**Machine Learning Methods for Solar Panel Power Estimation**

Invisible photovoltaic (PV) generation refers to the solar power capacity that is not monitored by, and thus not visible to, power grid. Mainly in the form of small-scale, photovoltaic modules, invisible solar power generation is a growing reality in power systems. Solar power basically acts as negative demand to the system, i.e., when it is available, it reduces the system’s total electricity demand. At a high penetration rate, solar power generation can thus impact the system’s network load pattern significantly. Network load is defined as the conventional electricity load minus the generated power by distributed energy resources (DERs). On the other hand, system network load is a key input when scheduling the short term operation of power systems. Thus, estimating the system net load in presence of significant invisible solar power generation is of interest. The project will seek to exploit this estimation to enable the design of scalable analysis of power systems.

The estimation can be done at different levels: appliance-level, transformer-level or distribution-level. Our focus is on transformer level. The goal is to use the estimation to manage the amount of load on each transformer. This procedure is called load/PV disaggregation. Our assumption is that under each transformer we have multiple PV panels and loads. The transformer-level disaggregation is performed for each individual transformer to forecast accurate total PV generation. The configuration of load and PVs could be very different from one another due to various dynamics (e.g., different area, the usage, etc.), which motivates the need for building more complex models than the simplified models currently used in the power engineering literature. This project will utilize context data such as weather condition and irradiance in order to perform context-aware disaggregation.

In order to realize the envisioned prediction models, the Study will develop and validate machine learning algorithms and tools with the following tasks:

* *Build Generalizable Models*: the heterogeneity of data makes it necessary to come up with models that are capable of learning useful features as well as capturing time dependency. Furthermore, we need to use methods with *confidence (or uncertainty) measures* in order to distinguish between anomalous and normal data. All of these make deep learning architectures a good choice to investigate and see what we can understand from field data. Deep learning introduces additional explicit and implicit learning priors in order to reduce the generalization error compared to traditional machine learning techniques. Specifically, the project will explore the use of convolutional deep networks to learn generalized features to overcome the heterogeneity of data for different cases (i.e. different load patterns, different level of DER injection).
* *Validation*: the project will validate the techniques to be developed both on simulated data and real datasets to assess the performance of the machine learning models. The project will specifically perform validation in the following order: i) load/PV disaggregation under each transformer; ii) reuse the disaggregation models to see how they can be generalized for other locations; and iii) repeat steps (i) and (ii) after each iteration of model refinement.

**Readings:**

[1] Zhang, Y., Beaudin, M., Taheri, R., Zareipour, H. and Wood, D., 2015. Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators. IEEE Transactions on Smart Grid, 6(5), pp.2253-2262.

[2] Shaker, H., Zareipour, H. and Wood, D., 2016. A Data-Driven Approach for Estimating the Power Generation of Invisible Solar Sites. IEEE Trans. Smart Grid, 7(5), pp.2466-2476.

[3] Ogliari, E., Niccolai, A., Leva, S. and Zich, R., 2018. Computational intelligence techniques applied to the day ahead pv output power forecast: phann, sno and mixed. Energies, 11(6), p.1487.

[4] Yona, A., Senjyu, T., Saber, A.Y., Funabashi, T., Sekine, H. and Kim, C.H., 2008, July. Application of neural network to 24-hour-ahead generating power forecasting for PV system. In 2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century (pp. 1-6). IEEE.

[5] Banerjee, P., Rizvi, S.M., Srivastava, A.K., Markham, P. and Patel, M., 2018, June. Data-Driven Static Load Model Parameter Estimation with Confidence Factor. In 2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS) (pp. 1-6). IEEE.

[6] Che, Z., Purushotham, S., Cho, K., Sontag, D. and Liu, Y., 2018. Recurrent neural networks for multivariate time series with missing values. Scientific reports, 8(1), p.6085.

[7] Cui, Z., Chen, W. and Chen, Y., 2016. Multi-scale convolutional neural networks for time series classification. arXiv preprint arXiv:1603.06995.