FEDERATED SHIFT-INVARIANT DICTIONARY LEARNING ENABLED DISTRIBUTED USER PROFILING

Qiushi Huang, Wenqian Jiang, Jian Shi, Chenye Wu, Dan Wang and Zhu Han

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

UNDERSTANDING USER BEHAVIOR IS A KEY ENABLER

- Active distribution network is crucial to the operation of a power grid with high penetration of renewable energies.

- And understanding user behaviors is a key enabler of active distribution network.

- How to understand user behaviors?
- What are the practical challenges?

Source: https://www.varonis.com/blog/what-is-user-behavior-analytics

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

CONVENTIONAL WISDOM

- The conventional wisdom is to conduct user profiling (i.e., clustering).

- Though effective, it faces several practical obstacles.

- Clustering requires massive data and each single data holder may not have enough data to understand all kinds of users.
- Also, data are the crucial assets for each data holder. Hence, data holders may not want to directly exchange the data with each other.

Source Up Inc.

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

OUR IDEA

- Distributed User Profiling Based on Federated Learning & Dictionary Learning

- Our idea is based on a simple observation:
  - Each user's data are the energy consumption aggregation of all its appliances.
  - Hence, the energy consumption patterns of those appliances can serve as the dictionary, which is used to construct the user's energy consumption data.
  - If the dictionary is rich, then the user profiling can be done more accurately.

- Hence, we propose that each data holder first conducts the dictionary learning based on its own data, and then exchange the dictionary with each other based on the federated learning framework.

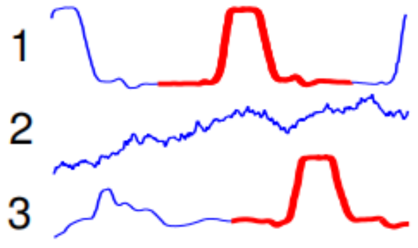
Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

STEP 1:
SHIFT-INVARIANT DICTIONARY LEARNING FOR USER PROFILING

Learn dictionary based on end users' load series

Cluster the derived pattern matrices D

Intuitive Idea



Time Warp

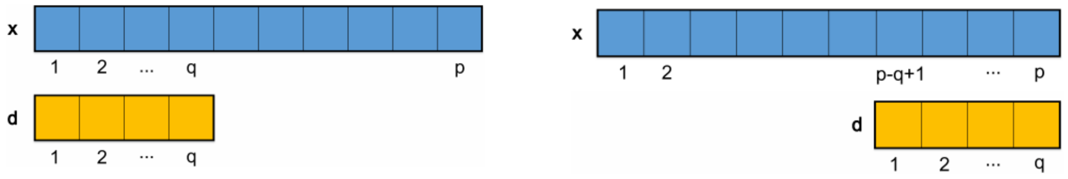
Making basis vector shift-invariant

Shifting Operator

$$\Psi(d_k,[\beta]_k)=\begin{cases}[d_k]_{t-[\beta]_k}, if\ 1\le t-[\beta]_k\le q\\ 0, otherwise\end{cases}$$

Shift-Invariant Dictionary Learning

$$\min_{D,\alpha_j,[\beta_j]_k}\sum_{j=1}^{J}\frac{1}{2}\left\|x_j-\sum_{k=1}^{K}[\alpha_j]_k\Psi\left(d_k,[\beta_j]_k\right)\right\|_2^2+\lambda\sum_{j=1}^{J}\left\|\alpha_j\right\|_1$$



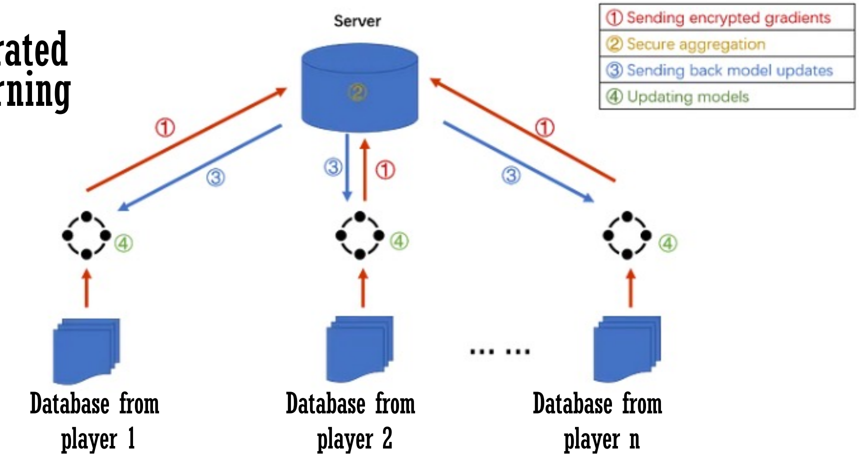
The same basis with different shifts

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

STEP 2:
FEDERATED SHIFT-INVARIANT DICTIONARY LEARNING CLUSTERING

- **Incorporate federated learning**
- **Enable distributed user profiling without exchanging the datasets with each other**

Federated Learning



Solution Procedure:

1. Local user dictionary learning

$$\min_{m^* \in [1,\cdots,M], \alpha_j} \sum_{x_j \in P_u} \frac{1}{2} \left\| x_j - \sum_{k=1}^{K} [\alpha_j]_k \Psi\left(d_k^{m^*}, [\beta_j]_k\right) \right\|_2^2 + \lambda \left\| \alpha_j \right\|_1$$

2. Local dictionary updating

$$\min_{d_k^{m^*}} \sum_{x_j \in P_u} \frac{1}{2} \left\| x_j - \sum_{k=1}^{K} [\alpha_j]_k \Psi\left(d_k^{m^*}, [\beta_j]_k\right) \right\|_2^2$$

3. Dictionary aggregation (federated averaging)

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.

NUMERICAL EVALUATION

Dictionary learning helps

TABLE I: Clustering Performance Comparison.

Evaluation	SIDL-clustering	K-means	GMM	K-shape	SC
h-score	0.9203	0.8714	0.7253	0.8942	0.8906
c-score	0.9453	0.9053	0.7451	0.9070	0.8934
v-measure	0.9322	0.8876	0.7349	0.9004	0.8920

- Dictionary learning better reveals the load series patterns.

Distributed implementation comes with a cost.

TABLE II: Performance Comparison between Centralized SIDL Clustering and Federated SIDL Clustering.

Method	*v-measure*	h-score	c-score	Comp. Time
Central	0.9322	0.9203	0.9453	264min
Federated	0.9115	0.8972	0.9269	52min

- The federated framework significantly improves the computation efficiency at the cost of slightly reduced v-measure, h-score and c-score.

Q. Huang, W. Jiang, J. Shi, C. Wu, D. Wang and Z. Han, "Federated Shift-Invariant Dictionary Learning Enabled Distributed User Profiling," in IEEE Transactions on Power Systems, doi: 10.1109/TPWRS.2023.3296976.