Privacy Preserving Synthetic Smart Meters Data

Ganesh Del Grosso
Joint work with: Georg Pichler, Pablo Piantanida
EDF: Georges Hebrail, Aurelien Delfosse, Amandine Pierrot, Benoit Grossin

September, 2020

Motivation

- Smart Meters Data is necessary for research and operational purposes.
- This data belongs to **clients** of a power company and cannot be released to third parties.

Motivation

- Smart Meters Data is necessary for research and operational purposes.
- This data belongs to **clients** of a power company and cannot be released to third parties.

We propose a method to generate "clean" synthetic data that:

- imitates the properties of the real data.
- cannot be traced back to particular **clients** present in the real dataset.

Motivation

- Smart Meters Data is necessary for research and operational purposes.
- This data belongs to **clients** of a power company and cannot be released to third parties.

We propose a method to generate "clean" synthetic data that:

- imitates the properties of the real data.
- cannot be traced back to particular **clients** present in the real dataset.

Today, we will briefly discuss:

- Data generation.
- Quality metrics.
- Privacy analysis.

Power Consumption Data

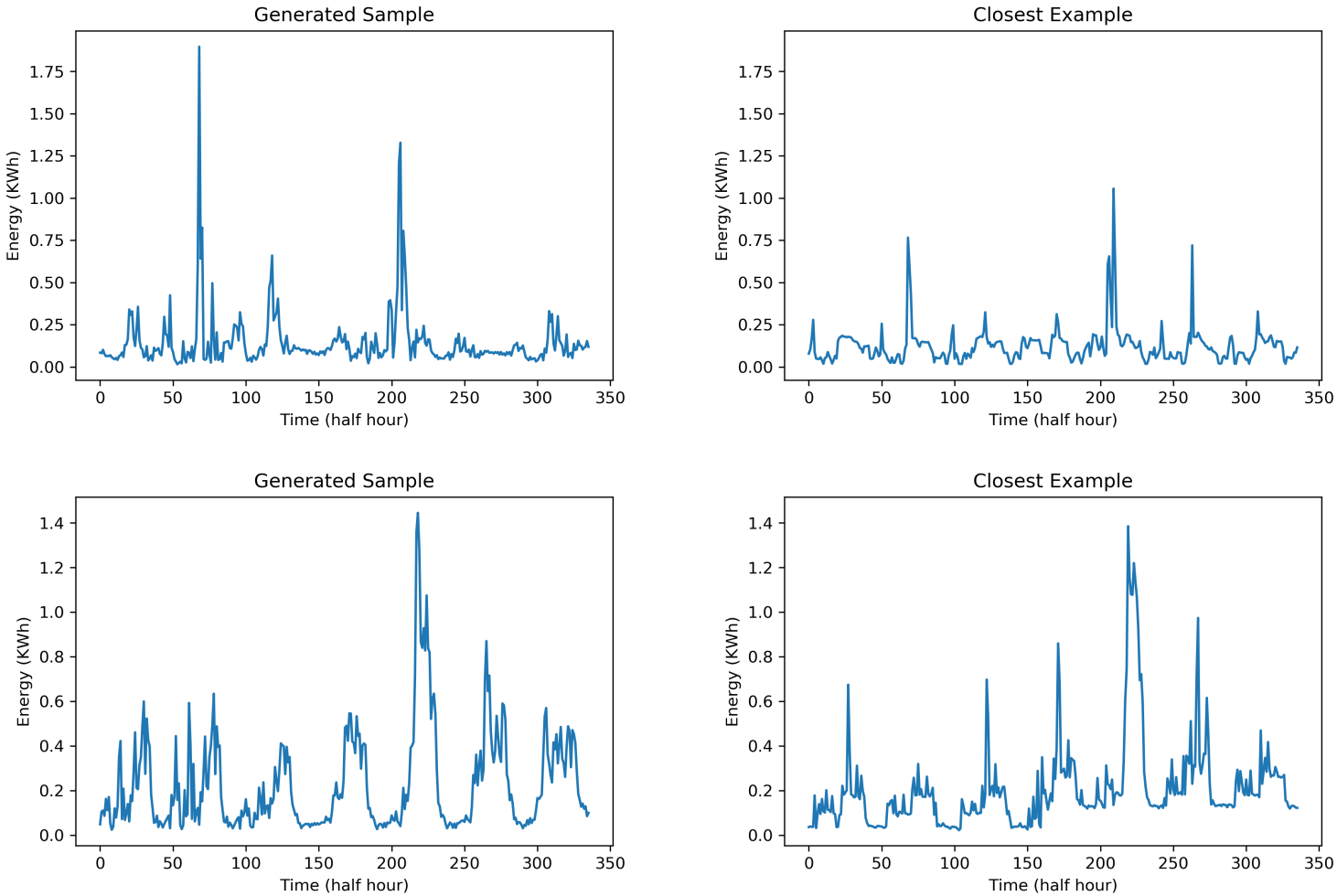
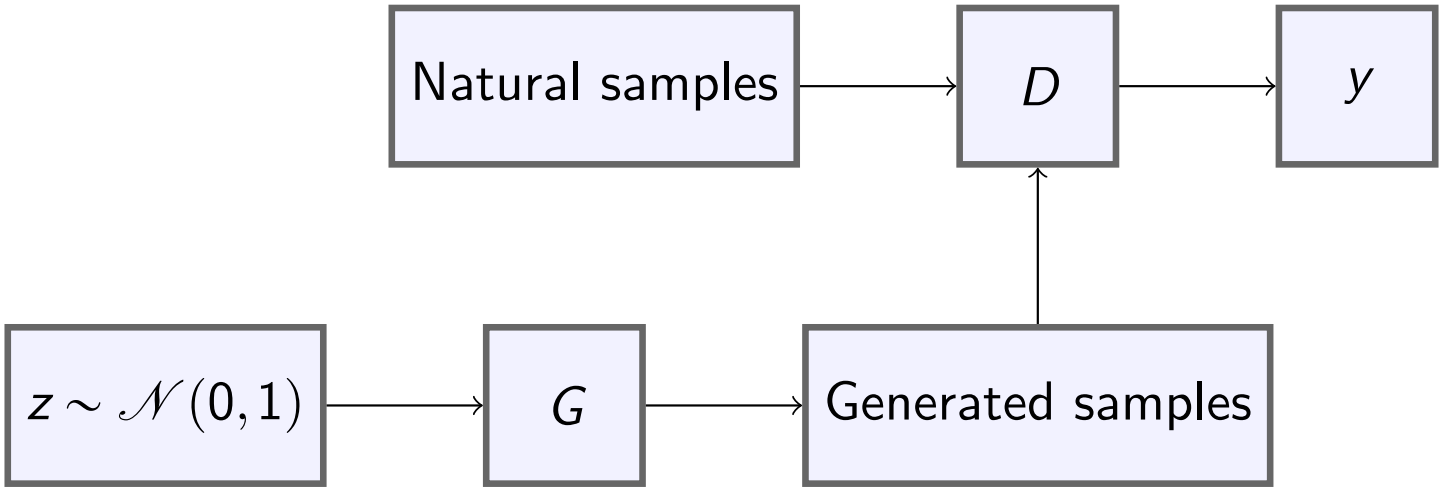


Figure: Generated curves (Left) compared to the respective closest Natural curves (Right).

Review: Generative Adversarial Networks

We approach the present problem via Generative Adversarial Networks (GANs) [1].



The GAN algorithm consists in a min-max game between a Generator and a Discriminator Network:

$$\min_G \max_D \mathbb{E}_{x\sim p_x}[\log(D(x))] + \mathbb{E}_{z\sim p_z}[\log(1 - D(G(z)))]$$

Quality Metrics

For evaluation of performance we consider 5 metrics or "indicators":

- Mean.
- Coefficient of Variation.
- Skewness.
- Kurtosis.
- Maximum-mean ratio.

Quality Metrics

We evaluate the quality of artificial curves via a prediction task [2].

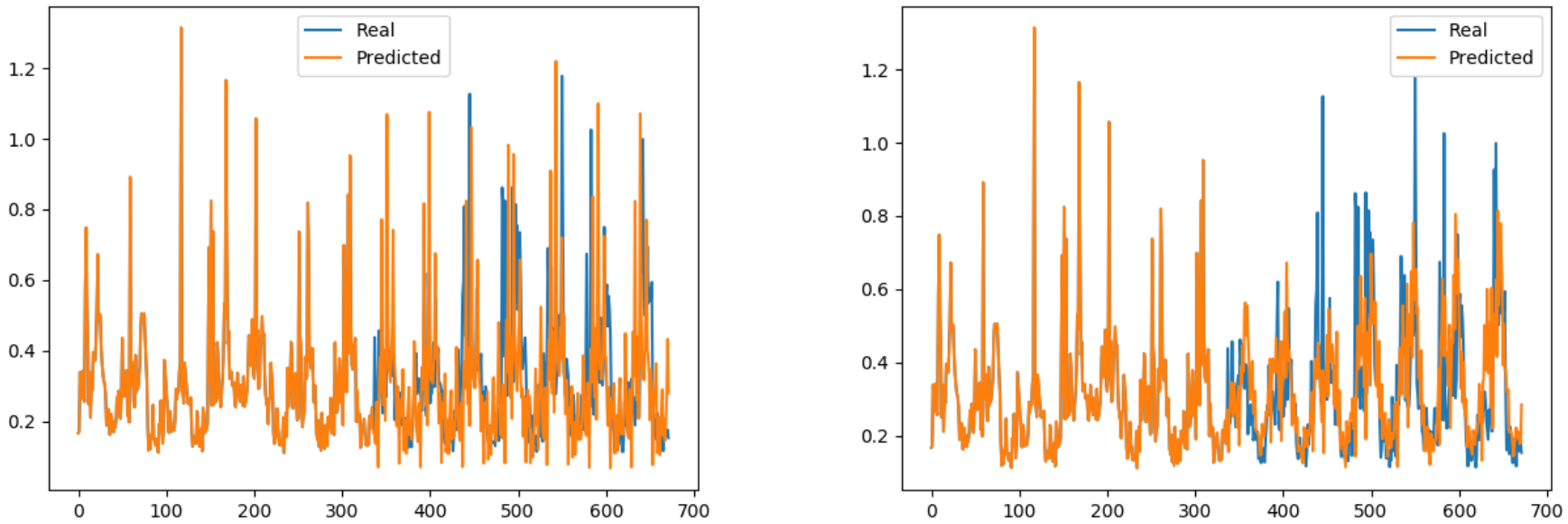


Figure: Prediction using a model trained with natural curves (Left). Prediction using a model trained with artificial curves (right).

Membership Inference Attacks

- Membership inference attacks measure the leakage of sensitive information from a model about its training set [3].

Membership Inference Attacks

- Membership inference attacks measure the leakage of sensitive information from a model about its training set [3].
- The attacker tries to determine whether or not a record, or set of records, belongs to the training set of the model.

Membership Inference Attacks

- Membership inference attacks measure the leakage of sensitive information from a model about its training set [3].
- The attacker tries to determine whether or not a record, or set of records, belongs to the training set of the model.

Algorithm 3 Likelihood Attack

Input: Universe $\mathscr{U}$, its partition $P_{\mathscr{U}}$, discriminator D.
Output: Index of the member of the training set i.
initialize $\mathrm{likelihoodList} \leftarrow []$
for each set **in** $P_{\mathscr{U}}$ **do**
    $\bar{p} \leftarrow$ **mean**($[D(\mathrm{curve})$ **for each** curve **in** set$]$)
    $\mathrm{likelihoodList}$**.append**$(\bar{p})$
end for
$i \leftarrow$ **argmax**$(\mathrm{likelihoodList})$
return i

Gradient-norm Regularization

Recall the loss function of the discriminator network,

$$L_{basic}(\theta_d) = -\frac{1}{m}\sum_{i=1}^{m}[log(D_{\theta_d}(x^{(i)}) + log(1 - G_{\theta_g}(z^{(i)}))],$$

Now consider the regularized loss function,

$$L_{reg}(\theta_d) = L_{basic}(\theta_d) - \gamma\|\nabla_{\theta_d} L_{basic}(\theta_d)\|_2 \ ,$$

where the second term is the L2-norm of the gradient of the original loss function with respect to the discriminator parameters.

More details in our upcoming paper...

References I

[1] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," arXiv:1406.2661 [cs, stat], June 2014.
arXiv: 1406.2661.

[2] A. Borji, "Pros and Cons of GAN Evaluation Measures," arXiv:1802.03446 [cs], Feb. 2018.
arXiv: 1802.03446.

[3] M. Nasr, R. Shokri, and A. Houmansadr, "Comprehensive Privacy Analysis of Deep Learning: Stand-alone and Federated Learning under Passive and Active White-box Inference Attacks," arXiv:1812.00910 [cs, stat], Dec. 2018.
arXiv: 1812.00910.