Topological Machine Learning Methods for Power System Responses to Contingencies

Brian Bush,[1] **Yuzhou Chen,**[2,5] **Dorcas Ofori-Boateng,**[3] **Yulia R. Gel**[4,5]

[1] Strategic Energy Analysis Center, National Renewable Energy Laboratory, Golden, CO 80401, USA
[2] Department of Statistical Science, Southern Methodist University, Dallas, TX 75275, USA
[3] Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR 97201, USA
[4] Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080, USA
[5] Energy Storage and Distributed Resources Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
{brian.bush@nrel.gov, yuzhouc@smu.edu, doforib2@pdx.edu, ygl@utdallas.edu}

Abstract

While deep learning tools, coupled with the emerging machinery of topological data analysis, are proven to deliver various performance gains in a broad range of applications, from image classification to biosurveillance to blockchain fraud detection, their utility in areas of high societal importance such as power system modeling and, particularly, resilience quantification in the energy sector yet remain untapped. To provide fast acting synthetic regulation and contingency reserve services to the grid while having minimal disruptions on customer quality of service, we propose a new topology-based system that depends on neural network architecture for impact metrics classification and prediction in power systems. This novel topology-based system allows one to evaluate the impact of three power system contingency types, namely, in conjunction with transmission lines, transformers, and transmission lines combined with transformers. We show that the proposed new neural network architecture equipped with local topological measures facilitates both more accurate classification of unserved load as well as the amount of unserved load. In addition, we are able to learn complex relationships between electrical properties and local topological measurements on the simulated response to contingencies for NREL-SIIP power system.

Introduction

Power-system simulations require detailed technical data on system components and involve significant non-linearities, so it is unclear whether topological measurements alone on an unperturbed power system (the "base case") will be highly correlated with the prediction of the impact of outages upon the system's operation, particularly the impact in terms of customer load that is unserved due to the contingency.

Oftentimes, resilience quantification involves the search over high dimensional spaces of system configurations, responses, and impacts using technical models, simulations, and analyses that typically are data and computation intensive. The regularly used brute-force approaches that enumerate and simulate each potential configuration of interest generally are not computationally efficient or even fully viable. In turn, AI and ML offer an opportunity to move beyond traditional techniques and to effectively explore the high dimensional space of system configurations and to create reduced-form approximations of complex power system models or the non-linear simulations involved. A classic example is the process of estimating the effect of contingencies (i.e., component outages) upon a power system's ability to meet customer demand. Power-system modeling requires detailed electrical data on the system and involves substantial computational resources, limiting the number of contingencies that can be routinely directly simulated. Machine-learning methods that are able to accelerate either the search or simulation of these models promise to substantially enlarge the number and interest of the contingencies that can be analyzed, at a fraction of the traditional computational burden.

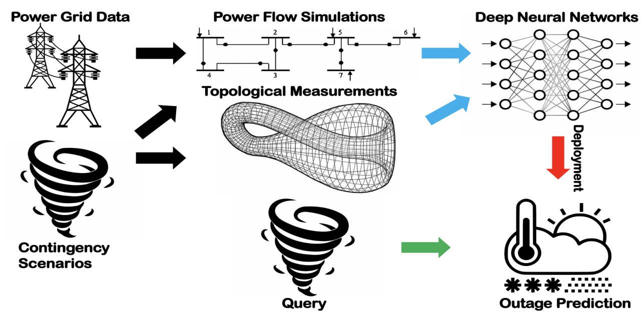

Figure 1: Illustration of training and applying contingency screening.

The purpose of this study is to assess these questions:

1. To what extent can basic graph-theoretic and topological measurements be used to qualitatively and quantitatively estimate the impacts of contingencies?
2. Which machine learning techniques are best suited for this?
3. Which graph and topology measurements are most informative?

4. By how much can measurements of static electrical properties enhance predictive power of topology-based models?

Related Work

Power Systems and Their Resilience Generally, there exist two main interrelated approaches for resilience quantification of power grid networks. The first direction is based purely on the topological properties of a grid network, and the second (hybrid) approach aims to incorporate electrical engineering (EE) concepts, such as impedance or blackout size into the complex network (CN) analysis. Both approaches offer important complementary insights on the hidden mechanisms behind the power system functionality, and neither approach can be viewed as a universally preferred method (Hines et al. 2010; Cuadra et al. 2015; Mohamed et al. 2019; Abedi, Gaudard, and Romerio 2019; Nakarmi et al. 2020; Bose, Chanda, and Chakrabarti 2020; Vugrin, Castillo, and Silva-Monroy 2017). In the CN context, the metrics for grid network vulnerability are mostly built with either the giant component, average path length or node degree distribution. The EE approach, on the other hand, utilizes power-flow models to simulate the effects of component failure or other forms of disruptions: a variety of approximations to the power-flow equations have been used (Abedi, Gaudard, and Romerio 2019). However, nearly all currently deployed resilience metrics are based on the global structure of power grid networks and tend to disregard important local structure information of the grid in its capacity to sustain adverse events.

AI for power systems AI tools are increasingly widely being used in power system analysis (Zhao et al. 2019; Jufri, Widiputra, and Jung 2019; Ibrahim, Dong, and Yang 2020), including, for instance, semi-supervised learning for load monitoring and scheduling (Gillis and Morsi 2016), Support Vector Machine (SVM) and deep learning for fault detection (Eskandarpour, Khodaei, and Arab 2017; Oh et al. 2017; Jenssen, Roverso et al. 2019), and Bayes networks and meta-action for cascading failure propagation (Pi et al. 2018; Huang et al. 2019). Nevertheless, as noted by Shuvro et al. (2019); Jufri, Widiputra, and Jung (2019), AI and ML algorithms remain underused in power system analysis. Essentially, the application of AI tools for resilience quantification of power grid networks is essentially in its infancy.

Topological Data Analysis (TDA) The recent decade has witnessed a steep rise in implementation of TDA machinery in diverse application settings including image detection (Asaad and Jassim 2017), system robustness analysis (Li, Ryerson, and Balakrishnan 2019) and spatial classification (Feng and Porter 2020). Most recently, Islambekov et al. (2018); Li et al. (2020) considered topological summaries as alternative metrics for transmission grid resilience assessment. However, studies that analyze the contribution of topological measurements quantitatively and qualitatively to contingency impact evaluation, or the study that harness TDA machinery and AI algorithms for predicting the power grid response to contingencies are yet to be explored.

Modeling Power Grids as Complex Networks

We use the following power-system terminology: A *bus* is a network node where lines, transformers, loads, and generators connect. The *flow* is the quantity of power passing through a network component. The *capacity* is the maximum flow that a network component can support. The *residual capacity* is the difference between the capacity and the flow. The *base-case flow* is the flow in the case where all components are in service.

Power grid systems can be inherently viewed as graph structure $\mathcal{G} = (V, E, \omega)$ where node set V represents buses, transformers and generators, while edge set E is formed by transmission lines and transformers. To incorporate the electrical properties (reactance, resistance, etc.) and power-flow characteristics (real and reactive power through components) of the grid, we use *(edge)-weighted* function $\omega : E \mapsto \mathbb{R}_{\geq 0}$ such that each edge $e_{uv} \in E$ has a weight ω_{uv}. Examples of weight functions for power grid networks include, for example, geographical distances, admittance, average power flow, and electrical conductance (Dörfler 2013; Sánchez-García et al. 2014; Xu, Gurfinkel, and Rikvold 2014). Quantifying grid connectivity with $\omega_{uv} \equiv 1$ for $e_{uv} \in E$ corresponds to an unweighted graph G, and in order to distinguish this from other weight kinds we relabel the variable ω_{uv} as $\omega_{uv}^{unweighted}$. In the weighted graph case, we work with the following as edge weights:

- Reactance weight: $\omega_{uv}^{\text{reactance}} = \mathcal{X}_{uv}$, where $\mathcal{X}_{uv}$ is the per-unit reactance of the line or transformer between u and v.
- Capacity weight: $\omega_{uv}^{\text{capacity}} = C_{uv}^{-1}$, where C_{uv} is maximum flow rating (in MW) of the line or transformer.
- Base-flow weight: $\omega_{uv}^{\text{base}} = \left(F_{uv}^{(0)}\right)^{-1}$, where $F_{uv}^{(0)}$ is the flow solution (in MW) across the line or transformer in the base (i.e., no contingency) case.
- Residual capacity weight: $\omega_{uv}^{\text{residue}} = \left(C_{uv} - \left|F_{uv}^{(0)}\right|\right)^{-1}$.

Power grid networks are imperiled by many factors, for example, extreme weather, human errors, and cyber-physical attacks. Such adverse events result in contingencies. Formally, a *contingency / outage* is a state where some network components are no longer in service, and a contingency set ($O_i \subset \{uv\}$ for scenario i) is a list of edges that are removed (outaged) from the graph in the event of the contingency. Generally, we label the number of contingencies as $N_i = |O_i|$, and denote the types of components in contingencies for scenario i as $t_i \subset \{\text{line}, \text{transformer}\}$. By principle, the total edge-weight lost in scenario i: $L_w^{(i)} = \sum_{uv \in O_i} \omega_{uv}^w$, where $w \in \{\text{unweighted}, \text{reactance}, \text{capacity}, \text{base}, \text{residue}\}$.

With D_u as the customer demand at vertex u, total demand will be $D^{\text{total}} = \sum_u D_u$, the power-flow solution for consumption at vertex u in scenario i as $S_u^{(i)}$, and the unserved demand/load (i.e., the shortfall between demand and actual consumption) in scenario i is $U^{(i)} = \sum_u \left(D_u - S_u^{(i)}\right)$. Armed with these notations, we define consumption (i.e., the amount of demand successfully served) in scenario i as $S^{(i)} = \sum_u S_u^{(i)}$, and make $I^{(i)} = 1_{U^{(i)}>0}$ the indicator variable for whether there is unserved demand in scenario i. The *unserved load* is the amount of customer load not being served in a contingency.

In this paper, contingencies are generated by removing from service a specific fraction (f_i) of the network edges (i.e., lines and transformers) within a subgraph of a given radius (r_i) around a randomly selected vertex. The radius of the subgraph is varied from 0 to 25 edges and the density of out-of-service components in this class is varied from 1% to 100%. This contingency-subgraph approach makes it possible to explore the impact of clustering the out-of-service components upon the operation of the power system. Note that the case of r_i of 0 and 100% outages simply correspond to uniformly and randomly selecting a single outage edge, whereas non-zero radii and less-than-complete selection of outage components correspond to localized outages of varying size and intensity.

Again in our case study, the experimental design varies the radius in integer increments up to 25 edges, considers the outage fractions 1, 2, 5, 10, 20, 50, and 100%, and samples 5000 randomly-chosen vertices for the center of the contingency subgraph. The dataset, which is available at Bush (2020), contains 281,273 solved power-flow simulations.

Preliminaries on Topological Data Analysis

The systematic evaluation of patterns and dynamics of multiscale network topological and geometric properties can be approached via an algebraic tool, based on the adaptation of a homology theory to applied data analysis known as *persistent homology* (Carlsson 2019). In brief, the idea is to capture the cycle of features which are long-lived, or persist over varying thresholds $\nu_j > 0$. Such persistent features are much likelier to significantly impact power grid network functionality.

Let an edge-weighted graph (G, ω) be a representation of a power-flow network. If we select a certain threshold $\nu_j > 0$ and keep only edges with weights $\omega_{uv} \leq \nu_j$, we obtain a graph $G_j \subseteq G$ with an associated adjacency matrix $A_{uv} = \mathbb{1}_{\omega_{uv} \leq \nu_j}$. By varying the threshold values in the order $\nu_1 < \nu_2 < \ldots < \nu_n$, we obtain a hierarchical nested sequence of graphs $G_1 \subseteq G_2 \subseteq \ldots \subseteq G_n$ that is known as a *network filtration*. With this filtration, we can associate each $G_j, j = 1, \ldots, n$ with a combinatorial structure called an *abstract simplicial complex*, and this allows to quantify the underlying geometry of (G, ω). From standard literature, the Vietoris-Rips (VR) combinatorial complex is the most popular complex choice in TDA due to its simplicity and computational benefits (Carlsson 2009; Zomorodian 2010). In particular, if we view weights ω_{uv} as some "distance" measure between nodes u and v, then a Vietoris-Rips complex at threshold ν_j is defined as $VR_j = \{u, v \in V | \omega_{uv} \leq \nu_j\}$. That is, VR_j contains all the sets of k node features, $k = 1, \ldots, K$, which are pairwise connected by an edge to form simplices of dimension $k-1$ (i.e., an edge for 2 nodes, a triangle for 3 nodes, etc).

Now, equipped with the associated *VR filtration*, $VR_1 \subseteq VR_2 \subseteq \ldots \subseteq VR_n$, we can trace the incidence and survival of qualitative features of various shapes such as connected components, triangles, tetrahedra, and higher-order structures, in terms of when they appear and disappear with an increasing threshold ν_j. In turn, analysis of evolution and lifespan of such topological features provides a multiscale quantitative insight into network topology and its role in network organization, functionality, and interactions among network components under varying operating conditions. We can quantify persistent features and their lifespans in power grid networks using *Betti* numbers and *persistence diagrams* (PD).

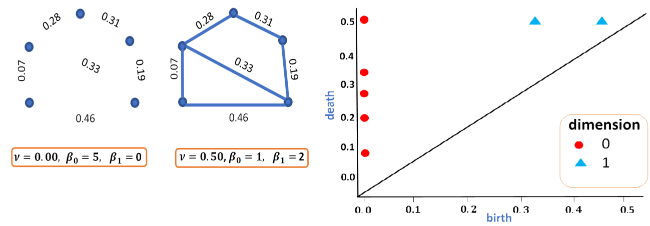

Figure 2: Illustration of Betti number (β_0, β_1) estimation and persistence diagram for VR-filtration.

Formally, the p-th Betti number β_p, $p \in Z^+$, of a simplicial complex (in our case the VR complex) is the rank of the associated p-th *homology* group. Informally, Betti numbers represent counts of connected components and p-dimensional holes. For instance, β_0 is the number of connected components; β_1 references the count of 1-dimensional holes, etc. Another visual tracker of the lifespan of topological structure is the persistence diagram (PD), and in this diagram each topological feature is denoted by a point (x, y). The x, y coordinates correspond to the birth and death times of the topological feature, respectively. Hence, features with a longer lifespan , i.e. a stronger persistence, are denoted by points that are far from the main diagonal. Figure 2 illustrates β_0 and β_1 estimation and the subsequent PD that is generated by the VR-filtration.

Numerically, two PDs D_a and D_b can be compared via a similarity measure known as the *bottleneck distance*, which is defined as: $b_i = \inf_\eta \sup_{x \in D_a} \|x - \eta(x)\|_\infty$. where η ranges over all bijections from $D_a \cup \Delta$ to $D_b \cup \Delta$, counting multiplicities, with $\Delta = \{(x, x) | x \in \mathbb{R}\}$ and $||z||_\infty = \max_r |z_r|$ (Wasserman 2018). Note that the proposed methodology is built with two kinds of bottleneck distances (i.e., b_0^i and b_1^i for scenario i).

Methodology

We introduce a new model **M**ulti-**C**hannel **D**eep **P**ortfolio **Net**works (MCDPNet). The MCDPNet has two main parts: *autoencoders* for reducing the total information to an information subset and this is applied to a large number of graphical, topological, and electrical features, i.e., $X = \{\text{graph} : r; \text{topology} : \beta_0, \beta_1, b_0, b_1, f; \text{electrical} : L_w, t\}$, and *convolutional neural network* (CNN) that (i) classifies *whether or not* any customer load is lost because of component outages and (ii) predicts *how much* customer load is lost because of those outages. Note that the dependent variable for the classification problem is I, and that for the regression (prediction) problem is U. Now, we will present the MCDPNet architecture in details and then elaborate on its implementation.

Autoencoder for feature learning With the graphical and topological features served as input, we use an autoencoder that has a fully-connected encoding layer with ReLU activation and a fully-connected decoder layer with linear activation. After auto-encoding the features in the original input, we consider the ℓ^2-norm difference between every feature and its auto-encoded version (i.e., $||X - \hat{X}||$, such that X represents the original features and $\hat{X}$ represents the auto-encoded version) and rank the features by this measure (*communal information*) from the lowest to the highest. The less ℓ^2-norm difference a feature has, the more important role it plays in the functionality of the power grid network. Since there is minimal to no benefit to include multiple features contributing the same information, we choose features based on ℓ^2-norm differences and fit them into CNN for the classification/regression tasks.

Convolutional neural network for optimal feature combination selection Through autoencoders, we obtain different feature combinations by using the ξ most *communal* features plus ζ-number of most *non-communal* features which we use to avoid adding unnecessary communal information, thus leading to ζ feature combinations. Let $X^\kappa = \{x_1^\kappa, x_2^\kappa, \ldots, x_{\xi+\kappa}^\kappa\} \in \mathbb{R}^{\xi+\kappa}$ denote the κ-th feature combination, where $\kappa \in \{1, 2, \ldots, \zeta\}$. That is, the set of feature combinations $\boldsymbol{X}$ is $\{X^1$: $\{\!\{\xi$ most *communal* features + the most *non-communal* feature$\}\!\}$, X^2: $\{\!\{\xi$ most *communal* features + 2 of most *non-communal* features$\}\!\}$, ..., X^ζ: $\{\!\{\xi$ most *communal* features + ζ-number of most *non-communal* features$\}\!\}\}$.

Since CNN models naturally accept image data, we construct image data before feeding the feature combination to the CNN model. First, we convert the κ-th feature combination into image data by calculating the normalized distance matrix such that the element in the matrix is the normalized difference between two features, i.e.,

$$\tilde{x}_{p,q}^\kappa = \frac{\exp\left(x_p^\kappa - x_q^\kappa\right)}{\sum_{q=1}^{\xi+\kappa} \exp\left(x_p^\kappa - x_q^\kappa\right)}$$

. Then, we obtain the corresponding feature distance matrix $\tilde{X}^\kappa \in \mathbb{R}^{(\xi+\kappa)\times(\xi+\kappa)}$, which is the input data that will

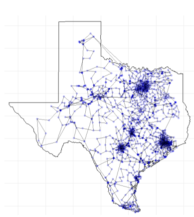

Figure 3: Map of the ACTIVSg2000 synthetic power system (Birchfield et al. 2017).

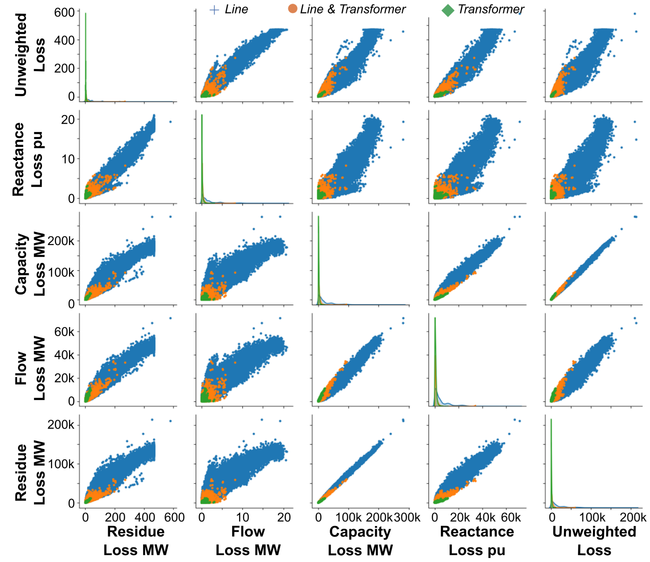

Figure 4: Scatterplot matrix of the relationship between the total weights of outaged components for the unweighted and four weighting schemes.

be given to the CNN model. After converting the ζ feature combinations $\boldsymbol{X}$ into ζ feature distance matrices $\tilde{\boldsymbol{X}}$, we fit these feature distance matrices into a CNN separately to estimate the impact of the contingencies on power systems and obtain the corresponding accuracy/ coefficient of determination (R^2). Objectively, the purpose of using CNN in the MCDPNet is (i) we do not need to model the underlying joint dependencies of the features via systematic conversion of non-image features into image forms and (ii) the input has two dimensions and CNN is faster at "deep learning" than the fully connected multilayer perceptron (MLP).

We combine the two steps above to learn the optimal feature combination to obtain the highest accuracy or the highest R^2. Equipped with the theory and structure of deep portfolio (Heaton, Polson, and Witte 2016), the framework of our proposed MCDPNet is able to select the final optimal feature combination based on preliminary results generated from multiple deep portfolios, and the deep portfolio theory provides a self-contained procedure for portfolio selection with deep neural networks. As our final step, we fit deep neural networks to the optimal feature combination (i.e., optimal feature distance matrix) as the input training data for classification and prediction tasks respectively.

Experiments

Dataset In this study we use the ACTIVSg2000 power-system model (Birchfield et al. 2017), shown in Figure 3, as a moderately complex, representative example of an electric power grid: it is a synthetic test system comprising 2000 buses, 544 generators, 1347 loads, 2345 lines, and 861 transformers. Static power-flow equations are solved using the the `PowerSimulations.jl` package of the Scalable Integrated Infrastructure Platform (SIIP) (Lara et al. 2018). This software, in turn, relies on the lossless DC load-

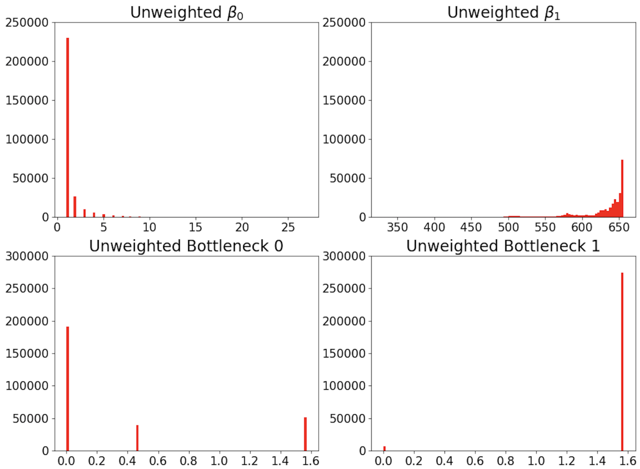

Figure 5: Histograms of Betti numbers and bottleneck distances in the contingency dataset.

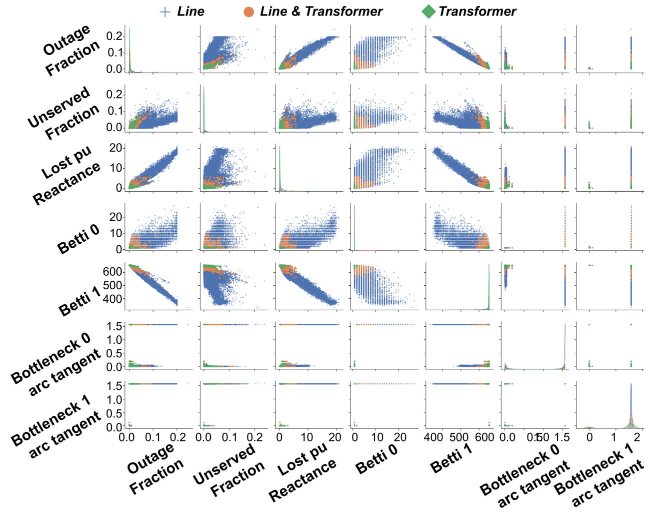

Figure 6: Scatterplot of the relationship between dependent variable (fraction of load not served, shown in the second row and second column) and independent variables in the contingency dataset under the electrical-distance weighting scheme.

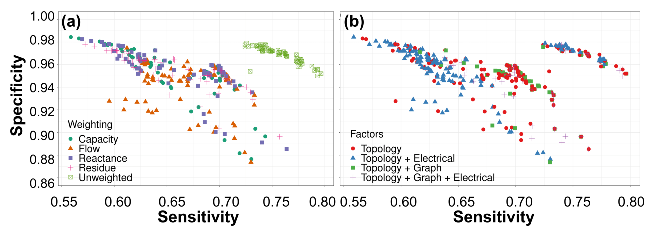

Figure 7: (a) Sensitivity and Specificity scores for predicting whether any customer load is lost due to a contingency, distinguished by the subset of independent used in the classification model; (b) Sensitivity and Specificity scores for predicting whether any customer load is lost due to a contingency, distinguished by how edges in the power-system graph are weighted.

flow algorithm implemented in the `DCPPowerModel` algorithm within the `PowerModels.jl` package provided by Los Alamos National Laboratory's Advanced Network Science Initiative (Coffrin et al. 2018). Graphical and topological measurements are extracted from both unweighted and weighted networks. For the weighted netwiorks, the following four separate edge-weighting schemes are considered: (1) electrical distance (Cotilla-Sanchez et al. 2013), represented by the per-unit reactance of the component, (2) capacity, expressed as the MW rating of the component, (3) base flow, equal to the MW flow in the component in the base case, and (4) residual capacity, the difference in MW between the capacity and base flow. Therefore for each of the four weighting schemes plus the unweighted network, we compute the total weight of the outaged components and each of the topological metrics. All metrics were computed using Python code and are published in the aforementioned case-study dataset (Bush 2020).

Baselines For the classification problem we consider 7 state-of-the-art baselines: classification trees (CART), deep neural networks (DeepNets), logistic regression, random forest, and support vector machine (a linear kernel [Linear SVM] and a radial-basis-function kernel [RDF SVM]), but for the regression problem we only consider DeepNets, ordinary least squares (OLS), and random forest. We reserve 30% of the dataset for testing and use sensitivity (true positives over total positives) and specificity (true negatives over total negatives) measures to score the classification problem and the fraction of variance explained to score the regression problem.

Parameter setting We implement MCDPNet on Amazon EC2 with 1 GPU (61 GB memory). With the autoencoding procedure for feature learning, the weight decay is set equal to 1×10^{-2} and the dimension of latent space is 5; we set the batch size to 256 (the mean training time per epoch is 1.27 seconds) and the autoencoder is trained for 500 epochs. To efficiently learn the features, we set the fixed number of the most communal features ξ to 5 and the number of least communal features ζ is selected from $\{1, 2, \ldots, 5\}$. With the CNN procedure of selecting optimal feature combinations, the optimal feature combination consists of 8 features and all convolution layers are followed by batch normalization, ReLU activation and Max-pooling. The dropout rate is 0.5, the weight decay is set to 0.01, and the number of hidden units is 16; we set the batch size to 32 (the mean training time per epoch is 17.6 seconds) and the CNN is trained for a 500 epochs.

Results Figure 4 shows the relationship between the unweighted and weighted schemes for the outaged component contingency simulations, and that they are highly correlated. The topological metrics computed are the Betti numbers (Betti-0 (β_0) and Betti-1 (β_1)) and the bottleneck distances, relative to the zero-contingency system. Figure 5 displays a histogram of the topological measurements: the Betti numbers tend to peak around the zero-contingency case (i.e., low

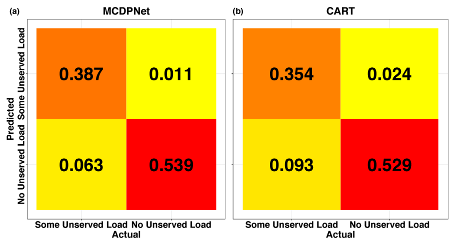

Figure 8: Confusion matrices of MCDPNet and CART (i.e., second-best baseline) for classification problem with unweighted edges and the full set of independent variables.

β_0 and high β_1), while the bottleneck distances are bi-modal, with the underlying persistence diagram being either very similar to or very different from the zero-contingency case. Since in many situations the bottleneck distance becomes extremely large (effectively infinite), we plot its arc-tangent value instead of the value itself, in order to better display the dynamic range of this variable.

Figure 6 shows scatterplots depicting the degree of correlation between the dependent variable and the independent variables (the fraction of components outaged, the total weight of outaged components, and the four topological measurements), under the electrical distance weighting scheme. Although the variables are clearly correlated, the several diagonal trends indicate the presence of non-trivial structure in the relationships. The relationships under the four other weighting schemes are quite similar to that for the electrical-distance weighting scheme presented in the same figure.

Figure 7(a) indicates that supplementing topological measurements with static electrical ones (i.e., the types of components outaged and their total edge weight) improves the predictions. Furthermore, Figure 7(b) demonstrates that using unweighted edges for the topological measurements significantly outperforms the other weighting schemes.

The confusion matrices in Figure 8 quantify how much the classifications tend to underestimate service disruptions more than they overestimate non-disruptions. Figure 8 shows that MCDPNet outperforms the second-best baseline (i.e., CART) with both its true positive rate (TPR) and false positive rate (FPR) by a large margin for the classification of unserved load.

Problem	Method	Score
Classification **(Accuracy)**	MCDPNet	**92.51 ± 0.50**
	CART	88.33 ± 0.11
	DeepNets	87.24 ± 0.47
	Linear SVM	87.66 ± 0.18
	Logistic regression	87.88 ± 0.09
	Random forest	88.23 ± 0.20
	RDF SVM	87.74 ± 0.21
Regression **(R^2)**	MCDPNet	**92.13**
	DeepNets	87.31
	OLS	70.61
	Random Forest	65.68

Table 1: Classification and regression scores (in percent) on the test set: accuracy (mean ± std) for classification and R^2 for regression.

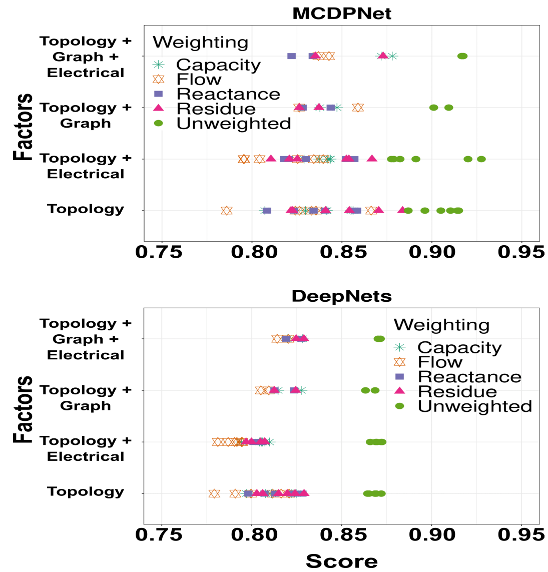

Figure 9: Score for amount of variance explained by two regression models (MCDPNet versus DeepNets) for different edge-weighting schemes and independent-variable subsets.

The regression results in Figure 9 indicate that the proposed MCDPNet outperforms all others. These results suggest that the unweighted-edge scheme is adequate (although weighting by electrical distance is better) and that supplementing the topological measurements with other independent variables (in this case, the radius of the contingency region and fraction of components outaged, which is a graph measurement) is sufficient.

Overall, we see that measurements of Betti numbers and bottleneck distances are modestly sufficient to predict whether load will not be served during a power-system contingency, and that not weighting edge works best for prediction. Under the MCDPNet, combining the topological measurements with basic graph-theoretic and static electrical ones supports coarse prediction of the quantity of load not served if edges are weighted by reactance (a measure of electrical distance), or they are left unweighted.

Lessons Learned on the Path to Deployment The case study demonstrated that basic graph connectivity information and topological measurements are sufficient to make coarse predictions of results for the focused, edge-based contingencies for a specific electrical network and that several ML techniques are adequate for this task. The addition of electrical information and the choice of deep neural networks improves predictability. Training the ML models, however, requires computing the impact of the contingency using power-flow simulation. Furthermore, the topological measurements and ML model training together typically take a similar amount of computation time as the power-flow simulation, although the ML model training times vary widely according to the specific algorithm.

The process situation of preparing the training dataset requires detailed electrical information and time-consuming simulations, performing the training requires moderately intense computation, but the prediction only uses minimal, non-electrical information and proceeds quickly. Such a system could be deployed in several ways:

- The holder of electrical data for a specific system could invest computational resources to train the ML contingency forecasting model that could later be used by themselves or a broader set of stakeholders to make rapid predictions of the impacts of hypothetical or impending contingencies. This *contingency screening* deployment would allow the stakeholders to focus computational resources such as detailed power-flow simulations on the most impactful contingencies identified by the coarse ML model. It also makes it possible to disseminate a contingency-prediction model to stakeholders without releasing information on the system's electrical properties—only the unweighted network graph and the ML model need be released. Hence, contingency screening can support rapid planning for natural disasters and human-caused events.
- If the methods of transfer learning could be successfully applied to create a contingency-screening model that is not tied to a specific system (see Cotilla-Sanchez et al. (2012) for comparison of graph characteristics of power-system networks in the U.S.), but instead applies to power systems in general, then such a model could be deployed to make coarse predictions of the impacts of contingencies in cases where the underlying electrical data for the system in question is not available to the user. One would just need the unweighted graph and that general-purpose model in order to estimate outages. Emergency-response and similar organizations could deploy such a model to make rapid assessments of potential or actual contingencies where they have limited information about or are unacquainted with the details of the power system. Some models could also be employed in the aforementioned contingency-screening role.
- In general, the performance requirement depends upon the particular user's computational budget and the value that they assign to cells in the confusion matrix (in the case of classification) or to mean-squared-error (in the case of regression). For a user interested in identifying contingencies involving loss of customer load, the technique must perform faster than the reciprocal of the sum of fraction of true and false positives for load loss, about $2.5\times$ for the case in the text; the prediction using our machine-learning technique is at least $1000\times$ faster than power flow-based contingency-analysis algorithms, so the model clearly performs adequately for this type of deployment. It would only be in situations where the user is extremely averse to false negatives that the model would be inadequate: here the figure of merit is the reciprocal of the false negative fraction, about $15\times$ in the case in the text, so a user would have to be 15 times more averse to a false negative relative to a true positive.

Conclusion and Future Work

The major findings of this analysis are that basic topological measurements such as Betti number and bottleneck distance are sufficient for coarsely predicting whether or not customer load is lost during a power system contingency, that a variety of deep learning and machining learning techniques are adequate for the task and that the graph's edges do not need to be weighted. Secondarily, using electrical properties in addition to the topological measurements allows coarse prediction of the quantity of customer load lost during the contingency, if the neural networks model is employed. These conclusions hold for contingencies generated by selecting a random outages uniformly within a fixed radius of a randomly chosen node within the test network. Our proposed MCDPNet model provides an optimal feature selection in a self-contained manner and can successfully be fine-tuned by incrementally enlarging the training dataset with newly features. The results show the highly competitive performance of our MCDPNet model in both classification and prediction tasks. Future work needs to focus on transfer learning and other aspects of generalizability, under a much wider (perhaps parameterized) characterization of contingency patterns (i.e., "attacks").

Furthermore, the current research direction could benefit from the systematic implementation of graph convolutional networks (GCN). As recently shown by (Chen, Gel, and Avrachenkov 2020), the Levy Flights Graph Convolutional Networks (LFGCN) deliver a highly competitive performance in node classification of power grid networks. In the future we plan to expand the results of the current project and explore utility of GCN coupled with TDA for power system learning under contingencies.

Acknowledgments

This work was authored in part by the National Renewable Energy Lab (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Dep. of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. Gel and Ofori-Boateng have been supported in part by NSF DMS 1736368. Chen has been supported by 2019 NSF Mathematical Sciences Graduate Internship (MSGI) and 2020 NSF INTERN under NSF ECCS 1824716.

References

Abedi, A.; Gaudard, L.; and Romerio, F. 2019. Review of major approaches to analyze vulnerability in power system. *Reliability Engineering & System Safety* 183: 153–172.

Asaad, A.; and Jassim, S. 2017. Topological data analysis for image tampering detection. In *Int. Workshop on Digital Watermarking*, 136–146. Springer.

Birchfield, A. B.; Xu, T.; Gegner, K. M.; Shetye, K. S.; and Overbye, T. J. 2017. Grid Structural Characteristics as Validation Criteria for Synthetic Networks. *IEEE Transactions on Power Systems* 32(4): 3258–3265.

Bose, D.; Chanda, C. K.; and Chakrabarti, A. 2020. Vulnerability assessment of a power transmission network employing complex network theory in a resilience framework. *Microsystem Technologies* 1–9.

Bush, B. W. 2020. Topology-Based Machine-Learning for Modeling Power-System Responses to Contingencies | NREL Data Catalog. URL https://data.nrel.gov/submissions/146.

Carlsson, G. 2009. Topology and data. *BAMS* 46(2).

Carlsson, G. 2019. Persistent homology and applied homotopy theory. *Handbook of Homotopy Theory* .

Chen, Y.; Gel, Y. R.; and Avrachenkov, K. 2020. LFGCN: Levitating over Graphs with Levy Flights. *2019 IEEE International Conference on Data Mining (ICDM)* .

Coffrin, C.; Bent, R.; Sundar, K.; Ng, Y.; and Lubin, M. 2018. Powermodels. jl: An open-source framework for exploring power flow formulations. In *2018 Power Systems Computation Conference (PSCC)*, 1–8. IEEE.

Cotilla-Sanchez, E.; Hines, P. D.; Barrows, C.; and Blumsack, S. 2012. Comparing the topological and electrical structure of the North American electric power infrastructure. *IEEE Systems Journal* 6(4): 616–626.

Cotilla-Sanchez, E.; Hines, P. D.; Barrows, C.; Blumsack, S.; and Patel, M. 2013. Multi-attribute partitioning of power networks based on electrical distance. *IEEE Trans. Power Syst.* 28(4): 4979–4987.

Cuadra, L.; Salcedo-Sanz, S.; Del Ser, J.; Jiménez-Fernández, S.; and Geem, Z. W. 2015. A critical review of robustness in power grids using complex networks concepts. *Energies* 8(9): 9211–9265. doi:10.3390/en8099211.

Dörfler, F. Bullo, B. 2013. Kron reduction of graphs with applications to electrical networks. *IEEE Transactions on Circuits and Systems* 60(1): 150–163.

Eskandarpour, R.; Khodaei, A.; and Arab, A. 2017. Improving power grid resilience through predictive outage estimation. In *IEEE NAPS*, 1–5.

Feng, M.; and Porter, M. A. 2020. Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence. *arXiv preprint arXiv:2001.01872* .

Gillis, J. M.; and Morsi, W. G. 2016. Non-intrusive load monitoring using semi-supervised machine learning and wavelet design. *IEEE Transactions on Smart Grid* 8(6): 2648–2655.

Heaton, J.; Polson, N. G.; and Witte, J. 2016. Deep portfolio theory. *arXiv preprint arXiv:1605.07230* .

Hines, P.; Blumsack, S.; Sanchez, E. C.; and Barrows, C. 2010. The Topological and Electrical Structure of Power Grids. In *43rd Hawaii International Conference on System Sciences*, 1–10.

Huang, G.; Chen, B.; Xiao, L.; Ran, Y.; and Zhang, G. 2019. Cascading Fault Analysis and Control Strategy for Computer Numerical Control Machine Tools Based on Meta Action. *IEEE Access* 7: 91202–91215.

Ibrahim, M. S.; Dong, W.; and Yang, Q. 2020. Machine learning driven smart electric power systems: Current trends and new perspectives. *Applied Energy* 272: 115237.

Islambekov, U.; Dey, A. K.; Gel, Y. R.; and Poor, H. V. 2018. Role of local geometry in robustness of power grid networks. In *IEEE GlobalSIP*, 885–889.

Jenssen, R.; Roverso, D.; et al. 2019. Intelligent monitoring and inspection of power line components powered by UAVs and deep learning. *IEEE Power and Energy Technology Systems Journal* 6(1): 11–21.

Jufri, F. H.; Widiputra, V.; and Jung, J. 2019. State-of-the-art review on power grid resilience to extreme weather events: Definitions, frameworks, quantitative assessment methodologies, and enhancement strategies. *Applied Energy* 239: 1049–1065.

Lara, J. D.; Krishnamurthy, D.; Barrows, C.; Research, L. D.; and Development. 2018. PowerSystems.jl and PowerSimulations.jl. doi:10.11578/dc.20181129.13.

Li, B.; Ofori-Boateng, D.; Gel, Y. R.; and Zhang, J. 2020. A hybrid approach for transmission grid resilience assessment using reliability metrics and power system local network topology. *Sustainable and Resilient Infrastructure* 1–16.

Li, M. Z.; Ryerson, M. S.; and Balakrishnan, H. 2019. Topological data analysis for aviation applications. *Transportation Research Part E: Logistics and Transportation Review* 128: 149–174.

Mohamed, M. A.; Chen, T.; Su, W.; and Jin, T. 2019. Proactive Resilience of Power Systems Against Natural Disasters: A Literature Review. *IEEE Access* 7: 163778–163795.

Nakarmi, U.; Naeini, M. R.; Hossain, M. J.; and Hasnat, M. A. 2020. Interaction Graphs for Cascading Failure Analysis in Power Grids: A Survey. *Energies* 13(9): 2219.

Oh, H.; Jung, J. H.; Jeon, B. C.; and Youn, B. D. 2017. Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis. *IEEE Transactions on Industrial Electronics* 65(4): 3539–3549.

Pi, R.; Cai, Y.; Li, Y.; and Cao, Y. 2018. Machine learning based on bayes networks to predict the cascading failure propagation. *IEEE Access* 6: 44815–44823.

Sánchez-García, R. J.; Fennelly, M.; Norris, S.; Wright, N.; Niblo, G.; Brodzki, J.; and Bialek, J. W. 2014. Hierarchical spectral clustering of power grids. *IEEE Trans. Power Syst.* 29(5): 2229–2237.

Shuvro, R. A.; Das, P.; Hayat, M. M.; and Talukder, M. 2019. Predicting Cascading Failures in Power Grids using Machine Learning Algorithms. In *IEEE NAPS*, 1–6.

Vugrin, E.; Castillo, A.; and Silva-Monroy, C. 2017. Resilience metrics for the electric power system: A performance-based approach. *Report: SAND2017-1493* .

Wasserman, L. 2018. Topological Data Analysis. *Annual Rev. of Stat. and Its Apps.* 5: 501–532.

Xu, Y.; Gurfinkel, A.; and Rikvold, P. 2014. Architecture of the Florida power grid as a complex network. *Physica A: Statistical Mechanics and its Applications* 401: 130–140.

Zhao, Y.; Li, T.; Zhang, X.; and Zhang, C. 2019. Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future. *Renewable and Sustainable Energy Reviews* 109: 85–101.

Zomorodian, A. 2010. Fast construction of the Vietoris-Rips complex. *Computers and Graphics* 34(3): 263–271.