Load Forecasting with Adversarial Attacks in Power Systems using DeepForge

Xingyu Zhou
Yi Li
Carlos A. Barreto
xingyu.zhou@vanderbilt.edu
yi.li@vanderbilt.edu
carlos.a.barreto@vanderbilt.edu
Vanderbilt University
Nashville, TN

Peter Volgyesi
Xenofon Koutsoukos
peter.volgyesi@vanderbilt.edu
Xenofon.Koutsoukos@vanderbilt.edu
Vanderbilt University, Institute for Software Integrated Systems
Nashville, TN

KEYWORDS

power systems, adversarial attacks, model-based design, testbed

ACM Reference Format:
Xingyu Zhou, Yi Li, Carlos A. Barreto, Peter Volgyesi, and Xenofon Koutsoukos. 2019. Load Forecasting with Adversarial Attacks in Power Systems using DeepForge. In *Hot Topics in the Science of Security Symposium (HotSoS), April 1–3, 2019, Nashville, TN, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3314058.3317729

1 INTRODUCTION

For electricity markets, the profit for a retailer is based on a dynamic balance between supply and demand. As a result, an accurate demand prediction system is essential for the pricing strategy to maximize the profit. The retailers may try to reduce the uncertainties by forecasting the demand of their customers. In particular, the accuracy of the forecasts improves aggregating the demand of large groups of customers [8].

With recent developments of machine learning techniques especially deep learning, neural network prediction models can be constructed to gain good results. From distributed meter readings of different areas, we can build a time series network model to predict the future demand after a certain amount of time.

However, the complexity of current black-box inference systems lead to vulnerabilities when exposed to adversarial attacks. Since the discovery of adversarial examples in CNNs from 2013 [9], security and robustness has become a hot topic [2] and valid concern. As a result, even though we can build a power load prediction model easily with current deep learning techniques, we still need to take a more cautious view on security issues beneath current machine learning applications.

One difficulty for viewing and testing these kinds of security risk lies in the domain-specific knowledge and state-of-art machine learning techniques. To address this issue in the power system CPS domain we followed a model-based design approach [4, 6, 7] and developed our forecasting method and security testing framework using DeepForge[1].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
HotSoS, April 1–3, 2019, Nashville, TN, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-7147-6/19/04.
https://doi.org/10.1145/3314058.3317729

2 MODEL-BASE FRAMEWORK DESIGN

DeepForge is a cloud-based development environment with deeply integrated domain specific modeling aspects created with WebGME[5]. It combines model integrated computing with rapid prototypical development of deep learning models.

DeepForge presents machine learning models with four hierarchical concepts: Pipelines, Operations, Executions and Jobs. Operations are atomic functions which accept inputs and return outputs.A Pipeline refers to a stack of tasks, such as data preprocessing, training or testing. Executing pipelines results in the creation of Executions. A Job corresponds to the selected operation combined with run status and running metadata.

As we can see, DeepForge provides a model-based representation to bridge the gap between domain-specific data and machine learning using the state-of-art Keras/TensorFlow framework.

To support the power domain we developed several general operations for data gathering and preprocessing from GridLAB-D, model training and testing. We developed a hybrid adversarial attack method based on FGSM [2].

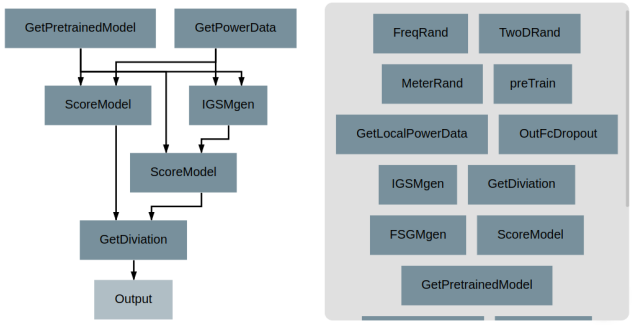


Figure 1: Atomic Operations and An Attack Pipeline Constructed.

Figure 1 shows the key atomic operations provided by our framework and the adversarial attack pipeline constructed using these components. For testing purposes we also embed a pre-trained prediction model.

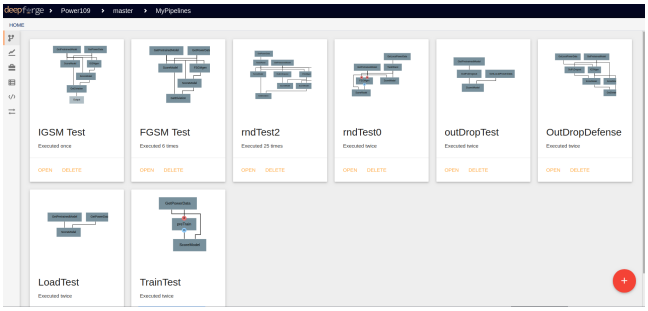


Figure 2: Fremework Pipelines.

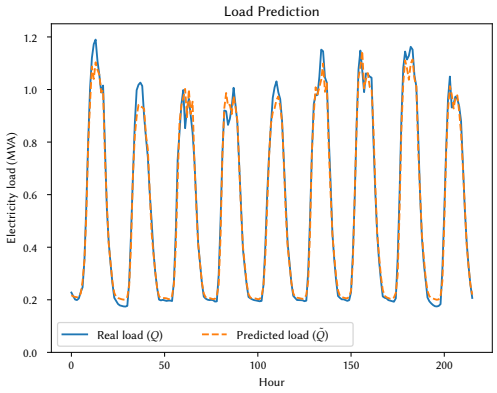


Figure 3: Illustration of the load prediction.

Figure 2 shows the developed pipelines ranging from simple train/test workflows to more complex attack/defense evaluations. All these pipelines can be created and modified easily.

3 PRELIMINARY RESULTS

Our internal prediction model is based on *recursive neural networks*. In particular, we use three Long-Short-Term Memory (LSTM) layers (150 units) and two fully connected layers (500 nodes) [3]. The input is the historical data from from 109 smart meters (both commercial and residential) and the temperature during two summer months. Measurements of the past 24 hours are used to predict the total load during the next hour, iteratively.

Fig. 3 shows an example of the load prediction, which uses 90% of the data for training and has a mean squared error (MSE) of 0.127 in the test data set.

We provide a hybrid attack combining L_0 and L_{inf} constraints. In practice, an attacker is allowed to modify Num meter readings each with a max deviation D. For each input $\boldsymbol{y}_0$, we choose the most sensitive meters using their gradient values. The algorithm is given in Algorithm 1. This is a simple approach combining different types of practical constraints and it provides a low computation complexity for fast prototypical tests.

Figure 4 shows the results using this simple attack scheme. The deviations (from true meter readings) increase from 1% to 10%. The ratio of the compromised meters increases from 0.05 to 0.50. This attack can serve as a baseline for further attack and defense settings.

Algorithm 1 l_0-FGSM Attack

Require: original observation $\boldsymbol{y}_0$, predictor f, l_∞ max deviation D, modification meter number limit Num.

1: $\Delta\boldsymbol{y} \leftarrow \mathbf{0}$, $i \leftarrow 0$
2: $grad \leftarrow \nabla f(\boldsymbol{y}_0)$
3: $grad \leftarrow allowModN(grad, Num)$
4: $\Delta\boldsymbol{y} \leftarrow D * sign(grad)$
5: **return** $\Delta\boldsymbol{y}$

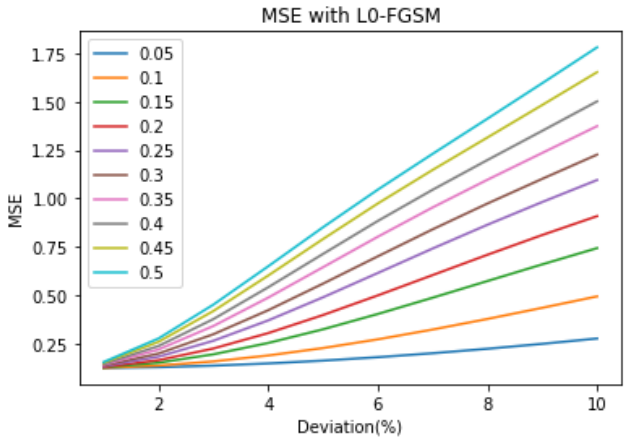


Figure 4: MSE under L0-FGSM Attacks.

The current attack models demonstrate the capabilities of the model-based framework and are good starting points for future explorations. We plan to develop more attack and defense strategies for practical applications.

REFERENCES

[1] Brian Broll and Jimmy Whitaker. 2017. DeepForge: An Open Source, Collaborative Environment for Reproducible Deep Learning. (2017).
[2] Ian J Goodfellow, Jonathon Shlens, and Christian Szegedy. 2014. Explaining and harnessing adversarial examples (2014). *arXiv preprint arXiv:1412.6572* (2014).
[3] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
[4] Xenofon Koutsoukos, Gabor Karsai, Aron Laszka, Himanshu Neema, Bradley Potteiger, Peter Volgyesi, Yevgeniy Vorobeychik, and Janos Sztipanovits. 2018. SURE: A modeling and simulation integration platform for evaluation of secure and resilient cyber–physical systems. *Proc. IEEE* 106, 1 (2018), 93–112.
[5] Zsolt Lattmann, Tamás Kecskés, Patrik Meijer, Gabor Karsai, Péter Völgyesi, and Ákos Lédeczi. 2016. Abstractions for modeling complex systems. In *International Symposium on Leveraging Applications of Formal Methods*. Springer, 68–79.
[6] Himanshu Neema, Janos Sztipanovits, Martin Burns, and Edward Griffor. 2016. C2WT-TE: A model-based open platform for integrated simulations of transactive smart grids. In *2016 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*. IEEE, 1–6.
[7] Himanshu Neema, Peter Volgyesi, Bradley Potteiger, William Emfinger, Xenofon Koutsoukos, Gabor Karsai, Yevgeniy Vorobeychik, and Janos Sztipanovits. 2016. SURE: An experimentation and evaluation testbed for CPS security and resilience: Demo abstract. In *Proceedings of the 7th International Conference on Cyber-Physical Systems*. IEEE Press, 27.
[8] Raffi Sevlian and Ram Rajagopal. 2018. A scaling law for short term load forecasting on varying levels of aggregation. *International Journal of Electrical Power & Energy Systems* 98 (2018), 350–361.
[9] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. 2013. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199* (2013).