.

The algorithm assigns each substation a base weight, which is defined as the sum of the loads and the generator capacities at that substation.

$$W_i = \sum P_{g,i} + \sum P_{l,i} \qquad (1)$$

For the two-level hierarchy, the weight in the second level is the same as the first level plus sum of the weights of all children from the first level.

The objective of the clustering is to minimize the weight of N clusters, where N is found from the total number of substations and Table I. The weight of a cluster is a function of the weight of its members and their distance from the parent.

$$W_c = \sum_i^{i \in c} W_i d_{ic} + \alpha_p \sum_i^{i \in c} W_i^2 + \alpha_d \sum_i^{i \in c} d_{ic}^2 \qquad (2)$$

where W_i for each child is defined in (1), d_{ic} is the geographic distance of the child to the cluster parent, and α_P and α_d are parameters, which can be used to tune the algorithm to focus more on limiting total power handled by a substation or total distance. For the implementation in this paper, $\alpha_p = \alpha_d = 0$ seemed to give an acceptable solution.

Perfect satisfaction of the optimality formulation

Fig. 3. Example cross-area connection. The 161 kV substation (gray, top) and the 500 kV substation (orange, center) each have a 115 kV bus added, which is not part of their area's network, so that they can connect to the 115 kV substation in another area (gray, bottom left). Transmission network is shown.

$$\min \sum_{c}^{N} W_c \qquad (3)$$

is not needed, but the definition in (2) provides a sufficient gradient to do the clustering. A greedy, steepest-descent approach is used, where each substation begins in its own cluster, with the total cost function equal to zero. At each iteration, one substation is selected to become the child of its nearest neighbor, and the cost function is updated. The substation is selected based on the smallest incremental change to the total cost function. The process continues until the number of clusters has been reduced to N. If at any iteration the selected substation already contains children, each child is assigned a new parent substation to which it is nearest. These changes are considered in the gradient analysis. To speed up this process, at the beginning each substation is assigned a parent priority list, which is the list of all other substations ordered by distance from it. Therefore, each time it needs to check for a new parent, it only needs to check at most a few substations on its list. The result is that the algorithm is relatively quick for up to a few hundred substations.

With this algorithm, each substation that is upgraded to a higher voltage level serves a purpose: it provides a network connection to the higher voltage level for a number of substations in its geographic neighborhood. The more load and generation in a neighborhood, and the more distance between substations in it, the more high-voltage substations it will have. Remote substations with very low loads may not need the full high-voltage transmission out there; similarly, a large urban area will not need

excessive high-voltage substations clumped together. This algorithm establishes a balance that covers the area in proportion to the load density and geographic density. At the same time, the largest generators are almost guaranteed to be upgraded to higher voltage levels, because making them children of another substation would increase the cost function too much. Thus, the results are similar to the original method of [2], but the approach is more systematic and less probabilistic.

2.2. Upgrades for cross-area connections

Focusing the voltage assignment problem on area needs individually is useful, not only to reduce the computation time, but also to avoid the confusion of voltage levels which span some but not all of the system. However, at the boundaries of the areas there must be ways for differing voltage levels to connect. Appropriately selecting these substations is the subject of this subsection.

The first issue is finding out which areas border each other and how strong that connection is. A simple metric employed by this paper is to use the Delaunay triangulation of all the system substations. Described in [1], the Delaunay triangulation is quick to calculate and provides a good approximation of neighborhood for geo-located points. For any two areas that are connected to each other by at least 10 Delaunay segments, there ought to be connections at high, medium, and low voltage levels. (Note that throughout this paper, the terms high, medium, and low voltage levels refer to relative magnitude within the high-voltage transmission range of 100+kV).

For any boundary where the two areas share voltage levels, no additional connections need to be made. They will be connected using these shared levels. For any boundary with mismatched voltage levels, additional buses are added to some substations along with transformers that connect them. These substations are selected as the shortest ones in a separate Delaunay triangulation using only the substations involved. Fig. 3 shows an example where additional buses have allowed lower-voltage connections between the areas that would not otherwise have been possible. So there will be 161:115 kV transformers at these boundary points but not anywhere else, similar to the actual grid. These inter-area connections are essential to the interconnected nature of the grid.

3. Creating the multi-level topology

This section reformulates the network synthesis algorithm outlined in [2] to make several changes. These modifications are related to the multiple voltage levels and multiple areas, but also contribute more generally to the quality of the cases, allowing flexibility to add additional constraints and objectives. Existing constraints fall in two categories, described below: the ones which consider power flow analysis and those which consider topology. The problem has some similarity to transmission expansion planning, but the starting point is an empty graph and the objective is realism rather than optimal performance.

As section 3.3 will describe in further detail, this new method begins with a minimum spanning tree and both adds and removes lines at potentially each iteration, rather than simply adding them from scratch without removal. This means that analyses can affect the reward and penalty structures differently in the removal and add stages.

3.1. Initial dispatch and power flow considerations

The method described in [2] runs an iterative dc power flow solution that estimates the power that would flow through each potential line and contributes a reward for that line proportional to the potential power. This reward competes with other rewards and penalties from other analyses to determine whether the line will be added. A similar structure is used here, with the power flow results affecting points for both removal (using calculated power flow in existing lines) and addition (using potential power flow as before). However, four important adjustments are made. First, the power flow part of the analysis is heavily dependent on the assumptions made with respect to the generation dispatch. Second, the power flow must be normalized so that it affects multiple voltage levels equally. Third, the power flow ought to affect addition much more than removal, as often lines with low power flow exist for other reasons. Fourth, near the end of the development, power flow considerations should not be allowed to swamp meeting the targets in other areas.

For the generator dispatch, several options are available. The solution proposed by [2] is reasonable, which pre-specifies inter-area exchanges and then dispatches all generators proportional to net load. An improvement upon this method is to set non-dispatchable load such as wind, hydro, and solar to some fixed proportion such as 25%, then dispatch the rest according to an equal-lambda economic dispatch.

TABLE II
CHARACTERISTIC PER-DISTANCE X AND MW LIMIT
FOR VOLTAGE LEVELS USED IN NETWORK SYNTHESIS

Nominal kV	Characteristic p.u. X/mile	Characteristic MW
115 kV	0.0054	160
138 kV	0.0040	223
161 kV	0.0029	265
230 kV	0.0015	541
345 kV	0.00058	1195
500 kV	0.00025	2598
765 kV	0.00012	4100

Creating synthetic cost curves was described in [17], and the local cost of coal and gas can be adjusted to tweak the system inter-area power flows as design goals require. Making these assumptions for the initial dispatch will aid in making the cases more applicable to optimal power flow (OPF) analysis.

With multiple voltage levels across many areas, the amount of MW flow that would be considered large through a transmission line varies greatly. With this in mind, it is important to ensure that the reward given for large power flow is not overly dominant at the high voltage levels and insignificant in the low voltage levels. To address this concern, a characteristic MW value is assigned to each voltage level, as shown in Table II. The actual or potential power value is divided by the characteristic value and capped at 1 to get a normalized value. This normalized value is then multiplied by some reward that can be adjusted, but will have a similar effect on all voltage levels. The reward used in the synthesis of this paper is 500 for addition and 50 for removal. These rewards are added for the first 80% of iterations, after which other considerations are allowed to dominate.

3.2. Delaunay triangulation and topological considerations

In [2], the Delaunay triangulation is noted for its usefulness in dramatically weeding out potential transmission lines as well as being a proxy for a variety of topological and geographic characteristics that are observed in actual grids. By restricting the search space to sections within three hops on the Delaunay triangulation, the computation time is dramatically reduced and over 99% of actual lines are considered. Moreover, by matching the percentages of first, second, and third neighbor Delaunay segments, the clustering coefficients, shortest path length, and node degree distribution naturally turn out similar to an actual grid. Meeting these objectives remains a key priority in building synthetic networks.
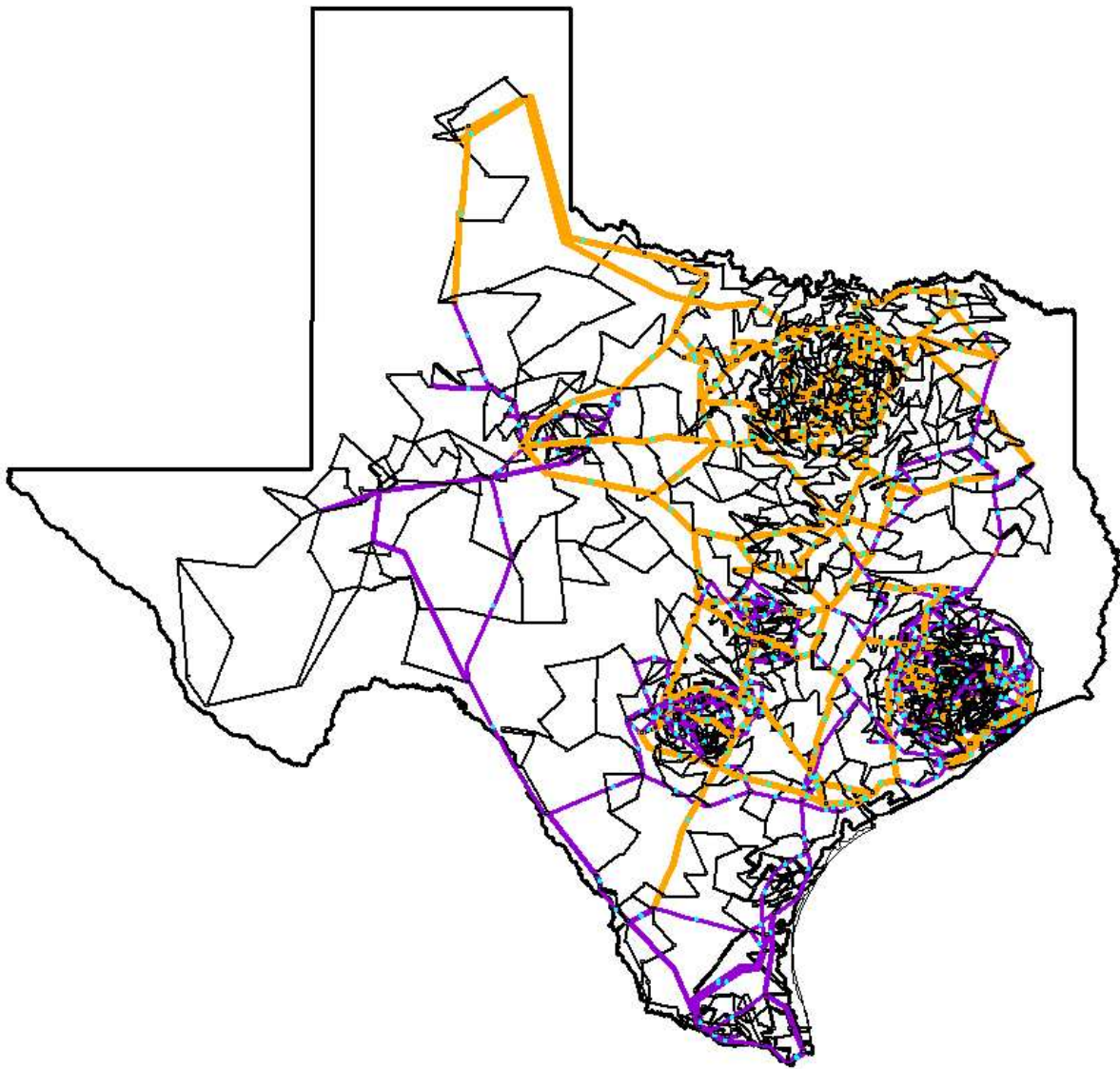
Fig. 4. The transmission network for the ACTIVSg2000 case. The 500 kV network (orange) covers all of the system except the far west. The 230 kV network (violet) is not present in the northern part of the grid. The 161 kV and 115 kV networks (black) split the grid between north and south. This case is entirely synthetic and does not represent any actual grid. It is publicly available online.

Additional topological requirements are heuristic, as bi-connectivity is enforced, along with connectivity of each voltage network and the forbidding of any radial loads. Each of these can be checked with a linear-time depth first search analysis. These are given rather high penalties or rewards, to ensure that the requirements are met. For example, all candidate lines that connect to a radial sub are given a reward of 200, and any line which is one of only two connections to a substation is given a removal penalty since removing it would create a radial substation.

3.3. Structure and implementation of topology algorithm

The network generation algorithm is an iterative process, and at each iteration there are several steps. A few analyses are performed for the network as a whole, including the dc power flow and depth first search for bi-connectivity, as well as the identification of radial substations. Then each voltage level is considered in turn. Analysis is made of the proportions of Delaunay first, second, and third neighbors matched, respectively, and the connectivity of the voltage level. Then 0 or 1 lines are removed and 1 or 2 lines are

TABLE III
Validation of ACTIVSg2000 case

#	Validation Metric	Criteria	ACTIVSg2000			
			Mean 1.7-3.5	1.6		
			Exponential decay	See Fig. 5		
			< 200 kV, 85-100%	100%		
			> 201 kV, 7-25%	16%		
3	Substations with load	75-90%	90%			
			Mean 6-18 MW	33.6 MW		
			Exponential decay	See Fig. 6		
5	Generation capacity / load	1.2-1.6	1.46			
6	Substations with generators	5-25%	15%			
			25-200 MW, 40+%	55%		
			200+ MW, 5-20%	30%		
8	Committed Generators	60-80%	79%			
9	Generators dispatched > 80%	50+%	67%			
10	Generator MaxQ/MaxP	0.40-0.55, > 70%	75%			
			500 kV	320 kV	161 kV	115 kV
			99 %	94 %	90 %	82 %
		40% below median	51/49	46/41	50/45	48/45
		40% above median	50/51	54/59	50/55	52/55
		80% within 10-90 range	97/98	100/94	100/98	99/82
13	Line per-unit, per-distance reactance, by kV level (Table III)	70% within 10-90 range	96	98	100	100
14	Line X/R ratio and MVA limits, by kV level (Tables IV and V)	70% within 10-90 range	98/97	100/95	100/100	85/99
15	Lines / Substations, by kV level	1.1-1.4	1.29	1.26	1.23	1.25
16	Lines on min. spanning tree	45-55%	49.0	49.7	50.0	49.5
		1, 65-80%	75.2	75.4	75.3	74.8
		2, 15-25%	20.3	19.9	19.8	20.2
		3+, 3-10%	4.6	4.7	4.9	5.0
18	Total line length / MST	1.2-2.2	2.15	1.71	1.62	1.66
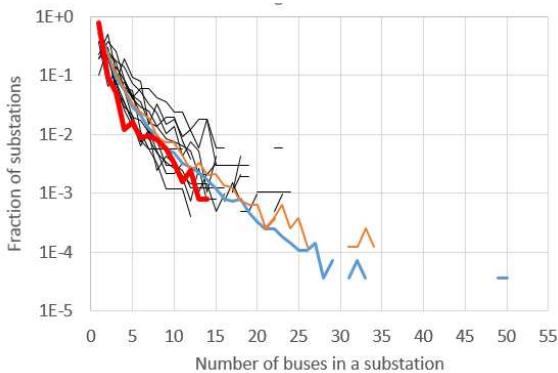
Fig. 5: ACTIVSg2000 case distribution of buses per substation (red). Actual case data is given for the Eastern Interconnect (blue), WECC (orange), and various subset cases.
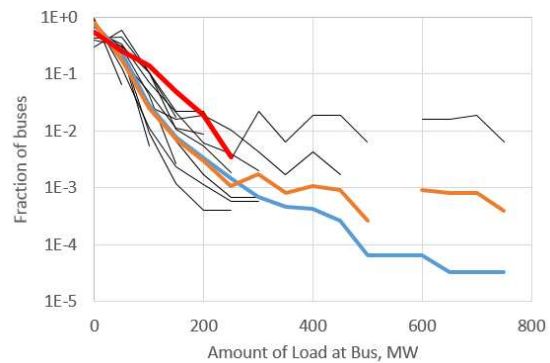
Fig. 6: ACTIVSg2000 case distribution of bus loads (red). The other lines are identical to Fig. 2. Actual case data is given for the Eastern Interconnect (blue), WECC (orange), and various subset cases.

added. Bridges (which would disconnect the network) can never be removed. Since the network is initialized to the minimum spanning tree, at the beginning all lines are bridges and therefore no lines can be removed. Until the designated number of lines is added, the number of lines added in an iteration is one more than the number removed. Each voltage network runs its own removal and addition at each iteration,

and the iteration continues for enough rounds to meet all the criteria.

4. Applications

The transmission network corresponding to the new 2000 bus case is shown in Fig. 4. This case was built with the methodology of this paper, revising the

framework of [2]. After the topology was automatically generated, additional parameters and modifications were manually made to meet ac power flow goals and contingency behavior.

Various statistics and validation metrics, as specified in [21], are shown in Table III and Figs 5-6. These validation metrics are based on analysis of the actual grid, with metrics taken from the Eastern Interconnect, WECC, and various subset cases of these. Note that the exception of the 2000 bus case from metrics 1, 4, and 7 are by design, since the number of buses was chosen to be smaller and more manageable than the usual size of a Texas grid. This case is publicly available online in a variety of data formats [25].

Fig. 7 shows the oneline diagram of an early draft of the ACTIVSg10K case, a 10,000 bus synthetic power grid situated on the geographic footprint of the Western Interconnection (WECC) in the United States. While the model is still under development, initial results show that the algorithms of this paper, designed for multiple areas and nominal voltage levels, are scalable to continental-size systems. In order to better match the richness of actual large-scale models, the 10,000 bus system includes additional complexities such as phase-shifting transformers and multiple reactive power devices remotely regulating a single bus voltage.

5. Conclusion

The main contribution of this paper is the set of updated algorithms for building synthetic power grid models that combine multiple voltage levels in different combinations across several areas in a system. The solutions to these problems lays groundwork for larger and more complex cases.

Synthetic power grids that are large, detailed, demonstrably realistic, and publicly available are valuable to the electric power research community. They fulfil a need to evaluate new innovations on realistic test cases, with the ability to fully publish and share all cases and results. The new 2000 bus case presented in this paper serves such a purpose.
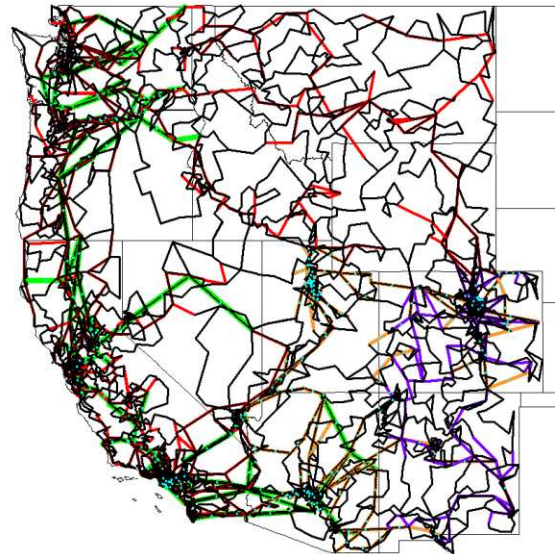
6. Acknowledgements

Fig. 7: Oneline diagram of ACTIVSg10K case early draft, demonstrating scalability of algorithms to 10,000 buses. This system is entirely synthetic and does not represent any actual grid.

7. References

[1] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye and T. J. Overbye, "Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1502-1510, March 2017.

[2] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, to be published.

[3] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998.

[4] M. Newman, "The Structure and Function of Complex Networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Jan. 2003.

[6] E. Cotilla-Sanchez, P. D. H. Hines, C. Barrows, and S. Blumsack, "Comparing the Topological and Electrical Structure of the North American Electric Power Infrastructure," *IEEE Systems Journal*, vol. 6, no. 4, pp. 616–626, Dec. 2012.

[7] G. A. Pagani and M. Aiello, "The Power Grid as a complex network: A survey," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 11, pp. 2688–2700, Jun. 2013.

[8] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, pp. 47-97, Jan. 2002.

[9] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Phys. Rev. E*, vol. 69, no. 2, Feb. 2004.

[10] P. Hines, S. Blumsack, E. Cotilla Sanchez, and C. Barrows, "The Topological and Electrical Structure of Power Grids," in *Proc. 2010 43rd Hawaii Int. Conf. System Sciences*, Koloa, HI, USA, Jan. 2010, pp. 1–10.

[11] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating statistically correct random topologies for testing smart grid communication and control networks," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 28–39, 2010.

[12] Z. Wang, R. J. Thomas, and A. Scaglione, "Generating random topology power grids," in *Hawaii International Conference on System Sciences, Proceedings of the 41st Annual*, 2008, pp. 183–183.

[13] 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, no. 1, pp. 28–39, 2010.

[14] 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," in *Proc. 2016 IEEE Power and Energy Conf. at Illinois,* Champaign, IL, Feb. 2016.

[15] B. Cloteaux, "Limits in modeling power grid topology," *2013 IEEE 2nd Network Science Workshop (NSW)*, West Point, NY, 2013, pp. 16-22.

[16] Q. Zhou and J. W. Bialek, "Approximate model of European interconnected system as a benchmark system to study effects of cross-border trades," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 782–788, May 2005.

[17] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Application of large-scale synthetic power system models for energy economic studies," *2017 50th Hawaii International Conference on System Sciences*, January 2017.

[18] T. Xu, A. B. Birchfield, K. S. Shetye, and T. J. Overbye, "Creation of Synthetic Electric Grid Models for Transient Stability Studies," *Bulk Power Systems Dynamics and Control Symposium (IREP 2017)*, Espinho, Portugal, September 2017.

[19] S. H. Elyas and Z. Wang, "Statistical analysis of transmission line capacities in electric power grids," *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Minneapolis, MN, 2016, pp. 1-5.

[20] M. H. Athari and Z. Wang, "Statistically characterizing the electrical parameters of the grid transformers and transmission lines," *Bulk Power Systems Dynamics and Control Symposium (IREP 2017),* Espinho, Portugal, September 2017.

[21] A. B. Birchfield, E. Schweitzer, H. Athari, T. Xu, T. J. Overbye, A. Scaglione, and Z.Wang, "Validation metrics to assess the realism of synthetic power grids," submitted to *IEEE Transactions on Power Systems, March 2017*.

[22] S. H. Elyas, Z. Wang, and R. J. Thomas, "On statistical size and placement of generation and load for synthetic grid modeling," *Bulk Power Systems Dynamics and Control Symposium (IREP 2017),* Espinho, Portugal, September 2017.

[23] S. Soltan and G. Zussman, "Generation of synthetic spatially embedded power grid networks," *2016 IEEE Power and Energy Society General Meeting (PESGM)*, Boston, MA, 2016, pp. 1-5.

[24] J. Hu, L. Sankar and D. J. Mir, "Cluster-and-Connect: An algorithmic approach to generating synthetic electric power network graphs," *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, 2015, pp. 223-230.

[25] http://electricgrids.engr.tamu.edu