NILMTK v0.2: A Non-intrusive Load Monitoring Toolkit for Large Scale Data Sets

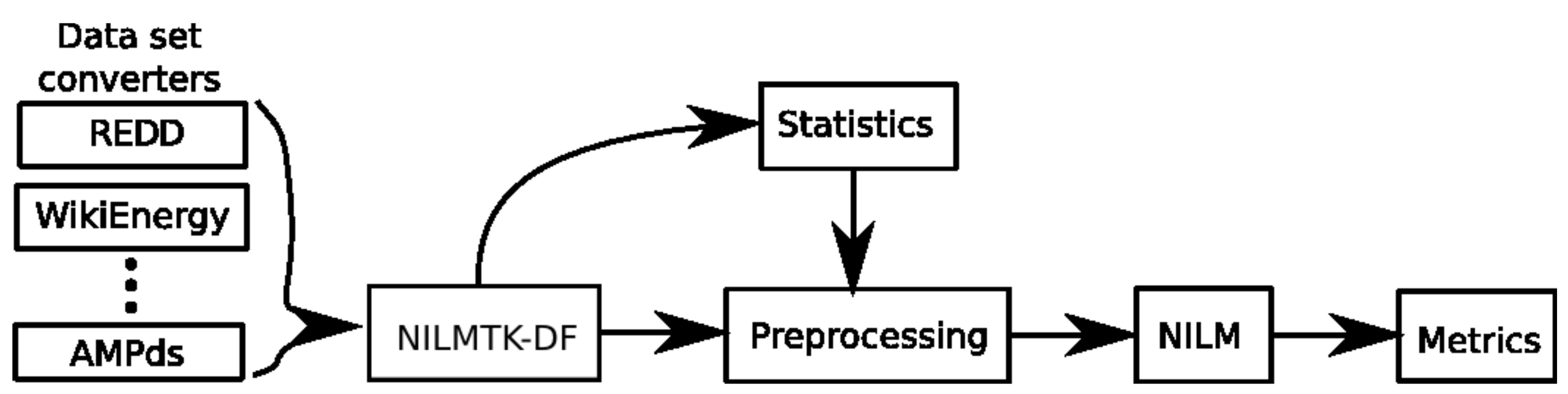
Jack Kelly[1], Nipun Batra[2], Oliver Parson[3], Haimonti Dutta[4], William Knottenbelt[1], Alex Rogers[3], Amarjeet Singh[2], Mani Srivastava[5]
[1]Imperial College London, [2] IIIT Delhi, [3] University of Southampton, [4] CCLS Columbia, [5] UCLA
nilmtk.github.io

Problems with NILM research

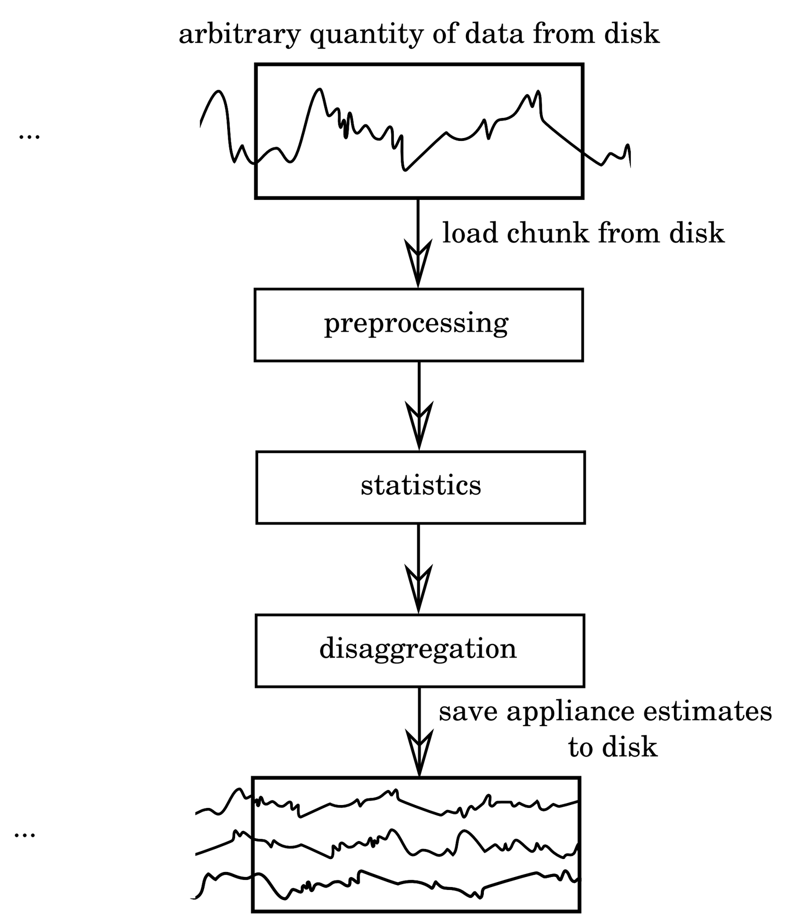
1. Different data sets used by each paper
2. No reference benchmark implementations available
3. Different metrics used by each paper

NILMTK: A toolkit for NILM research

NILMTK offers a complete pipeline from data sets to metrics:
- NILMTK defines a file format (NILMTK-DF) for NILM data
- Multiple dataset converters are included.
- Pipeline includes dataset statistics, preprocessing,
- training, disaggregation and NILM metrics.



Load arbitrarily large data sets

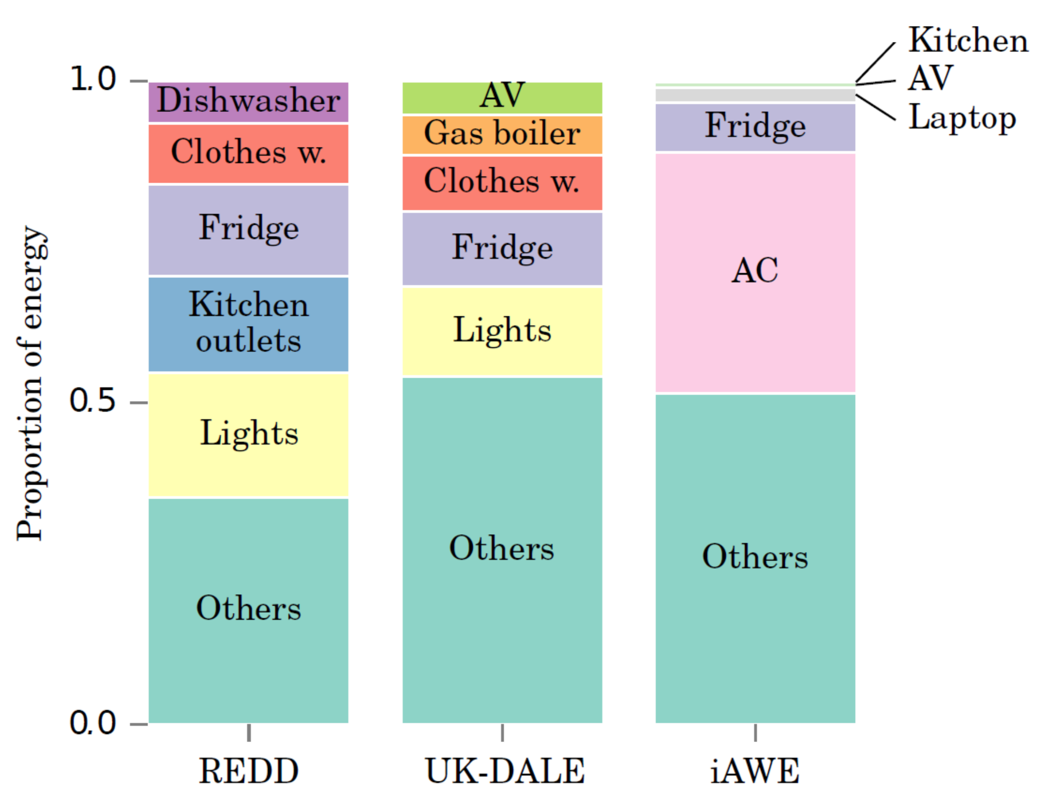


NILMTK v0.2 data set converters:

1. AMPds v2
2. COMBED
3. Dataport (was 'WikiEnergy')
4. ECO
5. GREEND
6. HES
7. iAWE
8. REDD
9. UK-DALE

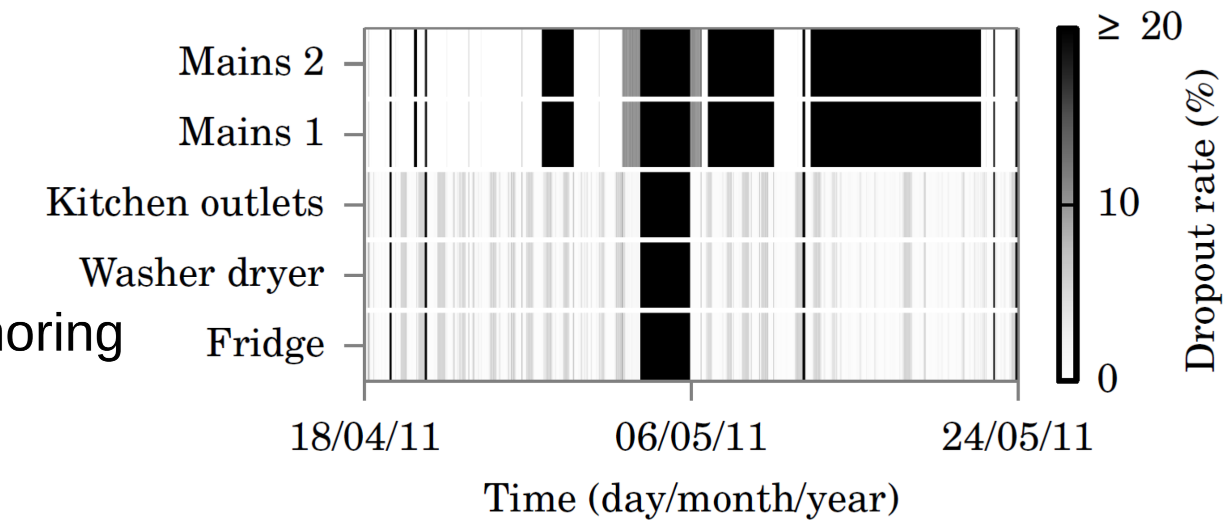
Data set statistics

- ON-OFF duration distribution
- Appliance usage distribution
- Appliance power distribution
- Correlation between sensor streams
- Find appliance contributions
- Percentage energy sub-metered
- Percentage of samples when energy sub-metered greater than threshold



Data set diagnostics

Common imperfections in data sets can be identified:
- Gaps
- Find continuous periods
- Dropout rate
- Dropout rate ignoring large gaps
- Up-time

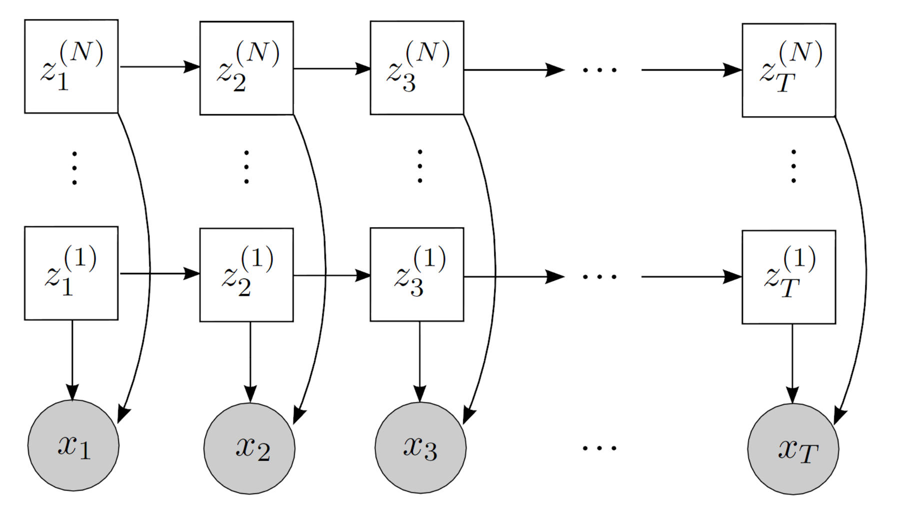


Benchmark algorithms

Combinatorial optimisation (CO): Finds combination of appliance states which sum to aggregate power demand

$$\hat{x}_t^{(n)} = \underset{\hat{x}_t^{(n)}}{\operatorname{argmin}} \left| \bar{y}_t - \sum_{n=1}^{N} \hat{y}_t^{(n)} \right|$$

Factorial hidden Markov model (FHMM): extends combinatorial optimisation to consider time dependencies between consecutive samples



George Hart 1985's algorithm

MLE

...and more coming...

```
disaggregator = FHMM()
disaggregator.train(training_data)
```

Performance metrics

- Error in total energy assigned
- Fraction total energy assigned correctly
- Normalised error in assigned power
- RMS error in assigned power
- Confusion matrix
- TP, FP, FN, TN
- Precision, recall
- F-score
- Hamming loss

$$\frac{\sum_t \left| y_t^{(n)} - \hat{y}_t^{(n)} \right|}{\sum_t y_t^{(n)}}$$

```
f_score(predicted_power, ground_truth_power)
```

Example results

- FHMM outperforms CO for 2 data sets
  - Uses state durations
- CO performs comparably to FHMM for 4 data sets
  - State durations add little value

Data set	F-score	
	CO	FHMM
REDD	0.31	0.31
Smart*	0.53	0.61
Pecan Street	0.77	0.77
AMPds	0.55	0.71
iAWE	0.73	0.73
UK-DALE	0.38	0.38