Multi-Label Consistent Convolutional Transform Learning: Application to Non-Intrusive Load Monitoring

Shikha Singh[1], Jyoti Maggu[1], Angshul Majumdar[1 2], Emilie Chouzenoux[3] and Giovanni Chierchia[4]

[1]Indraprastha Institute of Information Technology, Delhi.

[2]TCS Research Lab, India

[3]Université Paris-Saclay, CentraleSupélec, Inria, CVN, Gif-sur-Yvette, France

[4]Université Paris East, ESIEE, Noisy-le-Grand, France

ICASSP'20

Introduction: Non-Intrusive Load Monitoring(NILM)

1. NILM is a process to estimate the power consumed by the devices that are on in a given period of time, from the aggregate meter reading.
2. Broader goal of NILM is to feedback the appliance level energy consumption information to the users.
3. Most of the existing techniques use historical appliance-level data which makes the monitoring intrusive.
4. Recently, NILM has been framed as a multi-label classification problem to circumvent this problem.

Problem Statement

Energy disaggregation via simultaneous state detection

$AL \qquad X_{washer} \qquad X_{dishwasher} \qquad X_{desktop}$

'OFF' $state = 0$ 'ON' $state = 1$

X_{washer} = $0/1 \times$ Average power consumption of washer

$X_{dishwasher}$ = $0/1 \times$ Average power consumption of dishwasher

$X_{desktop}$ = $0/1 \times$ Average power consumption of desktop

Previously Proposed Techniques

- Multi-Label k-Nearest Neighbours(ML-kNN),
- Random k-Label sets(RaKel), and
- Multi-Label Consistent Deep Dictionary Learning(MLC-DDL)

Background: Convolutional Transform Learning(CTL)

- CTL was proposed for feature generation based on learnt convolutions, t_m

$$\underset{t_m, x_m^{(k)}}{\arg\min} \frac{1}{2} \sum_{k=1}^{K} \sum_{m=1}^{M} \left\| t_m * s^{(k)} - x_m^{(k)} \right\|_2^2 + \mu \left\| T \right\|_F^2 - \lambda \log\det(T) + \beta \left\| x_m^{(k)} \right\|_1 + \iota_+ \left(x_m^{(k)} \right) \quad (1)$$

where $T = [t_1 | \ldots | t_m]$, s^k is k^{th} data vector, $x_m^{(k)}$ is the corresponding feature vector and ι_+ denotes the indicator function of the positive orthant, equals to 0 if its input has positive entries, and $+\infty$ otherwise.

Convolutional Transform Learning

$$\begin{aligned}\underset{T,X}{\arg\min}\, F(T,X) = \frac{1}{2}\sum_{k=1}^{K}\left\|S^{(k)}T - X_k\right\|_F^2 + \beta\|X\|_1 \\ + \iota_+(X) + \mu\,\|T\|_F^2 - \lambda\log\det(T)\end{aligned} \tag{2}$$

where

$$\begin{aligned}S^{(k)}T &= \left[t_1 * s^{(k)}|\ldots|t_M * s^{(k)}\right] \\ X_k &= \left[x_1^{(k)}|\ldots|x_M^{(k)}\right]\end{aligned}$$

with $S^{(k)}$ being the Toeplitz matrix of the convolution and $X = [X_1^\top \;\ldots\; X_K^\top]^\top$

Proposed Algorithm: Multi-Label Consistent Convolutional Transform Learning(MLC-CTL)

$$\begin{aligned}\underset{T,X,M}{\arg\min}\, F(T,X,M) = \frac{1}{2}\sum_{k=1}^{K}\left\|S^{(k)}T - X_k\right\|_F^2 \\ + \beta\|X\|_1 + \iota_{[0,+\infty[}(X) + \mu\,\|T\|_F^2 - \lambda \log\det(T) \\ + \eta\|Q - MX\|_F^2. \end{aligned} \tag{3}$$

Q is the binary-encoded class labels and M is the mapping between labels, Q and the features, X .

Test Phase

- The test phase consists in applying the update rule for X on the test data.
- The generated features are projected by the learnt M onto the label space.
- In practice the generated label map may not be binary, but it is real valued.
- We threshold it to find the active classes, using a threshold value of 0.5.

Datasets and Metrics

- Datasets: REDD and Pecan Street
- F1 Score

$$F1 = \frac{2 \times TP}{2 \times TP + FN + FP}$$

  TP - number of true positives; FN - false negatives and FP - number of false positives.

- Appliance-level Energy Error/ Normalized energy error (NEE)

$$NEE = \frac{\sum_t |P_t^n - \hat{P}_t^n|}{\sum_t P_t^n}$$

  P_t^n - power consumption of the appliance n at any time instant t; $\hat{P}_t^n$ - predicted power consumption at the same instance.

Results

Table 1: **Performance Evaluation on REDD**

Method	Macro F1-Score	Micro F1-Score	Energy Error
RAkEL	0.6579	0.6616	0.7572
MLkNN	0.7153	0.7160	0.1101
MLC-DDL	0.6432	0.6435	0.1773
MLC-CTL	**0.7505**	**0.7611**	**0.1072**

Results

Table 2: **Performance Evaluation on Pecan Street**

Method	Macro F1-Score	Micro F1-Score	Energy Error
RAkEL	0.6520	0.6552	0.4349
MLkNN	0.6127	0.6133	0.0259
MLC-DDL	0.6847	0.6848	0.6373
MLC-CTL	**0.7240**	**0.7304**	**0.0021**

Results

Table 3: **Appliance-Level Evaluation on REDD**

Device	**RAkEL**		**MLkNN**		**MLC-DDL**		**MLC-CTL**	
	F1-Score	Error	F1-Score	Error	F1-Score	Error	F1-Score	Error
Lighting	0.6998	0.5690	0.6790	0.1432	0.7482	0.1818	**0.7629**	**0.1217**
Kitchen	0.6599	0.8110	0.6667	0.2390	0.5683	0.1807	**0.7027**	**0.1158**
Refrigerator	0.7025	0.3482	0.7368	0.4256	0.7049	0.1443	**0.7395**	**0.0062**
Washer Dryer	0.5693	0.7860	0.7527	0.0914	0.5513	0.1768	**0.7969**	**0.1074**

Results

Table 4: **Appliance-Level Evaluation on Pecan Street**

Device	**RAkEL**		**MLkNN**		**MLC-DDL**		**MLC-CTL**	
	F1-Score	Error	F1-Score	Error	F1-Score	Error	F1-Score	Error
Air Conditioner	0.6155	0.2051	0.6552	0.0494	0.6756	0.6231	**0.7096**	**0.0106**
Dishwasher	0.6993	0.8393	0.6564	0.1152	0.7048	0.6318	**0.7382**	**0.0082**
Furnace	0.6773	0.5557	0.5864	0.0764	0.7117	0.6639	**0.7540**	**0.0119**
Microwave	0.6164	0.3982	0.5510	0.0040	0.6467	0.6338	**0.6942**	**0.1235**

Conclusion

- One does not need appliance-level data to train the model.
- This model yields the best results in term of classification accuracy and comparable results regarding energy disaggregation.
- It is a generic approach and can be used to solve any multi-label classification problem.

Thank you!