2022 5th International Conference on Artificial Intelligence for Industries (AI4I) | 978-1-6654-5961-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/AI4I54798.2022.00016

Post-Fault Power Grid Voltage Prediction via 1D-CNN with Spatial Coupling

1st Carson Hu
Department of Mathematics
University of California
Irvine, CA 92697, USA.
hucf@uci.edu

2nd Guang Lin
Department of Mathematics
School of Mechanical Engineering
Purdue University
West Lafayette, IN 47907, USA
guanglin@purdue.edu

3rd Bao Wang
Department of Mathematics
Scientific Computing and Imaging Institute
University of Utah
Salt Lake City, UT 84112, USA
wangbaonj@gmail.com

4th Meng Yue
Sustainable Energy Technologies Department
Brookhaven National Laboratory
Upton, NY 11973, USA
yuemeng@bnl.gov

5th Jack Xin
Department of Mathematics
University of California
Irvine, CA 92697, USA.
jack.xin@uci.edu

***Abstract*—We propose a one-dimensional convolutional neural network (1D-CNN) with spatial coupling for post-fault power grid voltage prediction. Our proposed deep learning framework was inspired by the celebrated Prony's method in classical signal processing. Our spatio-temporal model significantly outperforms existing benchmarks, including long short-term memory model, and is applicable to other strong transients in power industries.**

***Index Terms*—Power grid, post-fault prediction, deep learning**

I. Introduction

Accurate power grid prediction and fault detection is crucial for decision making after anomalies (e.g. short circuits, weather related damages) occur. Computationally, it is challenging if a large physical system of circuit equations [6] has to be solved rapidly to produce accurate predictions. An alternative approach is to learn the system's response from short time observations and generate subsequent system responses. Classical signal processing uses linear time invariant filters, such as Padé and Prony's methods [3]. However, these methods are designed to fit responses with a rational system function. Besides requiring the user to select the degrees of numerator and denominator polynomials from observed data, the prediction is only on a scalar time series and can be sensitive to the choice of the degrees [5].

Prony's method is well-known for modeling impulse responses resembling post-fault responses of a power system. In this paper, we generalize Prony's method to 1D-CNN [1] with additional spatial information (e.g. currents in power lines near a node/bus). The presence of temporal (1D) convolution is a common ground of both methods, though 1D-CNN also contains nonlinear operations such as activation and max-pooling as well as dense layers. Related work [4] applied long short-term memory (LSTM) neural network to model transient dynamics of individual power generator and observed much faster inference than computational physical simulations [6]. However, we have not found LSTM to be effective in modeling post-fault responses. When varying the lookback of the LSTM model between 1 and 20, we found that the errors compound far too quickly in prediction, making the model fairly useless. Compared to the LSTM, a 1D-CNN is much better at extracting non-local dynamic features for prediction. Our main contributions are:

1) We trained 1D-CNN to successfully generalize the Prony's method to deep learning in the context of power-grid data, free from choosing integer hyper-parameters location by location.

2) The input to the 1D-CNN can be a concatenated vector consisting of node voltage and nearby line currents, consistent with measurements commonly used in power energy industry.

3) When applied to the prediction of voltage response shortly after a fault occurs on a realistic power system, 1D-CNN with neighboring current as joint input outperforms 1D-CNN with voltage input alone, while the latter improves on Prony's method.

The code can be found on github.

II. Dataset and Our Goal

Time-domain simulations using a software tool EPTOOL [2] generate a library of system dynamic responses to a variety of disturbances. The simulations are performed on the New York/New England 16-generator 68-bus power system. The data consists of 3638 different events (faults), 70 percent of which are selected as training data, 30 percent as test data. Each event contains data describing 1000 points of time (1000 ms) for each of the 68 buses. The first 200 points after the spike (estimated to be around the 105-th point) are used as input, the last 600 points are used as output. The voltage is recorded for each node, and the current is recorded for each line connecting the nodes to one another. Our goal is to predict

978-1-6654-5961-7/22/$31.00 ©2022 IEEE
DOI 10.1109/AI4I54798.2022.00016

the power grid's response immediately after the fault occurs for unseen scenarios, which assists real-time decision-making.

III. Model and Training

Our 1D-CNN model consists of a 1D-convolutional layer, a max-pooling layer, a 1D flattening layer, and two dense layers, with 392,466 parameters, of which 387,842 are trainable. For voltage data, the input is a vector $x \in \mathbb{R}^{200}$, representing the magnitude of the voltage for the first 200 time steps (200 ms); the fault occurs approximately 100 ms into the input, so this reflects the behavior of a node in the power grid immediately before and after a fault. The 1D-CNN outputs a prediction for the voltage of the next 600 time steps $y \in \mathbb{R}^{600}$, allowing us to predict voltage response on the power grid after the fault occurs. We found that incorporating neighboring current information can improve the predictions of the network. To accomplish this, we train a separate 1D-CNN for each node in the power-grid. The input is a multidimensional vector (a matrix) $x \in \mathbb{R}^{200\times(N+1)}$, where N is the number of neighbors. The first column represents the voltage at the node, whereas the subsequent columns represent the current from the neighbors. We train the nodes separately since the voltage behavior per node is highly independent. To train the model, we minimize MSE (mean-squares-error) between predicted voltage and actual voltage for the next 600 points by a stochastic gradient descent solver in Tensorflow.

IV. Experiments

On its own, the 1D-CNN generically outperforms Prony's method [3] as seen in comparisons of Fig. 1 on 4 buses. Prony's method requires hand-selecting two integer parameters to fit the observation at each node of the power network. This is a labor-intensive task that still results in an inconsistent method; even after optimal parameters are set, the method still fails to capture the overall voltage trend in many cases. *An advantage of our 1D-CNN approach is that it can model all nodes without tuning integer hyper-parameters.* The 1D-CNN is further improved by coupling neighboring currents (Table 1 and Fig. 2). Our model inference is fast, e.g. with an average prediction time of 0.046 seconds on a Tesla T4 machine.

In Table 1, we compare performance of the various models on voltage prediction. MSE results are grouped based on the number of neighbors for each node: the maximum number of neighbors in this graph is 5. The first column is the MSE when training a separate 1D-CNN for each node utilizing neighboring current information. The second column represents the results of the 1D-CNN with only voltage as input, no additional current information from the neighboring lines.

V. Conclusion

Motivated by the classical Prony's method, we adopted the 1D-CNN with spatial coupling for real-time long-term power grid prediction after the fault of the system occurs. Our proposed algorithm remarkably outperforms existing benchmark method, indicating the potential of 1D-CNN with spatial coupling for power industrial applications.

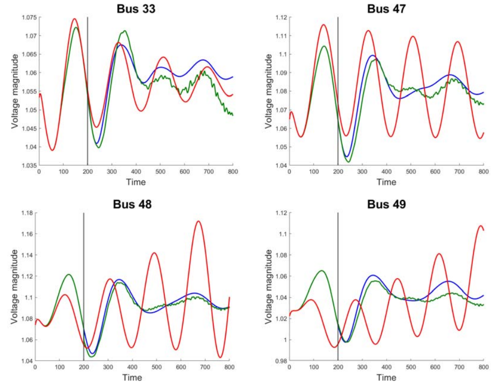

Fig. 1: Comparing 1D-CNN output (green), Prony's output [3] (red), and predictions vs. data (blue). Training/test timestamps: 0-200/201-800; separated by the vertical line at 200 ms.

TABLE I: MSE Comparison between different CNN models.

Neighbors	MSE (1D-CNN with neighbors)	MSE (1D-CNN)
1	1.6e-4	1.6e-4
2	1.7e-4	2.9e-4
3	1.5e-4	2.3e-4
4	1.9e-4	3.0e-4
5	1.1e-4	2.1e-4
Total	1.6e-4	2.3e-4

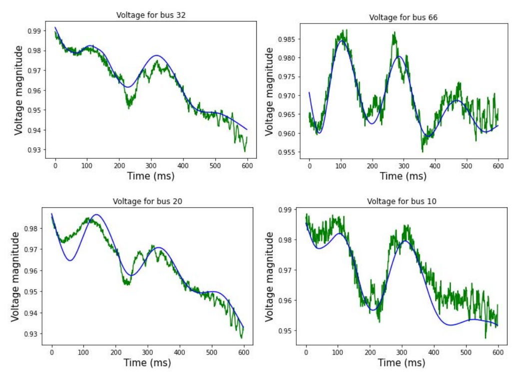

Fig. 2: Contrasting the predictions of 1D-CNN coupling 4 neighboring line currents (green) with ground truth data (blue).

Acknowledgments

The work was partially supported by NSF grants DMS-1924935, DMS-1952644, DMS-1924548, and DMS-1952339. And the work was also partially supported by the department of energy under the contract number DE-SC0021142.

References

[1] Z. Cui, W. Chen, Y. Chen, *Multi-Scale Convolutional Neural Networks for Time Series Classification*, https://arxiv.org/abs/1603.06995.

[2] "Power system toolbox", available at https://www.ecse.rpi.edu/chowj

[3] M. Hayes, "Statistical Digital Processing and Modeling", chapter 4, John Wiley & Sons, 1996.

[4] J. Li, M, Yue, Y. Zhao, G. Lin, *Machine-learning-based online transient analysis via iterative computation of generator dynamics*, in Proc. IEEE SmartGridComm, 2020.

[5] B. A. Sarray, S. Chrétien, P. Clarkson, G. Cottez, *Enhancing Prony's method by nuclear norm penalization and extension to missing data*, SIViP 11, 1089–1096 (2017).
[6] B. Stott, "Power system dynamic response calculations", Proceedings of the IEEE, vol. 67, no. 2, pp. 219–241, 1979.