Spatial Temporal Graph Neural Networks and Benchmarks for Node-level Energy Load Forecasting

Austin Bell
alb2307
Columbia University
austin.bell@columbia.edu

Ziyin Wang
zw2605
Columbia University
zw2605@columbia.edu

Malik Alix Drabla
mad2275
Columbia University
mad2275@columbia.edu

Abstract

Our project looks to leverage the state of the art in spatial-temporal graph neural networks for the purpose of node level energy load forecasting across Europe. Additionally, we acknowledge that there has been very little research on node level forecasting, therefore, we also develop a series of statistical, machine learning, and deep learning benchmarking algorithms for a more effective comparison. We utilize the RE-Europe dataset which consists of 1,494 nodes and their transmission connections plus hourly demand and solar / wind values and solar / wind forecasts every 12 hours.

1 Introduction

Since the industrial revolution, our earth has been warming due to the release greenhouse gases from burning 'dirty' fuel sources. A recent surge in clean energy investment has led to more widespread adoption of renewable, clean energy, however we note that there still remains many technical barriers to its adoption[1]. In particular, we highlight two technical barriers [1], [2], [3]:

- Geographic siting of plants: Fossil fuel plants are built remotely and energy production is centralized, whereas, renewable energy plants are decentralized and sometimes built in residential areas
- Storage or excess energy: the production of renewable energy is unpredictable leading to less efficient allocation of energy supply

While a wide net of solutions have been proposed and are currently being pursued, our project focuses on the use of AI to improve efficiency in the allocation of energy. We propose using spatial-temporal graph neural networks to accurately forecast load for each decentralized node. Through better predictions of energy demand, we can partially address each of these technical barriers.

- Spatial analysis enables us to better capture any geographic dependent features related demand (e.g., nodes providing electricity to cities are likely to be characterized by more extremes)

[1]We acknowledge that there are also significant economic, political, and social barriers to entry, but maintain that these are out of scope for our research.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

- Proximity of nodes will likely affect demand as one node reaches capacity they may divert the provision of energy to the nearest node that is below capacity

Overall, by better understanding our exact energy needs, we will be able to better optimize power flow across decentralized plants, ensure supply more closely meets demand, and thereby, lessening our reliance on advanced storage solutions.

1.1 Literature Review

Our project is novel in that it is the first to combine two ongoing lines of research: 1) deep learning for load forecasting and 2) spatial-temporal graph neural networks.

1.1.1 Load Forecasting

Prior to the resurgence of deep learning, most load prediction models utilized more traditional statistical algorithms. Within this group of statistical algorithms, we include linear regression[4], multiple regression models[5], [6], Autoregressive moving average (ARMA) models[7], and autoregressive integrated moving average (ARIMA) models[8], [9]. However, our focus will be on more black-box algorithms.

There are a variety of neural network approaches that look to capture the temporal features of load forecasting. The most popular approach is using RNN's with add-ons or adjustments to model day ahead or week ahead demand. Polson and Sokolov (2018)[10] create a new loss function leveraging extreme value theory to better model extreme peaks and troughs of energy demand . Bouktif et al (2018)[11] leverages a genetic algorithm to identify the optimal LSTM and feature configuration for load forecasting . Additional networks have been implemented including deep belief networks (Ouyang et al 2017 [12]) and more complex architectures combining various convolution and RNN layers (He 2017)[13].

In each of these, we note that similar features are used as inputs to the model including weather features (forecasted temperature, wind, sun), current and previous load time series, and temporal identifiers (holiday, day of week, season, etc.).

1.1.2 Spatial Temporal Graph Neural Networks

On the other side significant research is being conducted on spatial temporal graph neural networks ("STGNNs"), but we focus on advances within traffic forecasting. Originally developed by Bing Yu and his team in 2018[14], multiple improvements have occurred over the past two years. Their underlying model consists of multiple spatial-temporal blocks, which are each made up of three convolution layers (temporal + spatial + temporal).

Various improvements have since expanded on the original STGNNs. Guo et al (2019)[15] added an attention mechanism to better attend to road networks with more influence. Bai et al (2019)[16] combines the spatial temporal graph with a seq2seq architecture with attention for their predictions. Wu et al (2019)[17] utilizes both a graph convolution and a causal convolution. Finally, Song et al (2020)[18] first develops a localized spatial temporal graph by connecting each node with itself in the adjacent time steps. These localized graphs along with learnable temporal and spatial embeddings serve as input into the STGNN.

2 Methodology, Goal, and Criteria

The goal of our project is to improve forecast prediction of node demand through leveraging STGNNs. Our energy nodes will represent our graph nodes and their transmission connections will represent the edges. Each node includes a variety of data at each timestamp such as hourly demand, wind / solar forecasts, and wind / solar realizations. We will use spatial and temporal convolutions at each time step (i.e., each hour) to effectively forecast our time series.

The first step of our model will be to implement the base spatial temporal graph convolutional neural network proposed by Yu et al (2018). Once our baseline is set, we will seek to implement the state of the art model currently provided by Song et al. Where we connect nodes with themselves in the adjacent timesteps and incorporate spatial and temporal embeddings prior to running our graph network. Additionally, we will incorporate new variables such as various temporal markers (e.g., holidays, season, day of week) and other indicators of energy use (e.g., GDP and population).

Experiments will be run using both models to forecast day ahead and 7-day ahead energy demand. For each experiment, we will evaluate the time series output using mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE).

We will find it difficult to compare our results because this is the only dataset available (to our knowledge) that provides node-level load estimates due to confidentiality issues. Furthermore, we are unable to find any load forecasting results utilizing this data. To ensure that we can most effectively contribute to ongoing research in the field, our team will implement a series of bench marking statistical (e.g., ARIMA), machine learning (e.g., Random Forests and XGB), and deep learning algorithms (e.g., LSTM and CNN-LSTM) for fair comparison.

Therefore, we believe that our research will have two primary contributions to the field of load forecasting. First and foremost, to our knowledge, we are the first to explore the use of graph neural networks at the nodal level for load forecasting. Secondly, we will be the first to benchmark a series of algorithms leveraging this data set for the benefit of ongoing research.

3 Dataset

The dataset that we will be using is the RE-Europe dataset[19], which combines a transmission network model, plus data regarding electric generation and demand from 2012-2014. This transmission model identifies relevant nodes and their connections plus transmission capacity. For each hour of the day, demand values are provided for each node. Additionally, solar / wind forecasts are provided every 12 hours and their realized values are provided at the hourly level. All temporal data is provided over a period of three years.

The data covers the entirety of mainland Europe and includes 1,494 unique nodes with corresponding demand. These nodes make up 2,156 connections. While, node distance is not included, each node has a latitude and longitude, which can be used to identify the corresponding distance to other nodes. Additionally, the installed renewable capacity for each node is included.

References

[1] Seetharaman, Krishna Moorthy, Nitin Patwa, Saravanan, and Yash Gupta, Breaking barriers in deployment of renewable energy, Heliyon, 2019.

[2] Izadbakhsh M., Gandomkar M., Rezvani A., Ahmadi A. Short-term resource scheduling of a renewable energy based micro grid. Renew. Energy. 2015;75(March):598–606.

[3] Weitemeyer S., Kleinhans D., Vogt T., Agert C. Integration of renewable energy sources in future power systems: The role of storage. Renew. Energy. 2014;75:14–20

[4] A. Goia, C. May, G. Fusai, Functional clustering and linear regression for peak load forecasting, Int J Forecast, 26 (4) (2010), pp. 700-711

[5] R. Ramanathan, R. Engle, C.W. Granger, F. Vahid-Araghi, C. Brace, “Short-run forecasts of electricity loads and peaks”, Int J Forecast, 13 (2) (1997), pp. 161-174

[6] Close Amral N, Ozveren C, King D. Short term load forecasting using multiple linear regression. In: UPEC 2007. 42nd international universities power engineering conference, 2007, 2007. p. 1192–8

[7] S. Pappas, L. Ekonomou, D. Karamousantas, G. Chatzarakis, S. Katsikas, P. Liatsis, “Electricity demand loads modeling using autoregressive moving average (ARMA) models”, Energy, 33 (9) (2008), pp. 1353-1360

[8] C.-M. Lee, C.-N. Ko, “Short-term load forecasting using lifting scheme and ARIMA models”, Expert Syst Appl, 38 (5) (2011), pp. 5902-5911

[9] Cho M, Hwang J, Chen C-S. “Customer short term load forecasting by using ARIMA transfer function model.” In: Proceedings of EMPD ’95., 1995 international conference on energy management and power delivery, 1995. vol. 1; 1995. p. 317–22

[10] Michael Polson and Vadim Sokolov, Deep Learning for Energy Markets, Arxiv, 2018

[11] Salah Bouktif, Ali Fiaz, Ali Ouni, and Mohamed Adel Serhani, Optimal Deep Learning LSTM Model for Electric Load Forecasting using Feature Selection and Genetic Algorithm: Comparison with Machine Learning Approaches, Energies 2018, 11(7), 1636

[12] Tinghui Ouyang, Yusen He, Huajin Li, Zhiyu Sun, and Stephen Baek, A Deep Learning Framework for Short-term Power Load Forecasting, Arxiv, 2017

[13] Wan He, Load Forecasting via Deep Neural Networks, Procedia Computer Science, Volume 122, 2017, Pages 308-314

[14] Yu, B.; Yin, H.; and Zhu, Z. 2018. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In IJCAI.

[15] Guo, S.; Lin, Y.; Feng, N.; Song, C.; and Wan, H. 2019a. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, 922–929.

[16] Bai, L.; Yao, L.; Kanhere, S.; Wang, X.; and Sheng, Q. 2019. Stg2seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting. In IJCAI

[17] Wu, Z.; Pan, S.; Long, G.; Jiang, J.; and Zhang, C. 2019. Graph wavenet for deep spatial-temporal graph modeling. In IJCAI

[18] Chao Song, Youfang Lin, Shengnan Guo, and Huaiyu Wan, 2020, Spatial-Temporal Synchronous Graph Convolutional Networks: A New Framework for Spatial-Temporal Network Data Forecasting, AAAI.

[19] Jensen, T., Pinson, P. RE-Europe, a large-scale dataset for modeling a highly renewable European electricity system. Sci Data 4, 170175 (2017).