Convolutional Neural Network for Appliance Recognition in Energy Disaggregation (NILM)

Anthony FAUSTINE(NM-AIST), Prof.Nerey Mvungi(UDSM), Dr. Kisangiri Michael(NM-AIST) and Dr. Shubi Kaijage(NM-AIST).

Data Science in Africa Workshop (DSA2017) - Arusha, Tanzania.

21 July 2017

Outline

Presenter Bio

- PhD student at Nelson Mandela African Institution of Science and Technology,
- **Research :** Applied machine learning and signal processing for computational sustainability.
  - Hybrid HMM-DNN for energy dis-aggregation problem.
- co-founder pythontz [https://pythontz.github.io]
- ass.Lecturer : the University of Dodoma
- blog : [https://sambaiga.github.io]

Energy Disaggregation Problem.

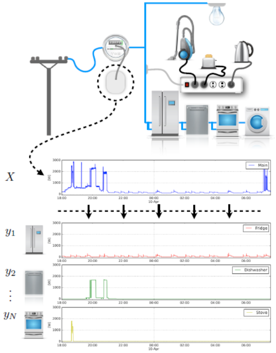
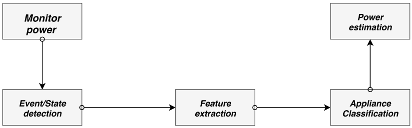


FIGURE – Huss A.(2015)

A source separation problem (signal processing problem) $\Rightarrow$ Separate aggregate power signal

$$y(t) = \sum_{t \in \{1,\ldots,T\}} x(t) + \sigma(t)$$

into all source (appliance) signals.
$x(t) : t \in \{1 \ldots T\}$



survey-paper :`https://arxiv.org/abs/1703.00785`

Outline

NILM Algorithm Development

State-of-the art NILM algorith : Hidden Markov Model (HMM) vs Deep neural networks (DNN)

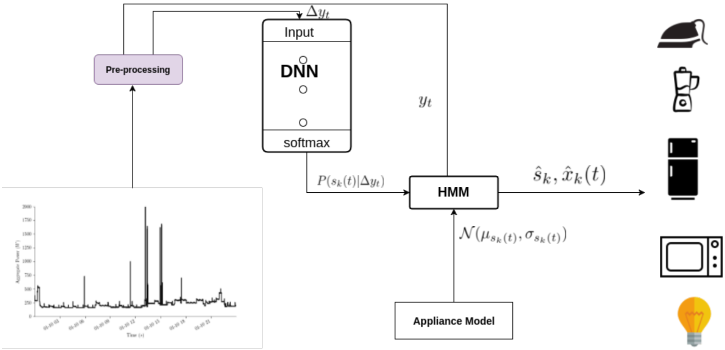
HMM [1]	DNN [2]
+ suitable for controlled multi-state loads + easy to train and can work in real-time - difficult to generalize to similar appliances - limited to few appliances	+ easier to generalize to similar appliances +very powerful - require lots of data for model training - training sensitive to hyperparameters

Open-Issue : Combine DNN and HMM for real-time and generalized energy disaggregation.

1. Makonin S., eta.l (2015), Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring.
2. Kelly J. eta.l (2015), Neural NILM : Deep Neural Networks Applied to Energy Disaggregation.

What Next

Hybrid DNN-HMM model



Aggregate Power

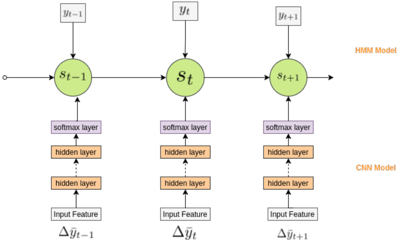
Hybrid CNN-HMM

Appliance Modeling

For each appliance k, the HMM parameters are :

$$\lambda^{(k)} = \{\pi^{(\mathbf{k})}, \mathbf{A}^{(\mathbf{k})}, \theta^{(\mathbf{k})}, \mathbf{B}^{(\mathbf{k})}\}$$

- $\pi^{(\mathbf{k})} \Rightarrow$ the initial probability of an appliance state $s_k(1)$ at time $t = 1$.
- $\mathbf{A}^{(\mathbf{k})} = P(s_k(t) = i | s_k(t-1) = j) \Rightarrow$ the transition probability.
- $\theta^{(\mathbf{k})} \sim \mathcal{N}(\mu_{s_k(t)}, \sigma_{s_k(t)}) \Rightarrow$ the appliance model.
- $\mathbf{B}^{(\mathbf{k})} = P(\Delta y_t | s_{k(t)} = j) \Rightarrow$ estimated from CNN $\Rightarrow \frac{P(s_k(t)|\Delta y_t)}{P(s_k(t))}$.



The CNN gets a window of $2B + 1$ of input features such that : $\Delta\hat{y}_t = [\Delta y_{max}(0, t - B), \ldots \Delta y_t, \ldots \Delta y_{min}(T, t + B)]$

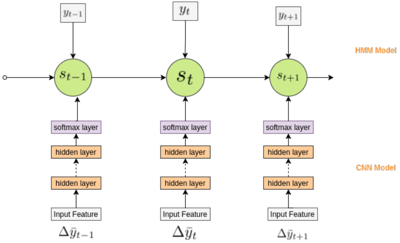
Hybrid CNN-HMM

Appliance Modeling

For each appliance k, the HMM parameters are :

$$\lambda^{(k)} = \{\pi^{(\mathbf{k})}, \mathbf{A}^{(\mathbf{k})}, \theta^{(\mathbf{k})}, \mathbf{B}^{(\mathbf{k})}\}$$

- $\pi^{(\mathbf{k})} \Rightarrow$ the initial probability of an appliance state $s_k(1)$ at time $t = 1$.
- $\mathbf{A}^{(\mathbf{k})} = P(s_k(t) = i | s_k(t-1) = j) \Rightarrow$ the transition probability.
- $\theta^{(\mathbf{k})} \sim \mathcal{N}(\mu_{s_k(t)}, \sigma_{s_k(t)}) \Rightarrow$ the appliance model.
- $\mathbf{B}^{(\mathbf{k})} = P(\Delta y_t | s_{k(t)} = j) \Rightarrow$ estimated from CNN $\Rightarrow \frac{P(s_k(t)|\Delta y_t)}{P(s_k(t))}$.



The CNN gets a window of $2B+1$ of input features such that : $\Delta\hat{y}_t = [\Delta y_{max}(0, t-B), \ldots \Delta y_t, \ldots \Delta y_{min}(T, t+B)]$

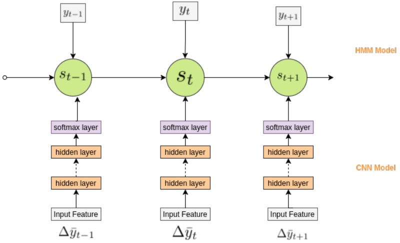
Hybrid CNN-HMM

Appliance Modeling

For each appliance k, the HMM parameters are :

$$\lambda^{(k)} = \{\pi^{(\mathbf{k})}, \mathbf{A}^{(\mathbf{k})}, \theta^{(\mathbf{k})}, \mathbf{B}^{(\mathbf{k})}\}$$

- $\pi^{(\mathbf{k})} \Rightarrow$ the initial probability of an appliance state $s_k(1)$ at time $t = 1$.
- $\mathbf{A}^{(\mathbf{k})} = P(s_k(t) = i | s_k(t-1) = j) \Rightarrow$ the transition probability.
- $\theta^{(\mathbf{k})} \sim \mathcal{N}(\mu_{s_k(t)}, \sigma_{s_k(t)}) \Rightarrow$ the appliance model.
- $\mathbf{B}^{(\mathbf{k})} = P(\Delta y_t | s_{k(t)} = j) \Rightarrow$ estimated from CNN $\Rightarrow \frac{P(s_k(t)|\Delta y_t)}{P(s_k(t))}$.



The CNN gets a window of $2B+1$ of input features such that : $\Delta\hat{y}_t = [\Delta y_{max}(0, t - B), \ldots \Delta y_t, \ldots \Delta y_{min}(T, t+B)]$

Hybrid CNN-HMM :Learning and Inference

Joint probability probability of all sequences :

$$P(Y, \Delta Y, S_k|\lambda) = \pi^{(k)}_{s_k(1)} \cdot \mathbf{B}^{(k)}_{s_k(1)} \cdot P(x_k(1) \leq y_1|s_k(1), \lambda)$$
$$\prod_{t=2}^{T} \mathbf{A}_{s_k(t), s_k(t-1)} \cdot P(x_k(t) \leq y_t|s_k(t), \lambda) \cdot \mathbf{B}_{s_k(t)}$$

Training :

- Bauch-welch algorithm under MLE to train initial GMM-HMM.
- Stochastic Gradient Descent (SDG) to train initial CNN.
- Embedded-Viterbi-algorithm to train CNN-HMM

Inference and Signal Extraction :

- virtebi algorithm : $\hat{s}_k = \arg\max_s [P(Y, \Delta Y, S_k|\lambda)]$
- Power estimation : $\hat{x}_k(t) = \mu_{s_k(t)}$

Hybrid CNN-HMM :Learning and Inference

Joint probability probability of all sequences :

$$P(Y, \Delta Y, S_k|\lambda) = \pi^{(k)}_{s_k(1)} \cdot \mathbf{B}^{(k)}_{s_k(1)} \cdot P(x_k(1) \leq y_1|s_k(1), \lambda)$$
$$\prod_{t=2}^{T} \mathbf{A}_{s_k(t), s_k(t-1)} \cdot P(x_k(t) \leq y_t|s_k(t), \lambda) \cdot \mathbf{B}_{s_k(t)}$$

Training :

- Bauch-welch algorithm under MLE to train initial GMM-HMM.
- Stochastic Gradient Descent (SDG) to train initial CNN.
- Embedded-Viterbi-algorithm to train CNN-HMM

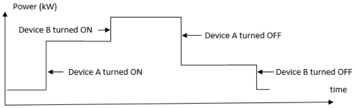
Inference and Signal Extraction :

- virtebi algorithm : $\hat{s}_k = \arg\max_{s}[P(Y, \Delta Y, S_k|\lambda)]$
- Power estimation : $\hat{x}_k(t) = \mu_{s_k(t)}$

Outline

Appliance recognition

Appliance recognition is an important sub-task of the NILM problem.



- Several approaches for this sub-task [3], [4]
- Deep-learning have received little attention.

---

3. Gao, Jingkun, et al (2015). "A feasibility study of automated plug- load identification from high-frequency measurements."
4. Karim Barsim, et al 2016. "Neural Neural Network Ensembles to Real-time Identification of Plug-level Appliance Measurement"

Objective

Goal Appliance recognition :
Apply CNN to recognize the labeled appliances once they are switched on.

Data : Plug load Appliance Identification Dataset (PLAID [5]).

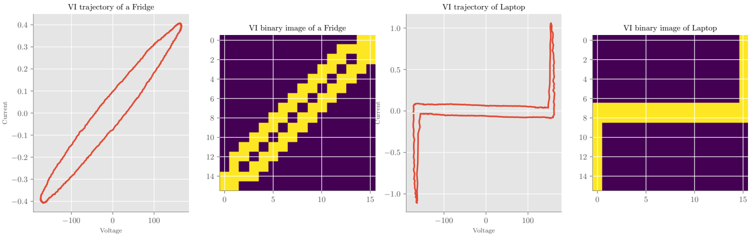
- 55 households in USA
- 11 different appliances
- sub-metered on events of the appliances (1074 total).
- Sampled at 30 kHz.

5. http ://plaidplug.com/

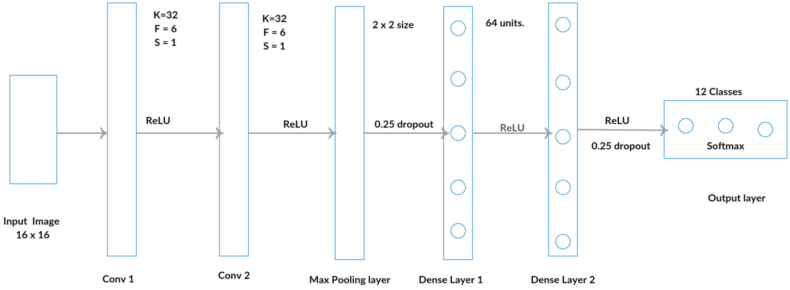
Appliance Signature

VI Binary Image Voltage-Current (VI) image feature during the steady state operation[6].

- Obtained by converting the VI trajectories into binary image.



Model Definition



Experment

- **Training** : Leave-house-out cross validation.
- **Metrics** : Precision (PR), Recall (RE), and F-Measure (F-1 score).

$$PR = \frac{TP}{TP + FP} \quad (1)$$

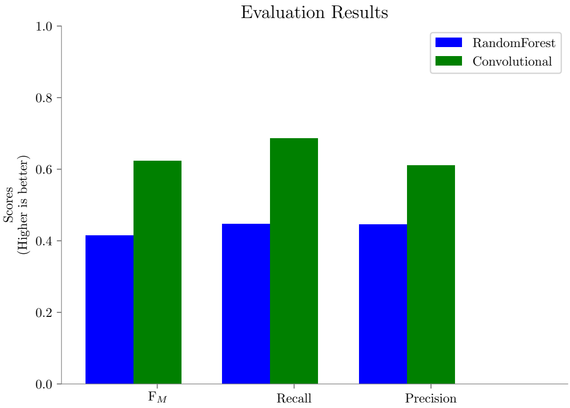
$$RE = \frac{TP}{TP + FN} \quad (2)$$

$$F_M = \frac{2 \times (PR \times RE)}{PR + RE} \quad (3)$$

where :

- TP $\Rightarrow$ correct claim the detected event was triggered by an appliance.
- FP $\Rightarrow$ incorrect claim that detected event was triggered by an appliance.
- FN $\Rightarrow$ indicates that appliance used was not identified.

Results



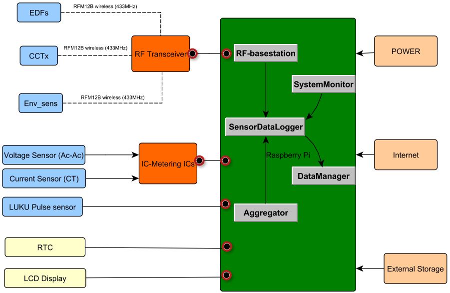
code : `https://github.com/sambaiga/cnn-appliance-detector`

Outline

Energy Data set Development

Develop tool and establish resource pertaining to residential electrical energy consumption-data set in Tanzania.

- RF-based WSN for individual appliances and aggregate power monitoring
- LUKU-pulse-sensor to collect aggregate power consumption using LED pulse found on existing LUKU meter.
- Experiment the tool in some buildings for one year ⇒establishment of the energy consumption data-set.



Outline

Conclusion

Open Challenges & Opportunities :

- Data, Data, Data
- NILM in renewable sources $\Rightarrow$ Improve battery energy storage.
- NILM to predict electrical fires accidents or solve electricity theft problem.
- Develop realistic simulators for simulating disaggregated electricity data.
- Explore different Deep Learning architecture for NILM problem.

THANK YOU