Electric Load Forecasting with Boosting based Sample Transfer

Di Wu [*1] **Can Cui** [*2] **Benoit Boulet** [1]

Abstract

With the increasing adoption of renewable energy generation and different types of electronic devices, electric load forecasting, especially short-term load forecasting, has attracted more and more attention. Accurate short-term load forecasting is of significant importance for the safety and efficiency of power grids. Deep learning-based models have shown impressive success on several applications, including short-term load forecasting. However, for several real-world scenarios, it may be very difficult or even impossible to collect enough training data to learn a reliable machine learning model. Specifically, in this paper, we propose a sample-based transfer learning algorithm to assist the learning performance of short-term load forecasting for houses with limited training data, e.g., newly built houses. The proposed method is evaluated on several real-world data sets and shows significant improvement over the baselines.

1. Introduction

Electric load forecasting is of significant importance for the secure and economic operation of the power grids (Wu, 2018; Wu et al., 2019). Depending on the forecasting horizons, electric load forecasting ranges from short-term (hours or minutes ahead) load forecasting to long-term (years ahead) load forecasting. *Short-term load forecasting (STLF)* is mainly used to assist real-time energy dispatching while long-term load forecasting is mainly applied for power grid infrastructure planning. Accurate short-term electric load forecasting can facilitate efficient residential energy management and power grid operation (Wu et al., 2017b). It is hard to store the electricity in a large quantity, thus is is important to keep the power generation close to the actual power demand. As reported in (Bunn and Farmer, 1985), even a 1% forecasting error increase can lead to more than £10 million increase on the operation cost of UK power grid.

The modern power grid is now facing fundamental changes from both the power supply side and the power demand side. The penetration of renewable energy generation is increasing very fast in recent years. The renewable energy generation (including wind and solar power generation) has been increasing exponentially in the last ten years (REN21, 2019),. The generation of renewable energy is highly influenced by weather conditions. Meanwhile, different types of electric appliances have been deployed in the power grids. The adoption of EVs has been growing very fast in the last few years. The annual EV sale increase in 2018 is 79% in Canada and 81% in US (InsideEVs, 2019). The EV charging demand could be directly affected by people's activities. Therefore, the consumption uncertainties will keep increasing very fast because of electric vehicles (EVs) and other electric appliances. Due to these reasons, electric load forecasting, especially short-term residential electric load forecasting, is very challenging. In this paper, we focus on tackling short-term load forecasting for single home houses.

Electric load forecasting has been an important research topic for a few years. Many approaches have been proposed for the forecasting of electric load (). The approaches can be generally categorized into two approaches: statistical methods and machine learning methods. The most frequently used classical statistical method is the autoregressive integrated moving average (ARIMA) model (He et al., 2012; Matsila and Bokoro, 2018). ARIMA-based models have been developed for load forecasting for distribution power grids (He et al., 2012) and hospital buildings (Matsila and Bokoro, 2018).

More recently, machine learning-based methods have appeared and due to their good performance, these approaches have drawn considerable attentions. There are a number of different kinds of machine learning-based approaches including support vector regression (SVR) (Ceperic et al., 2013; Ye and Yang, 2018; Chen et al., 2017), general additive models (Wu et al., 2017a), regression trees (Wu et al., 2016), neural networks (NNs) (Hippert et al., 2001; Kong et al., 2017; Kim et al., 2018), and Gaussian process regression (Di et al., 2018). SVR-based STLF methods are described in (Ye and Yang, 2018). In (Chen et al., 2017), a

[*]Equal contribution [1]Department of Electrical and Computer Engineering, McGill University, Montreal, Quebec, Canada [2]Google, New York City, New York, USA. Correspondence to: Di Wu <di.wu5@mail.mcgill.ca>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

SVR-based forecasting model is used to predict the hourly load of office buildings. The performance of the general additive model for residential STLF is studied in (Wu et al., 2017a).

However, most of the previous work on electric load forecasting assumes a large amount of training data for the house we are interested in, referred to as the target house. This assumption may not be satisfied for several real-world scenarios, e.g., newly built houses and buildings. Without enough training data, it will be very difficult to learn a reliable forecasting model. However, it is possible that we can have access to houses in the same area or with similar properties (e.g., building size), referred to as the source houses. Thus, we can apply transfer learning to use the information from the source houses to improve the forecasting performance of the target domain, which is the focus of this work. Specifically, in this paper, we propose a sample-based forecasting method that can help improve the forecasting performance of the target house by transferring samples from source houses.

The main contributions of this paper are summarized as follows: 1) a multi-source sample based boosting transfer method is proposed which can transfer the instances from source domain to improve the learning performance in the target domain; 2) The proposed method is investigated against other baselines on real-world data sets and has shown superior forecasting performances over other baselines. The rest of this paper is organized as follows. The technical background is presented in Section 2 and the proposed two forecasting algorithms are presented in Section 3. Section 4 presents the experimental results. Finally, the conclusion and future work for this paper are presented in Section 5.

2. Technical Background

2.1. Boosting

Ensemble is a type machine learning technique that can improve the performance of base machine learning learner by formulating an final model with several base learners. Boosting is an ensemble algorithm primarily focus on reducing the bias in a sequential manner. The boosting framework can be naturally combined with transfer learning. TrAdaboost (Dai et al., 2007) is designed to improve the classification accuracy by transferring samples from the source domain to the target domain in a boosting framework. Latterly, TrAdaboost was further extended to transfer with multiple source domains in (Yao and Doretto, 2010). Besides the classification tasks, boosting based knowledge transfer has also been adopted for the regression problems. In (Wu et al., 2019), a gradient boosting based model transfer was proposed for regression tasks. Different from these previous works, we propose to transfer samples from multiple sources to improve the target domain forecasting performance in a boosting framework.

Algorithm 1 TrAdaM: Transfer Adaptive sample based regression from Multiple sources

Input: Data set for target domain $\mathcal{D}_T$ (size of n), S data sets $\mathcal{D}_{S1}(totalsizeofm), \ldots, \mathcal{D}_{SM}$, number of iteration T, number of steps K, let T be the combination of all the source data sets and target data set, iteration number T

```
Initialize the weight vector w^1 such that w_i^1 = 1/(n+m)
TraAdaM
for s = 1, ..., K do
    call the LSTM base learner
    calculate the adjusted error e_i^t
    Update the weight vector following TrAdaS (Pardoe and Stone, 2010)
end for
Return the regression model
```

Output: the final forecasting model $F(x)$

2.2. Transfer Learning

Without enough training data, it will be very challenging to learn a reliable machine learning model. Transfer learning (Pan and Yang, 2010) is aimed to tackle this challenge. Specifically, transfer learning aims to improve the learning in the domain we are interested in (referred to as the target domain) via reusing the knowledge learned from other correlated domains (referred to as source domains). Depending on the reused knowledge from source domains, there are mainly two types of algorithms: sample-based transfer learning methods and model-based transfer learning methods. The sample-based transfer learning algorithms aim to reuse the samples from source domains, while the model-based algorithms aim to reuse the models learned from source domains. Transfer learning has been applied in various applications, including image recognition, language processing, and robotics. However, transfer learning has not been well studied for electric load forecasting. In (Wu et al., 2019), the authors proposed a multiple kernel-based transfer learning model for residential load forecasting. This paper proposes a boosting-based sample-based transfer learning algorithm with a deep learning model as the base forecasting model.

3. Methods

In this section, we first present the based forecasting model, i.e., an LSTM based forecasting model and then we present the sample-based electric load forecasting framework.

3.1. LSTM based Load Forecasting

Different types of neural networks have been applied for short-term electric load forecasting including feedforward neural networks, recurrent neural networks, and convolu-

Table 1. MAPE (%) and MAE (%) for short-term load forecasting

Metric	MAPE					MAE				
Method	House 1	House 2	House 3	House 4	House 5	House 1	House 2	House 3	House 4	House 5
Linear	7.61	10.89	7.63	8.67	6.69	2.12	3.67	2.04	1.62	2.12
SVR	7.86	15.68	6.64	9.28	6.56	2.21	5.36	1.81	1.80	2.16
NN	7.17	10.9	7.63	8.83	9.16	1.98	3.67	2.04	1.63	3.03
LSTM	6.92	13.91	5.36	8.18	7.05	1.86	4.77	1.52	1.55	2.38
TrAdaS	4.16	4.36	4.48	6.50	4.94	1.12	1.27	1.27	1.19	1.76
TrAdaM	**4.06**	**4.31**	**3.97**	**6.35**	**4.64**	**1.11**	**1.26**	**1.14**	**1.15**	**1.63**

tional neural networks. Among all these types of neural networks, RNNs have shown to be more suitable for STLF due to its sequential nature (Kong et al., 2017). However, RNNs will suffer from gradient vanishing and gradient explosion issues when dealing with long sequences. Compared with RNNs, the LSTM model is better on capturing long-term dependencies with additional gate mechanisms. In this paper, an LSTM based forecasting model is used as the base forecasting model. Specifically, the base model is consisted of an input layer, a set of LSTM layers, a dense layer and a output (corresponding to the predicted one hour ahead electric load consumption).

3.2. Boosting based Sample Transfer

In this paper, we use the sample transfer to tackle the concern on the limited amount of target house data. Specifically, our work is inspired by the two-stage TrAdaBoost.R2 (Pardoe and Stone, 2010) which transfers samples from one source house. For the simplicity, in this work, the two-stage TrAdaBoost.R2 is referred to as TrAdaS. In TrAdaS, samples are transferred from the source domain to the target domain and the sample weights will be reweighted based on the error of all the source samples. The final output model will be a model learned on all the target domain samples as well as re-weighted source domain samples.

Furthermore, in the real-world, we may have more than one source domains. Thus, we propose a simple extension of TrAdaS, i.e., extending it to the multiple sources scenario. As shown in Algorithm 1, we assume that we can have abundant data from multiple sources. Instead of transferring samples from one source, here we can transfer samples from multiple sources. The final output model will be a model learned on all the target domain samples as well as re-weighted source domain samples from all the available source domains.

4. Experiments and Simulation Results

The residential electric load data sets from OpenEI ($OPENEI$, 2021) are used to evaluate the effectiveness of proposed LSTM based forecasting algorithm. The data set include one-year (2014) hourly load consumption data (8760 data points) for 72 residential houses in New York.

In this paper, three types of features are used as the inputs to forecast one hour ahead electric load consumption of a single household. Specifically, features for the lagged electric load (electricity consumed in the last four hours), lagged temperature information (temperature in the last four hours), and weekday/weekend information (1 for weekday and 0 for the weekend) are used for short-term electric load forecasting. All the features are normalized between zero and one with min-max normalization.

4.1. Baselines and Evaluation Metrics

In this paper, Mean average percentage error (MAPE) as shown in Eq. 1 and mean absolute error (MAE) as shown in Eq. 2 are used to evaluate the effectiveness of the proposed algorithms. As shown in these two equations, y_i' is the predicted load consumption and y_i is the real value for load consumption at $i-th$ time slot. For benchmark results, apart from the base LSTM model, three other frequently used forecasting methods are used as baseline models: Linear regression, SVR and feed-forward Neural Network. All baseline models are tuned with parameters with the best performance for the data set.

$$MAPE = \frac{\sum_{i=1}^{N} \frac{|y_i - y_i'|}{y_i}}{N} \tag{1}$$

$$MAE = \frac{\sum_{i=1}^{N} |y_i - y_i'|}{N} \tag{2}$$

We randomly picked five houses from the data set as the target houses for our experiments. We use the last two days data of November with features described above as target house training data and evaluate the models' performance on the whole month of December. For the transfer learning methods, October and November data of other houses are used as source domain data. For single source setting (TrAdaS), one source house is randomly picked, for multiple sources setting (TrAdaM), 9 source houses are randomly picked.

Table 2. MAPE (%) and MAE (%) for short-term load forecasting with **noisy** data

Metric	MAPE					MAE				
Method	House 1	House 2	House 3	House 4	House 5	House 1	House 2	House 3	House 4	House 5
Linear	8.61	12.26	7.46	8.15	7.03	2.52	4.11	1.98	1.58	2.26
SVR	8.39	18.06	7.46	10.00	7.55	2.36	6.15	1.20	1.92	2.39
NN	8.35	12.26	7.48	8.16	7.03	2.43	4.12	1.99	1.58	2.26
LSTM	7.96	16.40	6.88	9.56	8.36	2.20	5.36	1.90	1.81	2.61
TrAdaS	4.52	4.61	4.23	6.45	**5.23**	1.23	1.34	1.20	1.19	**1.86**
TrAdaM	**4.41**	**4.44**	**4.15**	**6.40**	5.44	**1.20**	**1.30**	**1.20**	**1.17**	1.99

4.2. Evaluations on five houses

We first evaluate the performance of the proposed methods on the aforementioned five randomly chosen target houses. For the sample-based transfer algorithms, we have one randomly picked house as source house for TrAdaS and nine randomly picked houses as source houses for TrAdaM.

Table 1 shows the average forecasting performance of different algorithms on the five houses. As shown in this table, we can see that the proposed TrAdaM can show consistent superior average performance over other baseline models. Specifically, TrAdaBoost with multiple sources (TrAdaM) has the best performance for all houses. Compared with the base LSTM model, TrAdaM has 38.56% improvement for MAPE and 39.24% improvement for MAE on average. These results show that with the knowledge transfer, the forecasting performance can be significantly improved.

4.3. Evaluations on five houses with noisy data

To further showcase the robustness of the proposed method (TrAdaM), we compare its performance with other baselines on noisy data. Specifically, we add Gaussian noise of mean 0 and standard deviation 0.5 to all data.

Table 2 shows the average forecasting performance of different algorithms on the five houses with noisy data. As shown in this table, we can see that the proposed method can also significantly outperform other baselines. TrAdaM still has the best performance for most houses. Specifically, on average, TrAdaM had 45% improvement over the base LSTM model for MAPE and 43.42% improvement for MAE. These results show that the proposed methods are robust. Also, we can see that TrAdaM can outperform TrAdaS on four houses out of five target houses for MAPE and on all the five target houses for MAE.

4.4. Evaluations with varying number of training samples

To further analyze the benefits of the proposed forecasting methods. We analyze the effectiveness of the proposed method with a varying number of training samples for the target house. Specifically, we analyzed the performance House 5 with different numbers of training samples.

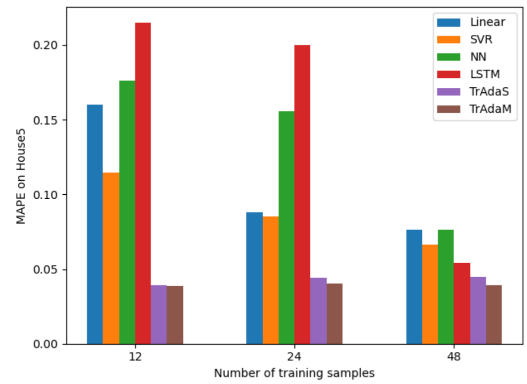

Figure 1. MAPE on House 5 with varying training samples

Fig. 1 shows the MAPE (note that the Y axis is not in percentage) on House 5 of all models with varying training samples. We can see that our proposed transfer learning method can achieve the best forecasting performance in all three cases. Also, it can get higher performance gain with fewer training samples in the target domain. This implies that knowledge transfer is more beneficial when little data is available in the target domain.

5. Conclusion and Future Work

With the fast development of different types of electric appliance, the electric consumption is increasing fast and its pattern is becoming more complex. Meanwhile, the penetration of renewable energy generation is also increasing very fast. Short-term electric load consumption is of crucial importance for the safe and secure operation of modern power systems. However, most of the current work on load forecasting assumes that there is a large amount of training data available which could be very challenging in the real world. In this work, we proposed an instance based transfer forecasting framework. Specifically, we fist propose a single source based transfer regression algorithm and we lately extend it to multiple sources. Experiment results show that with the proposed methods, the forecasting algorithms can be significantly improved. In the future, we plan to investigate the possibility of jointly transferring both the learned models and as well as samples.

References

D Bunn and E Dillon Farmer. 1985. Comparative models for electrical load forecasting. (1985).

Ervin Ceperic, Vladimir Ceperic, and Adrijan Baric. 2013. A strategy for short-term load forecasting by support vector regression machines. *IEEE Transactions on Power Systems* 28, 4 (2013), 4356–4364.

Yongbao Chen, Peng Xu, Yiyi Chu, Weilin Li, Yuntao Wu, Lizhou Ni, Yi Bao, and Kun Wang. 2017. Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy* 195 (2017), 659–670.

Wenyuan Dai, Qiang Yang, Gui-Rong Xue, and Yong Yu. 2007. Boosting for transfer learning. In *Proceedings of the 24th international Conference on Machine Learning (ICML)*. ACM, 193–200.

LU Di, WANG Xinghua, ZY DONG, and PENG Xiangang. 2018. Online Gaussian Process Regression for Short-term Probabilistic Interval Load Prediction. In *2018 International Conference on Power System Technology (POWERCON)*. IEEE, 173–178.

Hongming He, Tao Liu, Ruimin Chen, Yong Xiao, and Jinfeng Yang. 2012. High frequency short-term demand forecasting model for distribution power grid based on ARIMA. In *2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, Vol. 3. IEEE, 293–297.

Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza. 2001. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems* 16, 1 (2001), 44–55.

InsideEVs. 2019. Avaliable [Online] https://insideevs.com/news/343633/global -electric-vehicles-sales-are-rising-exponentially/. (2019).

Nakyoung Kim, Minkyung Kim, and Jun Kyun Choi. 2018. LSTM Based Short-term Electricity Consumption Forecast with Daily Load Profile Sequences. In *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*. IEEE, 136–137.

Weicong Kong, Zhao Yang Dong, Youwei Jia, David J Hill, Yan Xu, and Yuan Zhang. 2017. Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid* (2017).

Hulisani Matsila and Pitshou Bokoro. 2018. Load forecasting using statistical time series model in a medium voltage distribution network. In *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*. IEEE, 4974–4979.

OPENEI. 2021. Available [Online]http://en.openei.org/doe-opendata/dataset. (2021).

Sinno Jialin Pan and Qiang Yang. 2010. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* 22, 10 (2010), 1345–1359.

David Pardoe and Peter Stone. 2010. Boosting for regression transfer. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*. 863–870.

REN21. 2019. Renewables 2019 Global Report. (2019).

Di Wu. 2018. *Machine Learning Algorithms and Applications for Sustainable Smart Grid*. McGill University (Canada).

Di Wu, Boyu Wang, Doina Precup, and Benoit Boulet. 2017a. Boosting based multiple kernel learning and transfer regression for electricity load forecasting. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 39–51.

Di Wu, Boyu Wang, Doina Precup, and Benoit Boulet. 2019. Multiple Kernel Learning based Transfer Regression for Electric Load Forecasting. *IEEE Transactions on Smart Grid* (2019).

Di Wu, Haibo Zeng, Chao Lu, and Benoit Boulet. 2017b. Two-stage energy management for office buildings with workplace EV charging and renewable energy. *IEEE Transactions on Transportation Electrification* 3, 1 (2017), 225–237.

Xiaoyu Wu, Jinghan He, Tony Yip, Ning Lu, et al. 2016. A two-stage random forest method for short-term load forecasting. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 1–5.

Yi Yao and Gianfranco Doretto. 2010. Boosting for transfer learning with multiple sources. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE conference on*. IEEE, 1855–1862.

Jianhua Ye and Li Yang. 2018. A Comparative Study of Ensemble Support Vector Regression Methods for Short-term Load Forecasting. In *2018 5th International Conference on Systems and Informatics (ICSAI)*. IEEE, 139–143.