ECEN 615
Methods of Electric Power System Analysis

Synthetic Power Grids

Adam Birchfield
Dept. of Electrical and Computer Engineering
Texas A&M University
abirchfield@tamu.edu

Complex networks are everywhere!

- There has been a lot of interest since the late 1990s in various "complex networks" or graphs and how their properties are important to the systems' function and reliability
  - These are systems with many nodes (vertices) connected by branches (edges)
- Examples are: biological (neurons), ecological (food chain), social (Facebook), professional (author citations, movie collaboration), and technical (internet, world wide web, airline routes, roads, power grids)

Types of complex network models

- Random graphs: Erdos-Renyi (1960)
  - As number of edges increases, system undergoes phase transition to one large connected component
- Small-World networks: Watts-Strogatz (1998)
  - Have the property of a very regular graph in that there is a high degree of clustering, but also have the random graph property of few hops between any two vertices
  - Watts and others said that power grids are Small-World, but others have pointed out limits to this characterization
- Scale-Free networks: Barabasi-Albert (1999)
  - Some very large degree nodes, such as the internet. System looks the same at all zoom levels.
  - These graphs are created by “preferential attachment” where high-degree nodes become more likely to attach new edges

The “small-world” concept

- Begin with regular lattice
- Rewire each edge with probability p
- L is the average shortest path length in hops between all pairs of vertices
- Each node with k neighbors has $k_v(k_v - 1)/2$ possible connections among them. C is the fraction of these that actually exist, averaged over all nodes.

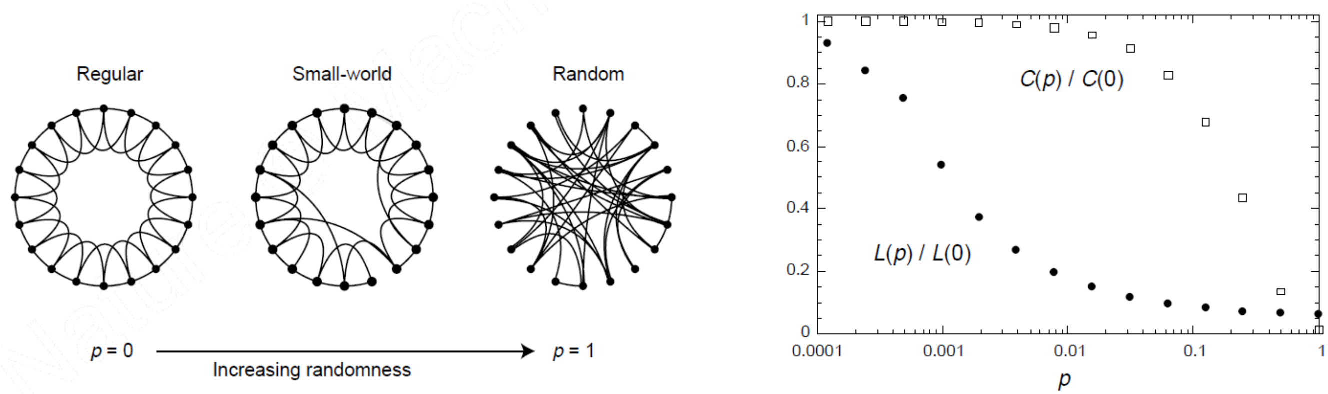


Image source: Watts & Strogatz, *Nature* 1998.

What's unique about the power grid as a complex network?

- The power grid is not a random network
  - It has been carefully designed over many decades
- Highly geographically constrained
  - Physically impossible to have substations connected to hundreds or thousands of transmission lines
- Building edges (lines) is expensive
- Reliability priority means that there are usually at least two lines into each transmission substation (N-1 reliability requirements)

The need for test cases in power systems research

- Some historic test systems exist (IEEE 14 bus, IEEE 300 bus) but in general these are small and do not represent actual, modern grids
- Power systems are critical infrastructure, and transmission grid data is often confidential
- Scientific rigor requires researchers to promote the reproducibility of results by publishing their data
  - but power systems data is often restricted by non-disclosure agreements
  - Even more so with more sensitive data like costs, geography, protection

Synthetic grids

- Synthetic grids are fictitious test cases that match characteristics of real grids but do not represent any actual grid
- This is a new, very active research area with several approaches under development
  - Geographic-security iterative (TAMU)
  - Small-World based topology-only approach (Wang)
  - Learning-based with spatial correlation (Soltan)
  - Fragmenting anonymizing actual grids (SDET)
  - Geographic-based two-stage approach (Comillas Spain, Wisconsin)

Synthetic power grids

- Large

This case is 10,000 buses (we are going up to 70,000)

- Complex

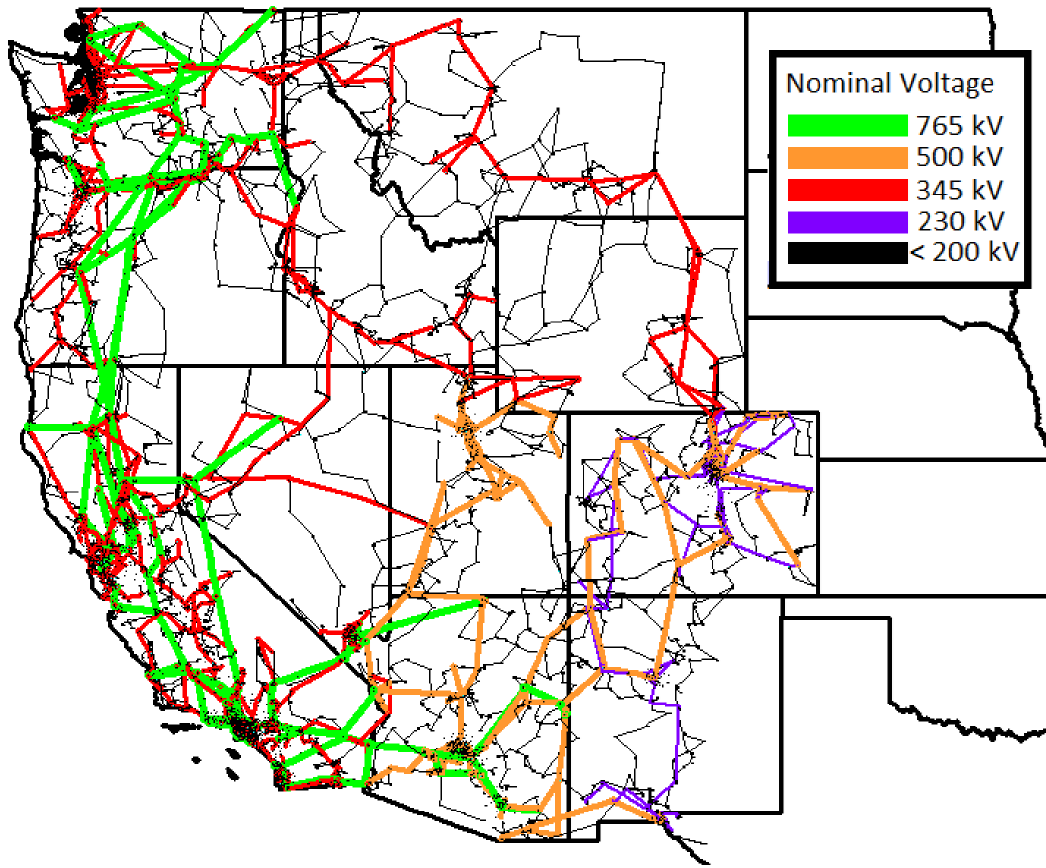
Multiple interacting voltage levels, remote regulation, and phase-shifters

- Realistic

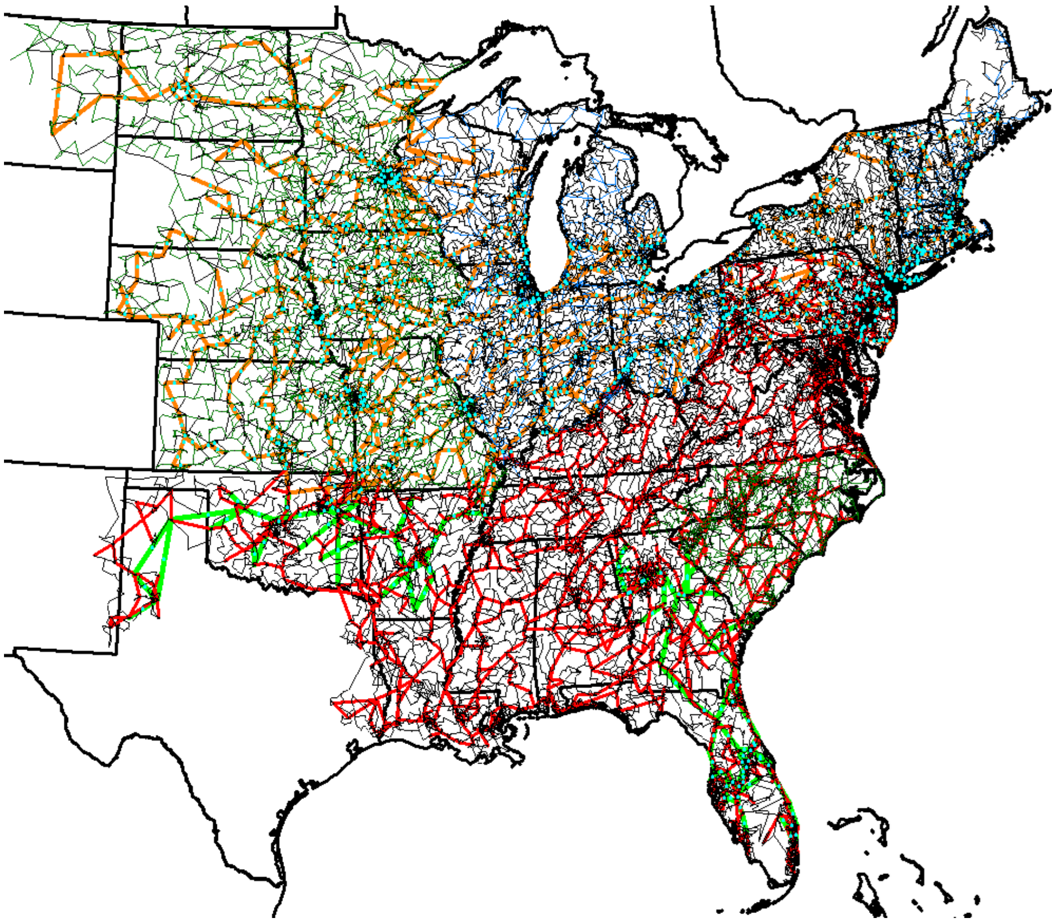
Matching a growing suite of valid-ation metrics against actual systems

- Fully public

It does not correspond to any actual grid or contain any confidential information



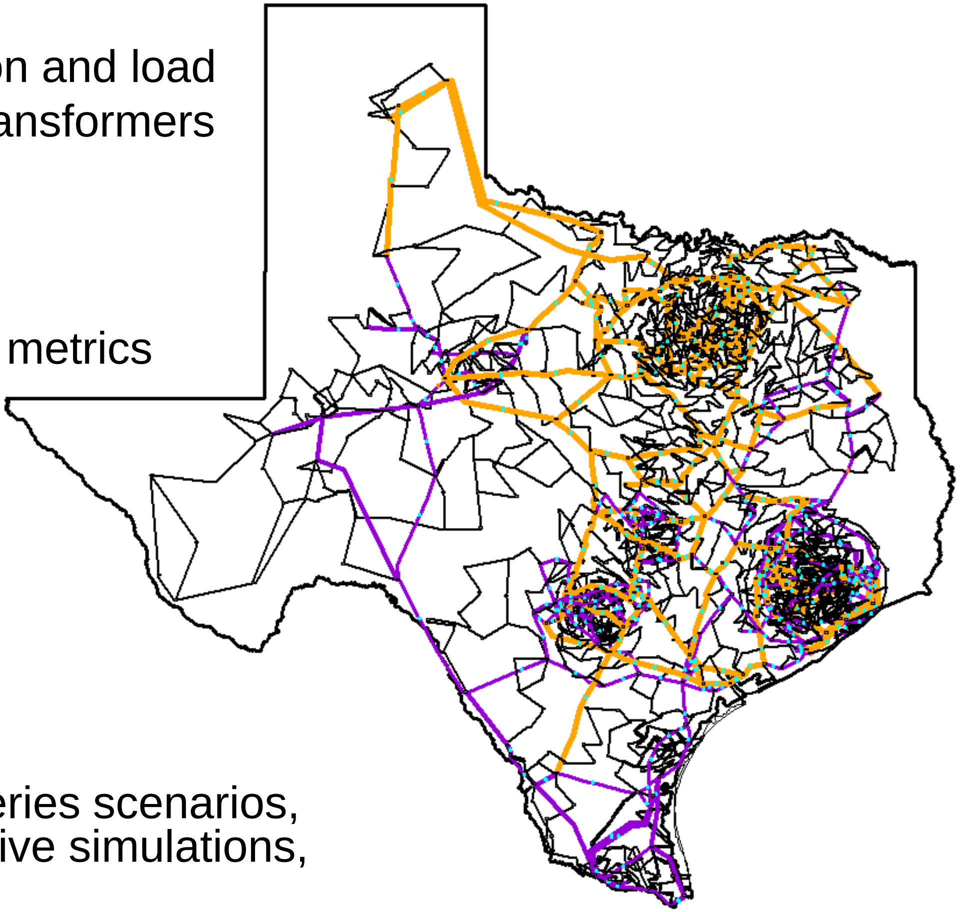
Building very large synthetic grids



- Built by TAMU team
- 70,000 bus synthetic grid on footprint of U.S. portion of eastern interconnect
- Grid has ac power flow solution, N-1 reliability, transient stability models, single-line diagrams

Building synthetic grids: overview

- Substation Planning
  - Start with public data for generation and load
  - Cluster substations, add buses, transformers
- Transmission Planning
  - Place lines and transformers
  - Iterative dc power flow algorithm
  - Match topological and geographic metrics
- Reactive Power Planning
  - Power flow solution (ac)
  - Voltage control devices
- Extensions
  - Transient stability
  - Geomagnetic disturbances
  - Single-line diagrams
  - Optimal power flow (OPF), time series scenarios, security constrained OPF, interactive simulations, …



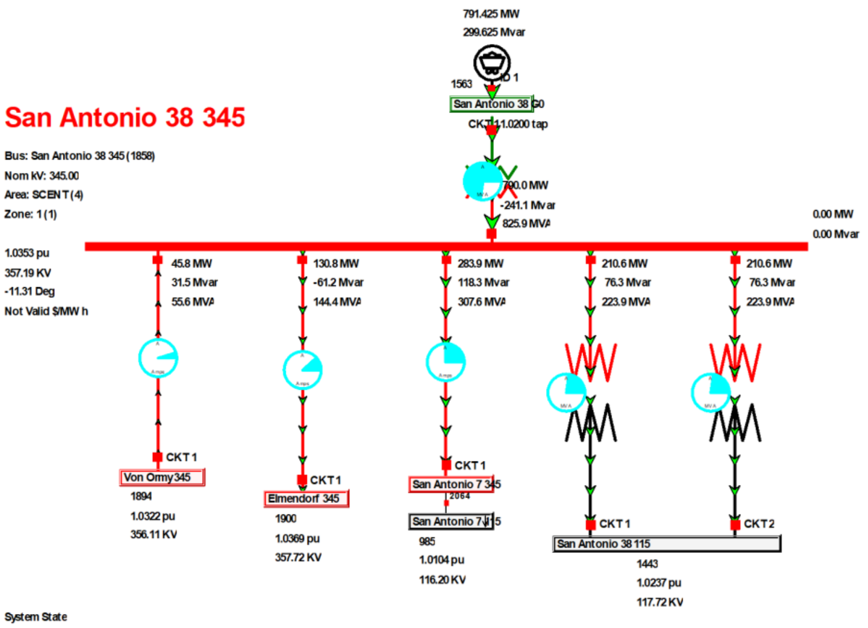
2000-bus synthetic grid on the Texas footprint

Substation Placement

- Generators come from U.S. Energy Information Administration
- Loads come from U.S. Census data, at zip code level
- Hierarchical clustering is used to condense substations
  - Metric is product of MW load or generation with geographic distance
  - Want to minimize for subs i clustering load/gens j:

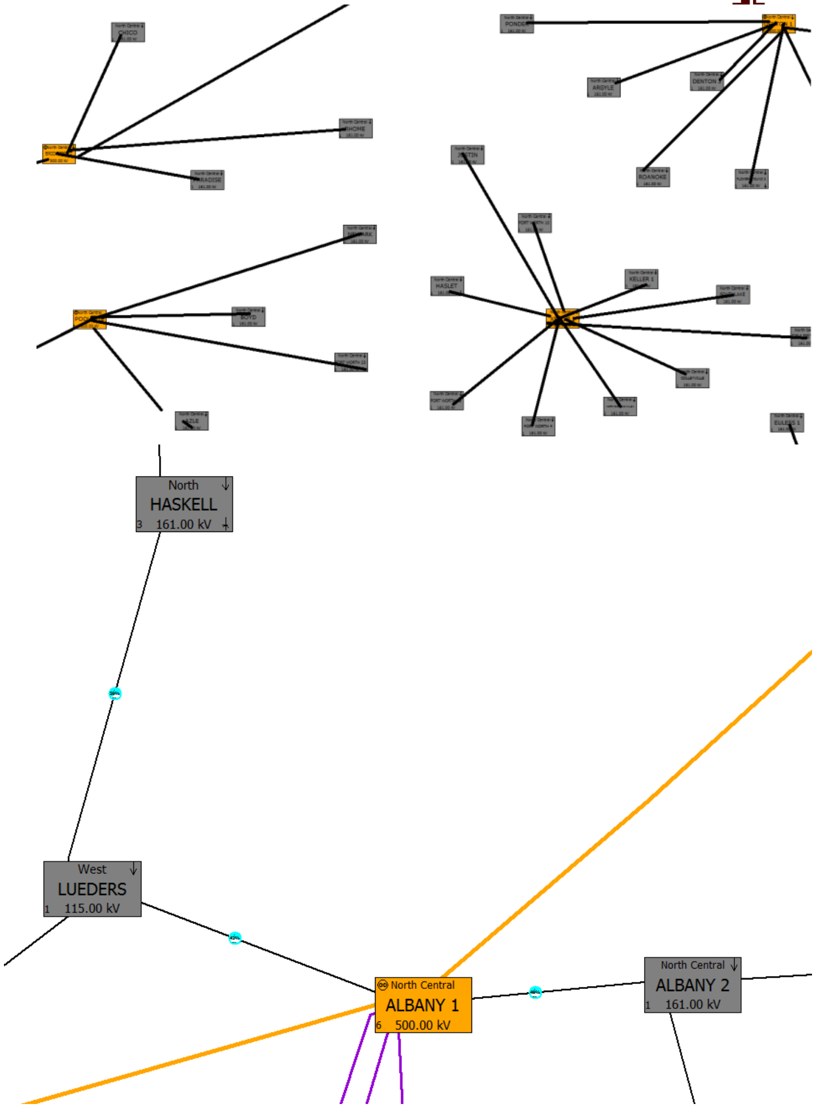
$$\sum_{i}\left(\sum_{j\in C_i}\left|P_{jL}\right|\cdot d_{ij}\right)$$

  - Balances density in urban areas



Assigning substation voltage levels

- Substation planning in large systems can be decoupled by area
- Each area is assigned 2-3 voltage levels
  - Most subs have lower voltage buses
  - 10-20% will have higher voltage buses
  - Higher load/generation more likely to have higher voltage buses
  - Modified hierarchical clustering
  - Need cross-area connection points for neighboring areas that do not share kV levels
- Generation dispatch
  - Assume peak planning load
  - Cost curves are determined statistically
  - Use economic dispatch
  - Adjust fuel costs per area as needed



Modified hierarchical clustering

- Input: list of N substations, target M
- Initialize each substation to be its own parent
- While # of parents > M:
  - Find least-cost parent to remove
  - Remove that parent and assign it to closest neighbor as parent
- Cost function to be minimized is:

$$C = \sum_{i=1}^{N} D_{ip} \cdot P_i$$

D_{ip} is the distance from substation to parent

P_i is the sum of generation and load at substation

Modified hierarchical clustering

- Hence the cost to remove one parent j is:

$$D_{jp} \cdot P_j + \sum_{c \in C_j} P_c \cdot \left(D_{cp} - D_{cj}\right)$$

  where p is the new parent to c
- The key computational bottleneck is calculating all the distances D : $O(n^2)$
  - But it can be done area by area (decoupled)
- This hierarchical approach is used both for the initial clustering, to make the substations, and the voltage clustering to identify the high-voltage subs

Building the transmission grid: key considerations

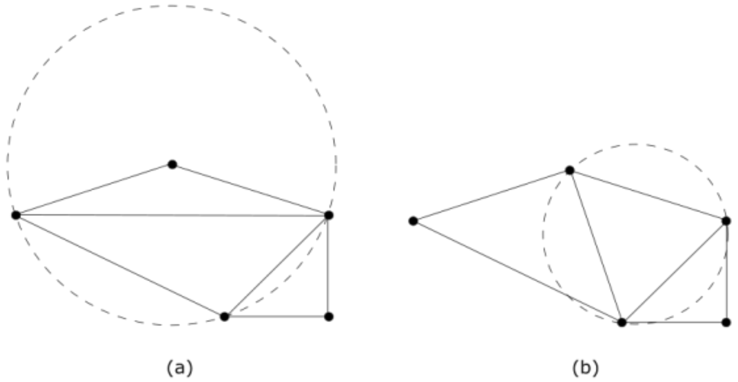
- Geography drives transmission planning, and is central to the approach (substations are geo-mapped)
- Network topology parameters
  - Graph metrics such as degree distribution, clustering, and diameter
  - Consider both individual voltage level networks and combined bus-branch topology
- Power flow feasibility
  - Avoid line limit violations
  - Consider contingency conditions

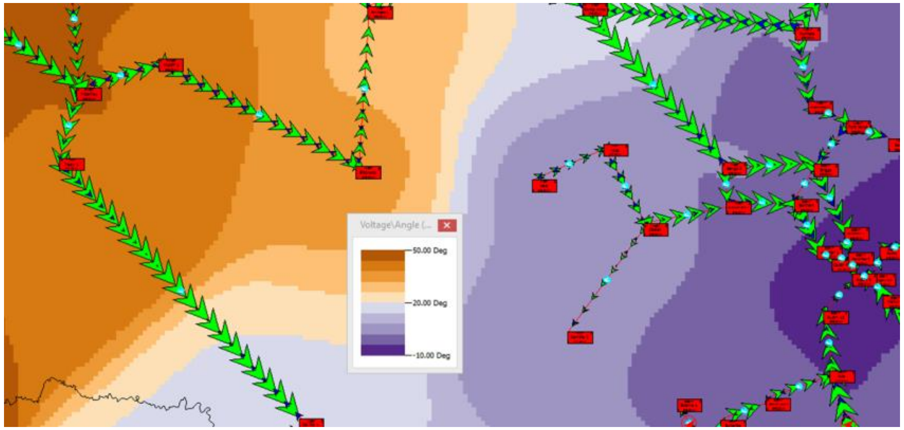
A difficult problem – especially for large grids

- Possible branches is n^2, possible combinations of branches is intractable
- Many competing metrics to meet (or objectives to optimize)
- Large grids have many overlapping voltage networks that can only connect at substations
- Consideration of different operating points and contingency conditions increases computation even more
- Manual adjustments grow with system size

Basic approach

- Reduce search space from n^2 to $21n$ with Delaunay triangulation (up to 3rd neighbors = 99% of lines)
- Begin with empty graph, add one branch to each network per iteration
- Estimate power flow in candidate lines with DCPF
- Only consider base case (peak load)
- Significant manual adjustment





Better, detailed approach

- Geographic constraints by voltage level
  - Delaunay triangulation reduces search space
  - Favoring shorter lines
- Depth first search to check connectivity
- DC Power flow N-1 contingency analysis, determine sensitivity of candidates lines to contingency overloads
- Iterative process of random removal, analysis, targeted addition for each same-voltage subnet

Derivation of N-1 security planning sensitivity calculations

- We want $\frac{\partial P_k}{\partial y_i} = X_k \frac{\partial \phi_k}{\partial y_i}$, that is, the sensitivity of the power P or phase angle difference ϕ across right-of-way (ROW) k with respect to the admittance y of ROW i.
  - If we add some admittance between one pair of buses, how does it affect an overload between another?
- DC power flow assumptions

$$Y\theta = P$$

$$d\theta = -Y^{-1}(dY)\theta$$

- For adding some admittance to ROW i

$$dY = e_i e_i^T dy_i$$

- Where e_i is 1 at from bus of ROW i, -1 at to bus, and 0 elsewhere

Derivation of N-1 security planning sensitivity calculations, cont.

- Thus,

$$d\theta = -Y^{-1} e_i e_i^T \theta dy_i$$

- We care about the difference between two bus angles

$$d\phi_k = e_k^T d\theta = -e_k^T Y^{-1} e_i e_i^T \theta dy_i$$

- But the red part is just ϕ_i, so

$$\frac{\partial \phi_k}{\partial y_i} = -e_k^T Y^{-1} e_i \phi_i$$

- You can switch k and i and the derivation is the same so

$$\frac{\partial \phi_k}{\partial y_i} = \frac{\phi_i}{\phi_k} \frac{\partial \phi_i}{\partial y_k} = \phi_i e_i^T (-Y^{-1} e_k)$$

- So for each monitored ROW k we only need 1 forward/backward substitution, and then 3 FLOPs per candidate line; so we can try lots of candidates.

Security constraints sensitivity

- Each subnet chooses the worst contingency overload to target in the addition step
- Calculate sensitivity to overload (e_{mon} has 1 at from bus and -1 at to bus of monitored)

$$\bar{s} = \boldsymbol{B}^{-1} e_{mon}$$

- Now each candidate line sensitivity is

$$s_{mon \to add} = (\bar{s}_{add1} - \bar{s}_{add2}) \cdot (\theta_{add1} - \theta_{add2})$$

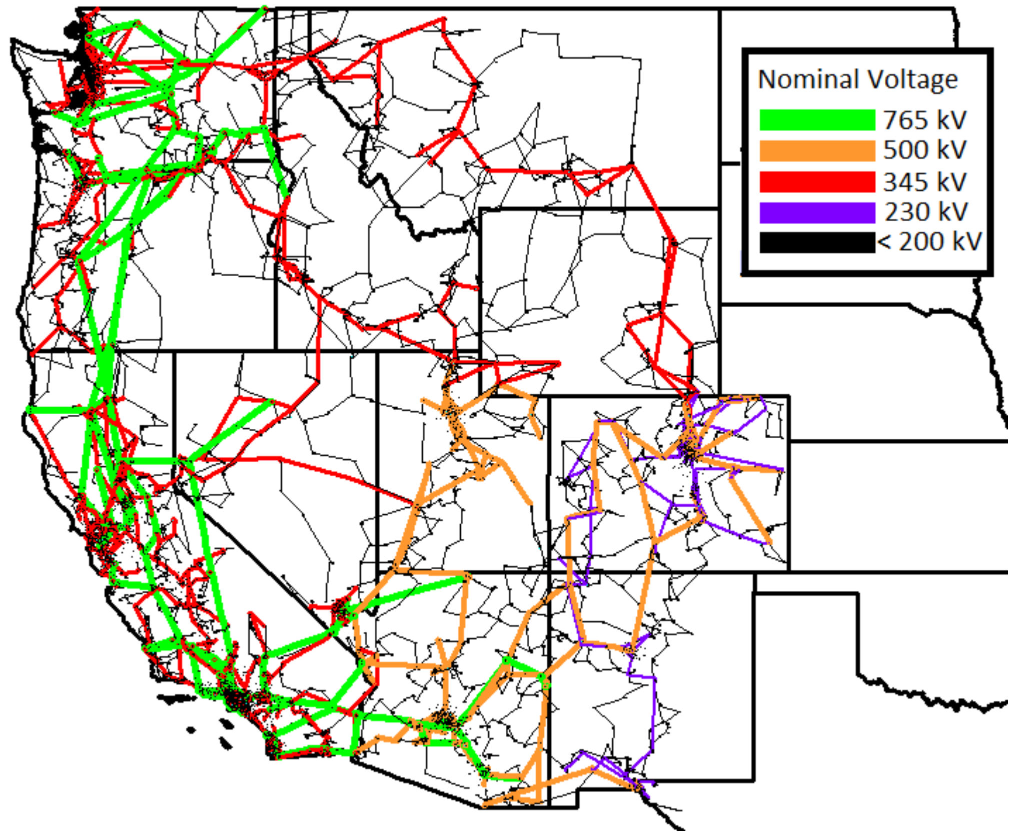
- Computational order is linear with number of candidate lines (three FLOPs each) so we can check all our potential lines for which ones contributed best to N-1 security

Practical considerations

- Stopping criteria and reward/penalty for each metric are the main parameters
- Line length and power flow are normalized so that the same penalty structure can be used for multiple voltage levels
- Other heuristics: reduce radial subs, forbid removing bridges, encourage biconnectivity
- Each voltage level is a network, with lower voltage levels divided into sub networks by area (cuts computation time)

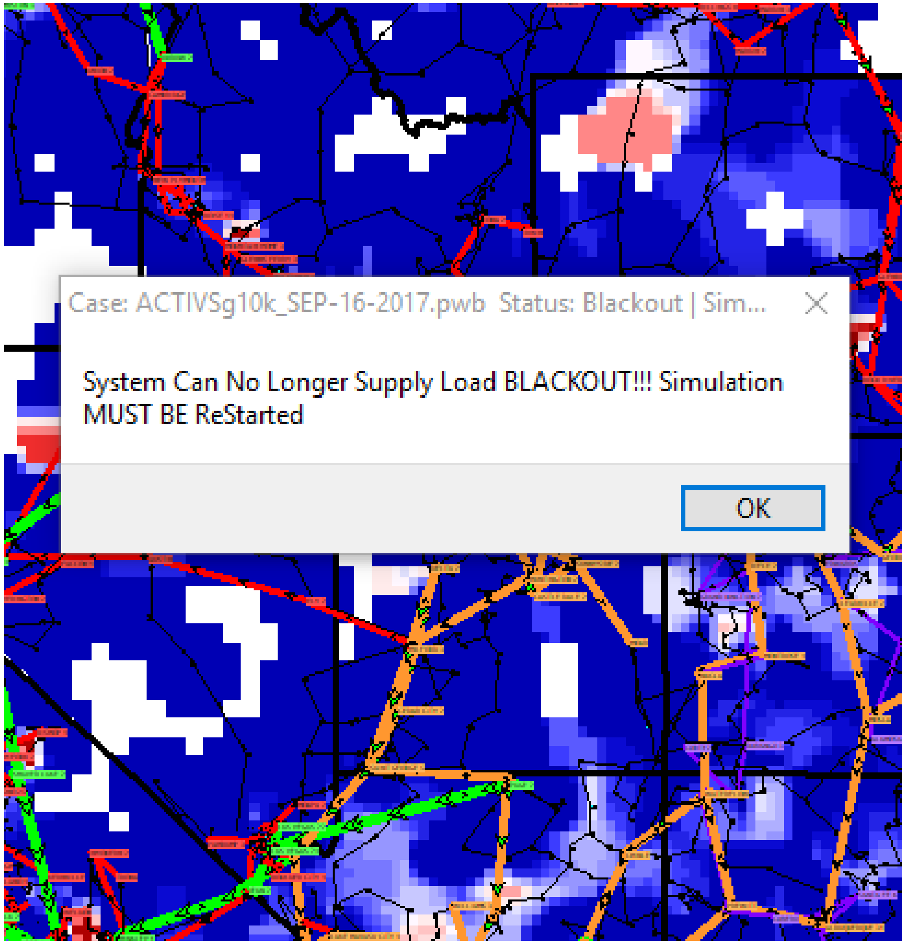
Results: large base cases

- Metrics are met more consistently
- Good dc solution
- Contingency performance is better
- Multiple voltage levels interconnect naturally
- Requires less manual adjustment



Motivation for reactive power planning: Large system power flow solution

- Flat start often does not converge!
- For real interconnects, just start with a previous solution
  - Doesn't work for new synthetic grids
  - Also synthetic grids, without reactive compensation, might not even have a solution
- So what do we do?



Work backwards

- Since we have a good dc solution, iteratively move from that to a realistic ac solution
- Step 1: Add a temporary generator to the highest voltage bus of every substation with 0 MW, controlling the bus voltage
  - Not every bus because they will fight each other
  - Start out all controlling to the same voltage
  - Some of these will eventually become shunts
- Step 2: Solve the ac power flow solution with this large number of PV buses (need to dispatch generators to anticipate 2-4% losses)

Iterative improvement

- Step 3, first stage iterations: remove most of the temporary generators
  - Divide into 100 groups, uniformly at random (since reactive power is localized)
  - Do 100 iterations of the ac power flow
    - Remove most of the generators in the group
    - Do the new solution
      - If it diverges, restore all the temporary generators
    - Add a few back to fix voltage violations

Removing temporary generators

- Point system used to decide which generators to remove in first stage

Condition	Points
Nominal kV < 200	2
Nominal kV < 400	2
Substation generation > 100 Mvar	1
Substation generation > 10 Mvar	1
No. tie lines = 2	1
At least one tie line is sending MW	1
Nearest Q resource is 1 hop away	3
Nearest Q resource is 2 hops away	2
Nearest Q resource is 3 hops away	1
Second-nearest Q resource is 1 hop away	5
Second-nearest Q resource is 2 hops away	4
Second-nearest Q resource is 3 hops away	3
Second-nearest Q resource is 4 hops away	2

Matching the voltage schedule

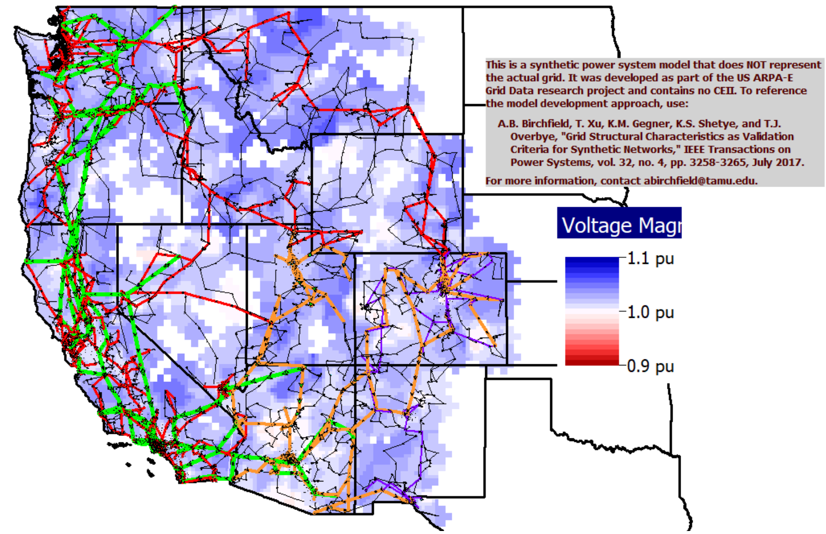
- Step 4, second stage iterations: adjust the generator voltage setpoints and transformer taps (LTCs)
  - Select which transformers have LTCs according to statistics
  - Do twenty iterations (LTCs have +/- 16 steps)
  - At each step, each substation adjusts its generator set points and LTC taps

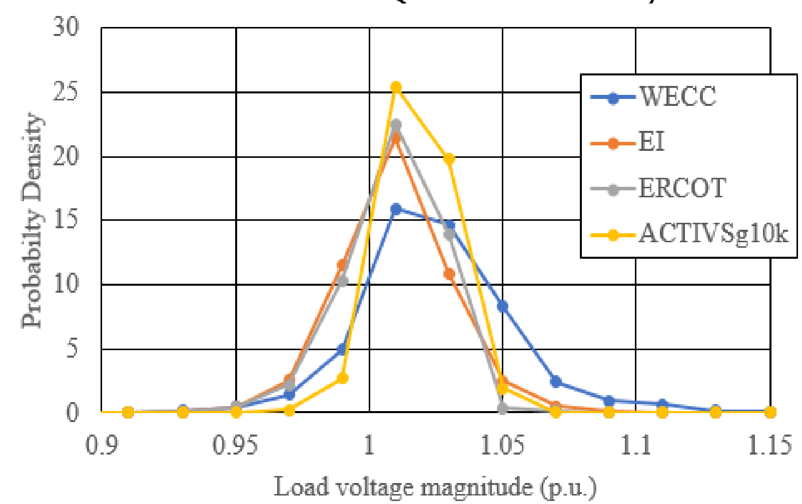
Substation voltage schedule

- Each substation is initialized with a scheduled voltage magnitude in the range [1.035, 1.045]
  - Higher voltage buses have a slightly higher schedule
- At each iteration, try to match this with small changes to control parameters
  - Temporary generators (shunts) and actual generators at network buses: gradually adjust set point in discrete steps
  - Network transformers: 0.00625 step changes in tap, to control the low-side bus voltage
  - GSUs with taps: control the high-side voltage if no other devices is controlling it, otherwise work with corresponding generator to regulate generator Mvar output
- As a post-processing step, convert temporary generators to shunts and have actual generators regulate the GSU high-side bus

Example: 10K Case

- Initial power flow solution diverges!
- Algorithm previously described was applied
- 387 shunt capacitors remained for 4762 substations
  - This is 8%, actual grid has 10-20% (good)
- Voltage profile matches actual interconnect observations





Validating synthetic grids

- Validation is key in building synthetic grids!
- Key question is how to quantify what makes a power grid realistic.
  - Thus our approach is to define validation metrics
- Because of the variety in engineering design and practice, actual grids are quite diverse
  - Challenge is to capture metric distribution and realistic ranges
  - Thresholds are set that should not be violated unless justified with an engineering design choice. Thus this is a screening process.

TAMU approach to validation

- Anchored in statistical analysis of 3 actual North American interconnects and 12 subset cases, from FERC 715 data
- All created synthetic cases are designed to meet these validation metrics
- Categories of metrics
  - Size and structure: ratio and proportions of elements
  - Parameter distributions and correlation
  - Topological network structure
  - Power flow metrics and voltage control devices

List of validation metrics

- Overall size and structure
  - Number of buses per substation
  - Substation voltage levels
  - Percent of substations containing load
  - Load at each bus
  - Ratio of total generation capacity to load
  - Percent of subs containing generation
  - Capacities of generators
  - Percent of generators committed
  - Generator dispatch percentage
  - Generator reactive power limits
  - Load power factor
- Branch parameters
  - Transformer per-unit reactance
  - Transformer MVA limit and X/R ratio
  - Transmission line reactance
  - Transmission line X/R ratio and MVA limit
- Network topology
  - Ratio of lines to substations
  - Percent of transmission lines on the minimum spanning tree
  - Distance of transmission lines along the Delaunay triangulation
  - Ratio of total length of all lines to length of minimum spanning tree
  - Cycles
  - Surge impedance loading
- Voltage control and complexities
  - Voltage profile
  - Reactive power elements
  - Three-winding transformers
  - Impedance correction tables
  - Transient stability
  - Substation grounding resistance

Example validation metric: X/R ratios and MVA limits

- For line and transformer X/R ratios and MVA limits, the branches are categorized by nominal voltage level, and a percentage of parameters must fall within 10/90 percentiles.

Transformer High Voltage Level (kV)		MVA Limit			X/R Ratio		
		10%	Median	90%	10%	Median	90%
EI system	69	10	42	115	10	20	50
	115	22	53	140	16	25	48
	138	39	83	239	19	30	54
	161	48	100	276	18	32	68
	230	63	203	470	25	44	84
	345	200	444	702	35	60	157
	500	215	812	1383	44	70	119

Line XR Ratios

Voltage Level (kV)	90%	Median	10%
500	26.0	17.0	11.0
345	16.0	12.0	9.0
230	12.5	9.0	6.4
161	10.0	6.0	4.1
138	9.1	5.7	3.0
115	8.3	4.6	2.5

Line MVA Limits

Voltage Level (kV)	90%	Median	10%
500	3464	2598	1732
345	1494	1195	897
230	797	541	327
161	410	265	176
138	344	223	141
115	255	160	92

Example validation metric: Delaunay Triangulation

- Delaunay neighbor ratios are consistent among networks, and capture geographic and topological properties together.
  - Implies degree distribution, clustering coefficient

Transmission Line Category	Average percentage for Eastern Interconnect	Average percentage for Western Interconnect
Minimum Spanning Tree	47.8%	44.3%
Delaunay Triangulation	75.6%	71.1%
Delaunay 2 neighbor	18.3%	21.5%
Delaunay 3 neighbor	4.6%	5.3%
Delaunay 4 neighbor	1.1%	1.4%
Delaunay 5+ neighbor	0.4%	0.7%

Complex network properties

Metric	Actual Systems			Synthetic Systems		
	EI	WECC	ERCOT	70K	20K	5000
n	36,187	9398	3827	34,999	9524	2941
$\bar{d}$	2.61	2.58	2.61	2.74	2.67	2.71
$\bar{c}$	0.044	0.058	0.032	0.048	0.034	0.031
$\bar{\ell}$	29.2	18.9	14.2	36.7	20.3	13.8
$\bar{b}$	0.083	0.21	0.40	0.11	0.22	0.50

n : number of substations

$\bar{d}$: average node degree

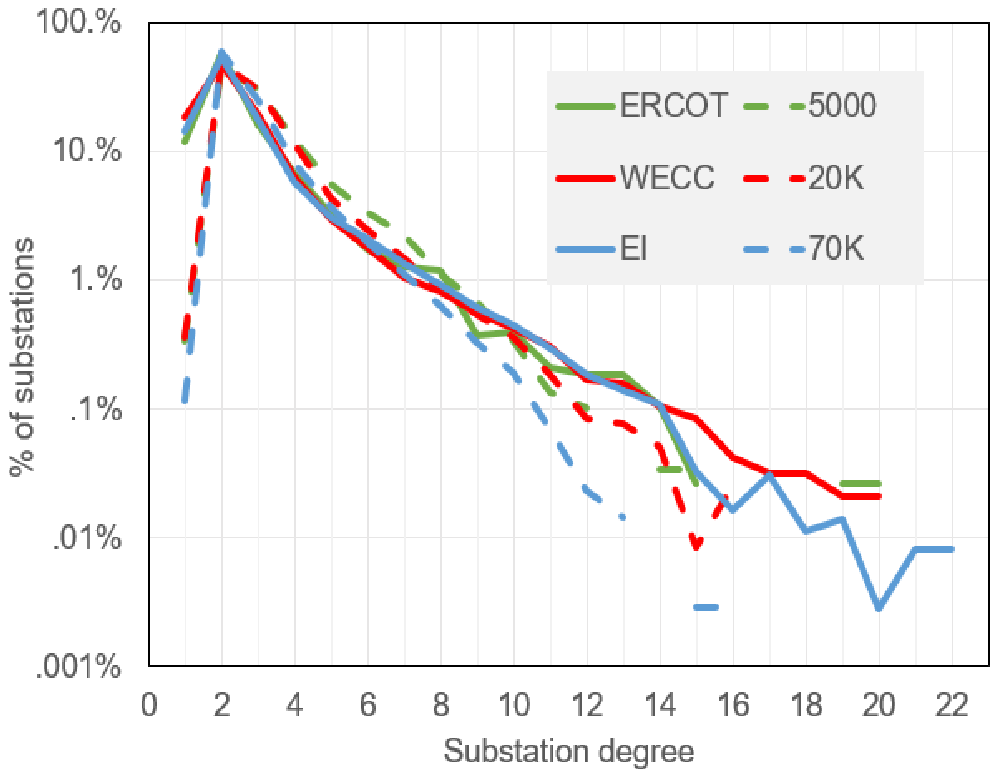
$\bar{c}$: clustering coefficient

$\bar{\ell}$: average shortest path

$\bar{b}$: average betweenness centrality measure (%)

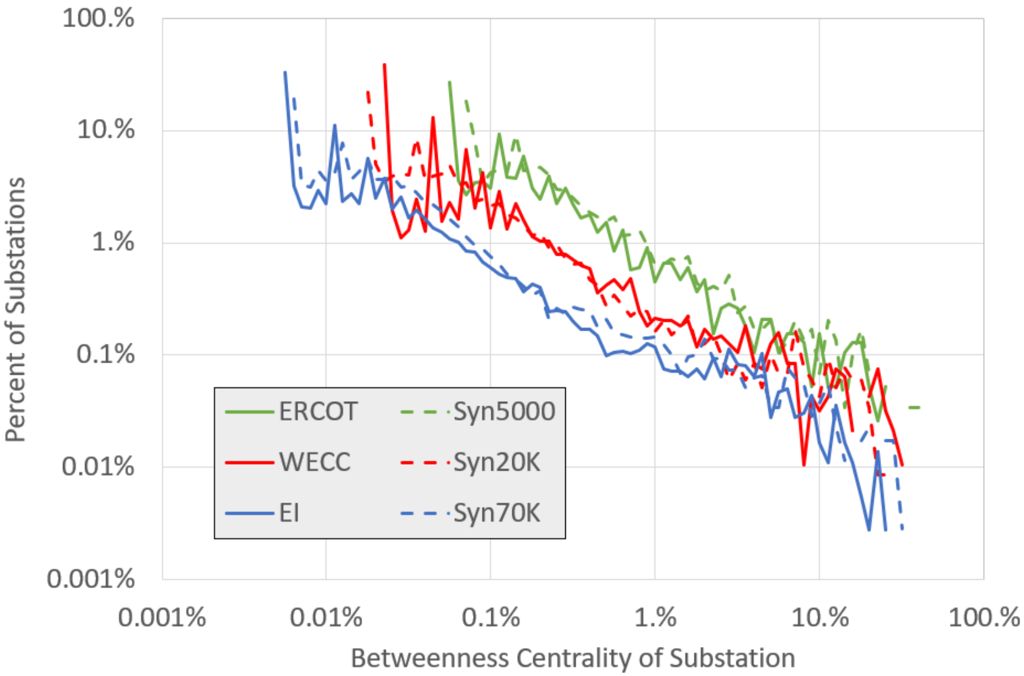
Degree distribution

- Measures how many lines attach to each substation
- Exponential distribution (logarithmic vertical axis)
- Horizontal axis is not logarithmic (as in scale-free)
- Does not change much with system size
- Degree=1 nodes are less than expected

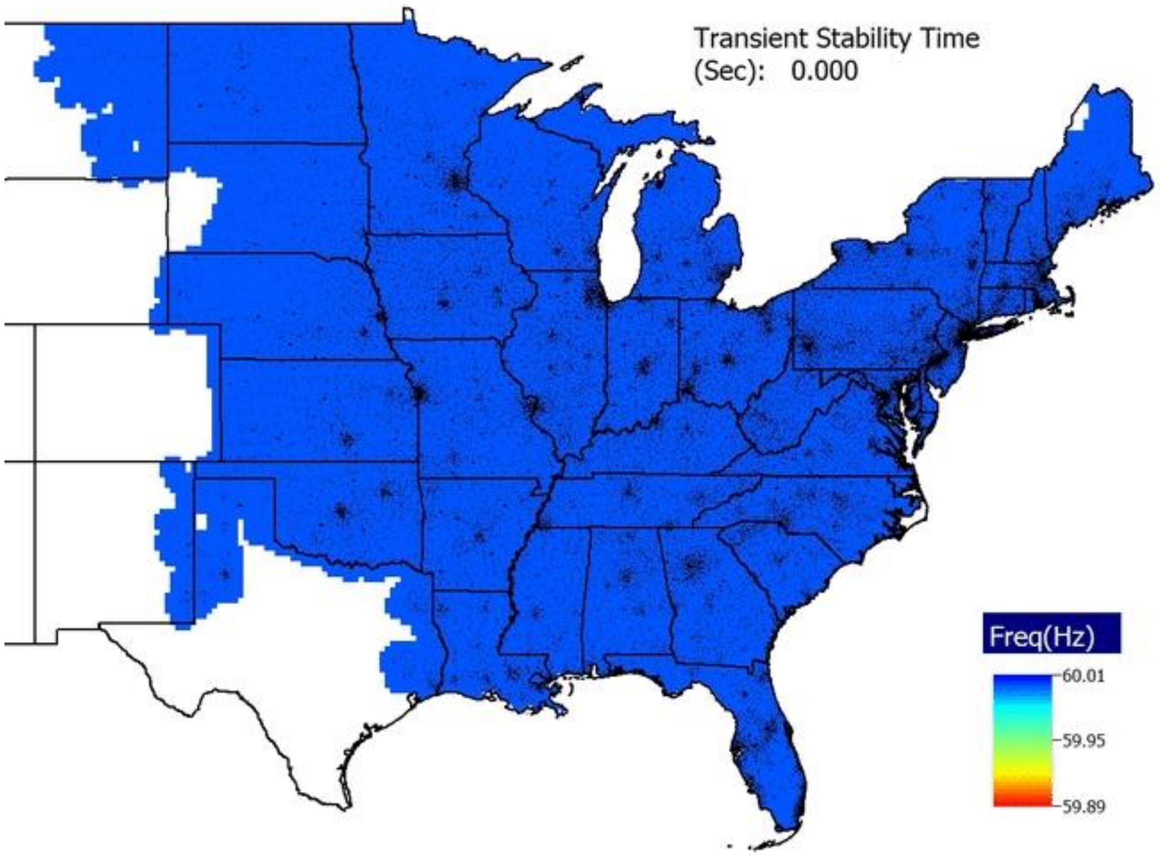


Betweenness centrality

- For all the n^2 shortest paths between pairs of vertices, how many pass through a given vertex?
- For some nodes, as much as 25%
- Similar properties on a log-log plot



Transient stability

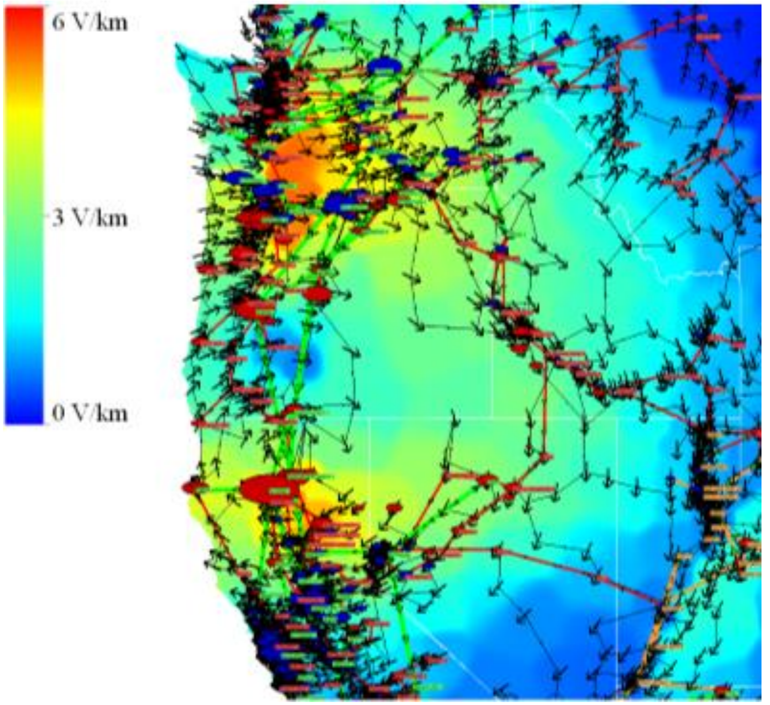


These cases can be extended with dynamic models for transient stability studies that match statistics and dynamic behavior of actual grid.

These are used to produce synthetic PMU data that can be publicly released, along with the underlying model.

Geomagnetic disturbance parameters

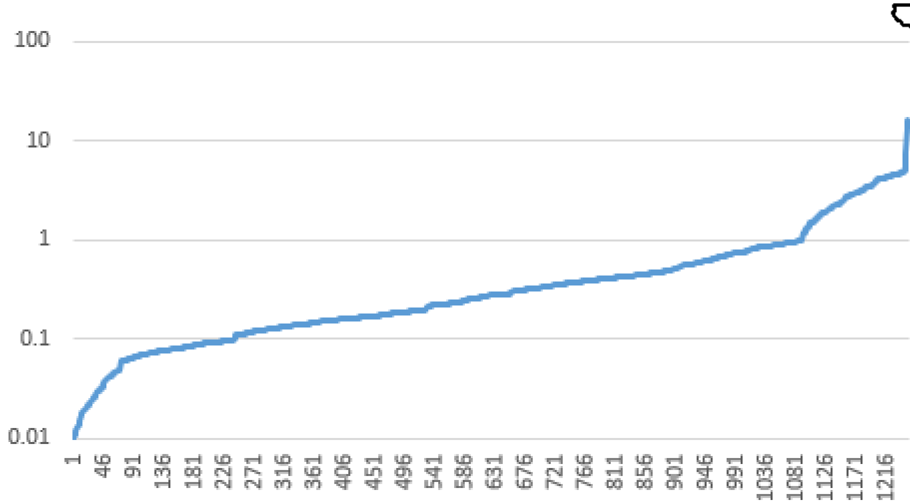
- Geomagnetic disturbances (GMDs) from solar storms and electromagnetic pulses (EMPs) from high-altitude nuclear blasts can induce quasi-dc GICs on power transmission systems
- Synthetic grids can be augmented with data necessary to perform these studies
- Particularly useful for EMPs, due to sensitive nature of the topic
- GIC studies require substation definitions and geo coordinates, which are not usually included with system models or test cases
  - But we already have them for synthetic grids



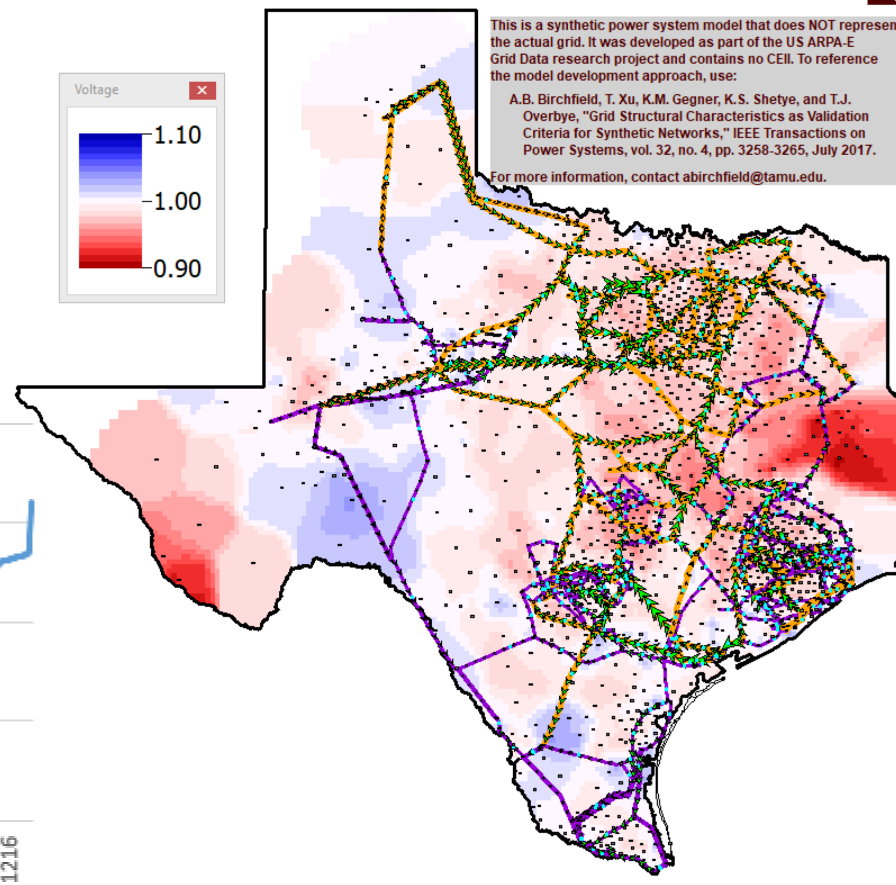
Contour shows electric field magnitude, with arrows showing direction and ovals showing induced transformer current to EMP.

GMD Parameters for 2000 Case

- Substation grounding added
- Transformer parameters
- Run uniform GMD analysis



Ordered substation grounding in Texas case

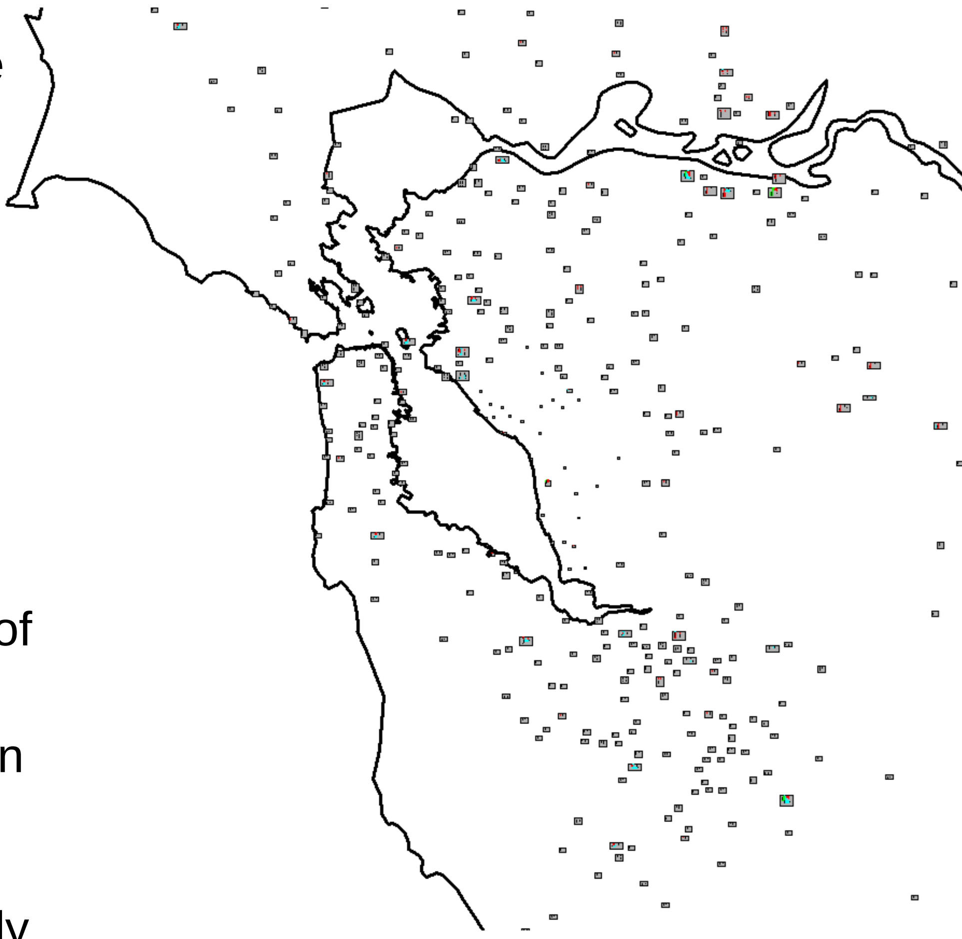


The voltage begins to sag significantly under a 5 V/km field

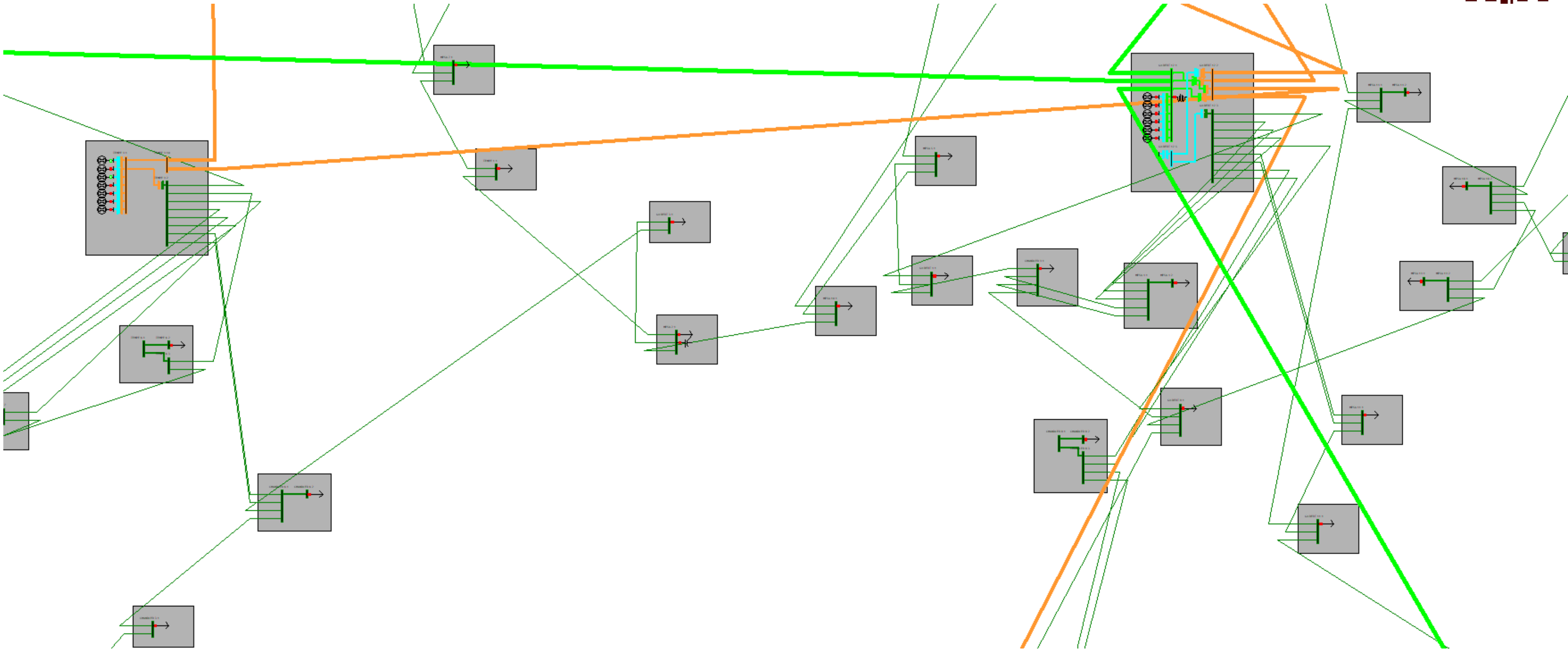
Automatic one-line diagrams

Greedy approach to substation spacing

- Challenge is that we want to balance geographic correctness with readability
- Two approaches:
  - First is force-directed, which models substations as particles with Coulomb/Hooke forces interacting
  - Better is greedy approach:
    - In dense areas, reduce size of substations
    - Iteratively put each substation as close as possible to original location while respecting spacing for already placed substations

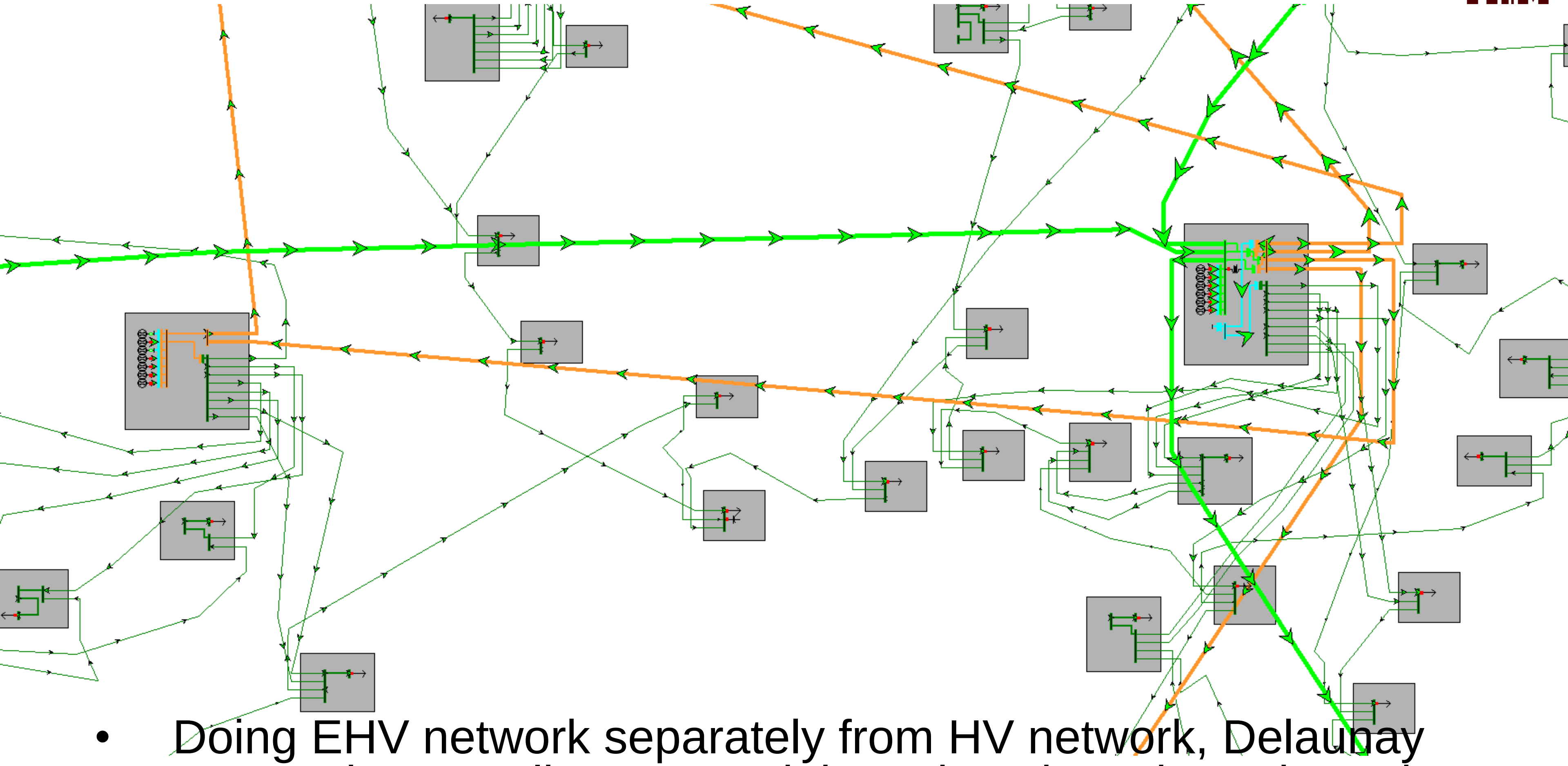


Transmission line routing



- Straight line approach can cause confusion and be difficult to read
- Objective is to minimize lines overlapping substations

Two-layer Delaunay approach



- Doing EHV network separately from HV network, Delaunay approach routes lines around the substations through routing channels
- Two layers keeps EHV lines from having too many bends