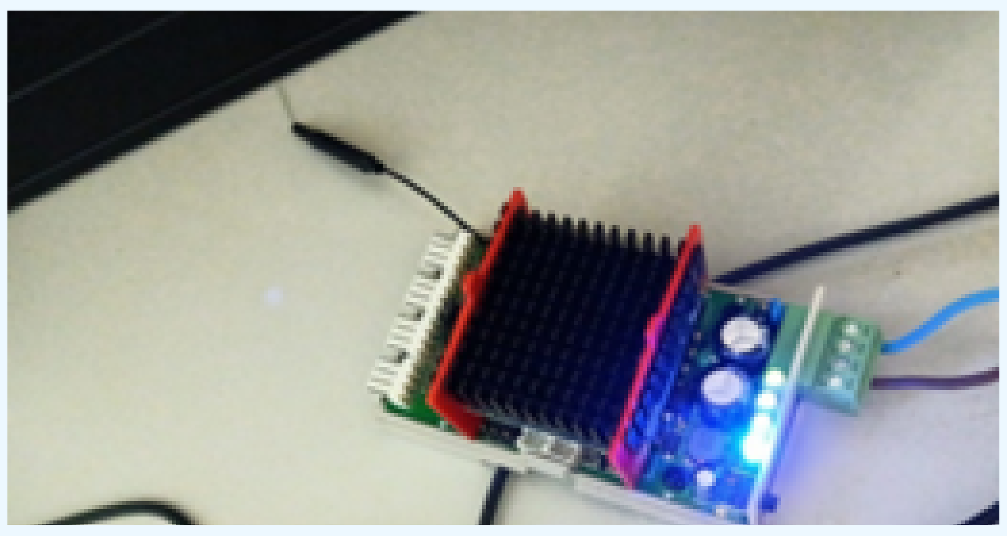
DSUALMH-A new high-resolution dataset for NILM



The OpenZMeter (OZM), is an open source, open hardware, multi-purpose precision smart meter capable of measuring a wide range of electrical variables at a high sampling rate. The aim of this work is to showcase the use and potential applications of the new high sampling frequency data provided by the OZM device, which are much more accurate than those obtained with other low-cost electrical meters. The study includes a comparison of the performance of two widely used disaggregation algorithms, Combinatorial Optimisation (CO) and Factorial Hidden Markov Model (FHMM), using metrics for different cases, as well as the incorporation of transients and comparison with other public datasets.

Authors

C. Rodriguez-Navarro 1, A. Alcayde 1, V. Isanbaev1, L. Castro-Santos 2, A. Filgueira-Vizoso3, F.G. Montoya 1.

Affiliations

1 Universidad de Almería, Escuela Superior de Ingeniería, La Cañada de San Urbano, 04120, Almería, Spain; email: crn565@inlumine.ual.es, aalcayde@ual.es, vs613@ual.es, pagilm@ual.es

2 Universidade da Coruña, Campus Industrial de Ferrol, Departamento de Enxeñaría Naval e Industrial, Escola Politécnica de Enxeñaría de Ferrol, Esteiro, 15471 Ferrol, Spain; email: laura.castro.santos@udc.es

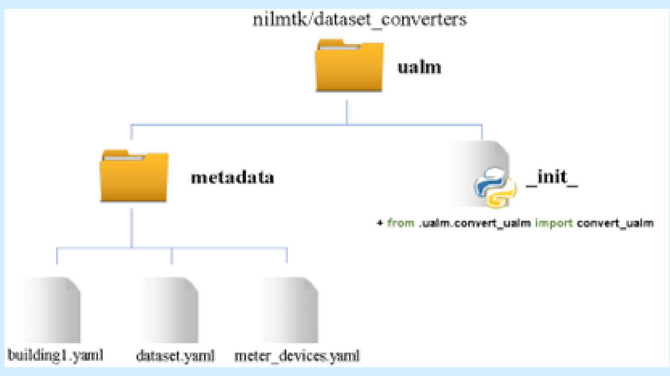
3 Universidade da Coruña, Campus Industrial de Ferrol, Departamento de Química, Escola Politécnica de Enxeñaría de Ferrol, Esteiro, 15471 Ferrol, Spain; email: almudena.filgueira.vizoso@udc.es

Introduction

NILM is a technique that estimates individual appliance consumption from a centralised meter, gaining importance in the current energy crisis. The NILMTK framework reduces the entry barrier for new researchers.

OZM is a sophisticated, open-source multi-purpose power meter with IoT capabilities that measures various electrical variables at high frequency, including harmonics up to order 50.

This work aims to demonstrate the potential of energy desegregation by adapting NILMTK with OZM data. Two new converters were developed to create two new datasets, allowing consideration of voltage, current, power harmonics, and association with OZM metadata.



Related work

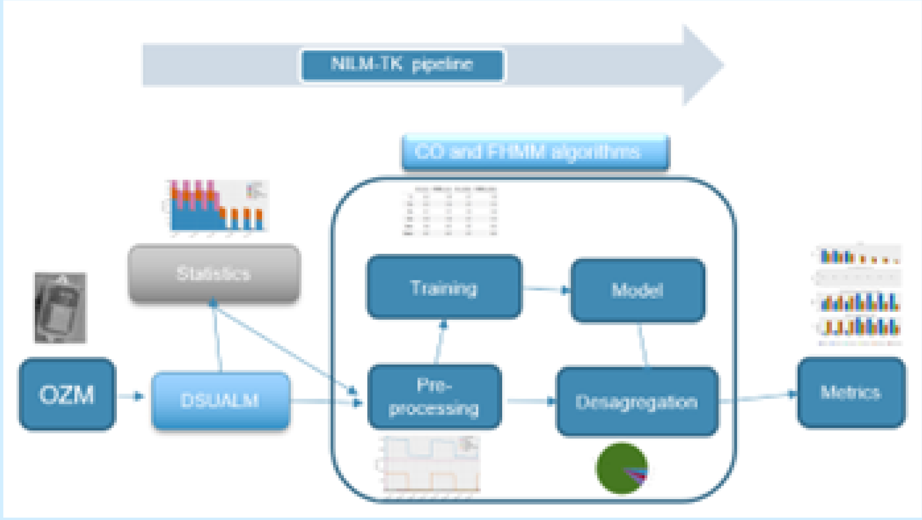
Algorithms and datasets for testing energy disaggregation methods can be classified into three groups: optimization methods, supervised methods, and unsupervised methods. Optimization methods include OBSA, genetic algorithms, and PSO. Supervised methods include Bayesian Classifiers, SVM, DDSC, ANN, and their extensions. Unsupervised methods include CO, HMM, and FHMM.

Public datasets available for testing include IAWE, REDD, UK-DALE, and DEPS.

The new DSUALM and DSUALMH datasets offer over 150 electrical variables at a high sampling rate for improving NILM research.

Methodology

In this work present a process for disaggregating data using the NILMTK Toolkit, which involves dataset curation and generation, data analysis, pretreatment, and validation. The OZM API is used to collect electrical characteristics data from devices, which is then processed into a new data format for use with NILMTK. The resulting datasets are analyzed using CO and FHMM supervised algorithms to train models, and the best model is used to predict appliance energy consumption with acceptable results.



Results

NILMTK has a calculation engine for evaluation metrics through the use of the MeterGroup for the validation of the results by means of the validation set. For this purpose, it is necessary to run different metrics on the models obtained, such as FEAC, F1, EAE, MNEAP and RMSE, which produce a output like this:

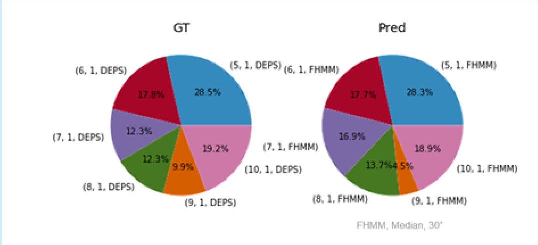
	fan	freezer	television	vacuum cleaner	boiler
F1	0.679	0.749	0.842	0.989	0.996
EAE	0.000	0.000	0.000	0.000	0.000
MNEAP	0.660	1.815	0.802	0.029	0.022
RMSE	20.600	62.253	24.520	27.240	32.839

Incorporating transients improves almost all metrics for most appliances. The fan and freezer stand out in particular. Regarding the TV, it would only get worse for F1, and for the hoover or the kettle, it would only get worse for RMSE.

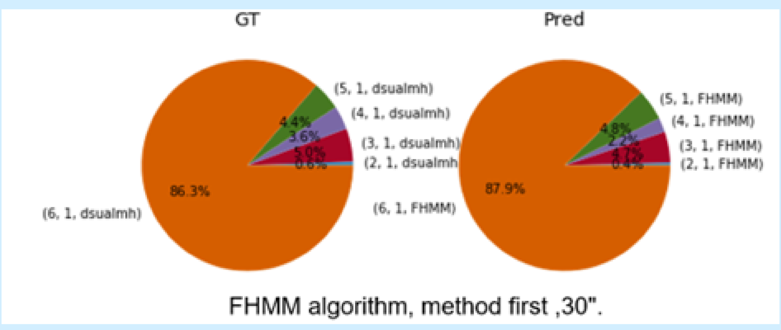
The CO algorithm with the Mean method is found to be the most efficient for the IAWE dataset, whereas for the DEPS dataset, the CO algorithm with the Mean method at a sampling time of half an hour performs better compared to only 10 seconds for the OZM data

	CO_mean	FHMM_mean	CO_median	FHMM_median	CO_first	FHMM_first
1s	11.08	69.63	10.72	54.70	10.33	59.98
10s	10.52	17.83	10.83	16.95	10.81	16.89
30s	10.69	13.27	10.65	12.92	10.52	13.20
60s	10.17	13.31	10.45	11.57	10.60	12.04
5min	8.84	9.88	8.85	9.97	8.86	9.60
10min	8.16	9.59	8.53	9.33	8.53	9.48
15min	7.61	8.62	7.18	8.64	7.08	7.62
30min	6.60	7.20	7.27	7.87	7.85	7.16

Likewise, if we compare GT and Pred for the DS of DEPS, the divergences are very important, ranging between 1.4%, 4.6% and 4.9% compared to 0 and 1.6% for DSUAL.



As can be seen in next figure the results for DSUALMH are quite acceptable in terms of predictions. For example, for the kettle (in blue) there is only a small deviation of 0.2% from the actual data. Likewise, both the fan (in red) and the light (pink) show minimal variation and the hoover (in orange) only shows a deviation of 1.6%



FHMM algorithm, method first ,30".

Conclusion

This work provides new tools and a dataset that facilitates access to Non-Intrusive Load Monitoring (NILM) for researchers using the OZM. The dataset includes voltage, current, power transients up to order 50, and 13-digit timestamp, which improves NILMTK metrics when compared to other datasets.

Results show better disaggregation and lower error rates. The aim is to involve the research community in using the dataset and tools for future improvements and continue incorporating new applications. The inclusion of high-frequency measurements and transients enhances NILMTK results for specific applications, providing opportunities for further research.

Acknowlegment

Research supported by Spanish Ministry of Science (grant PGC2018-098813-B-C33) and UAL-FEDER 2020 (Ref. UAL2020-TIC-A2080).

References

[1] H. George W., “Nonintrusive Appliance Load Monitoring,” Proceedings of the IEEE, vol. 80, no. 12, pp. 1870–1891, 1992, doi: 10.1109/5.192069.

[2] B. Nipun, K. Jack, P. Oliver, D. Haimonti, and K. William, “NILMTK: An Open Source Toolkit for Non-intrusive Load Monitoring,” Apr. 2014, doi: 10.1145/2602044.2602051.

[3] K. Jack, B. Nipun, P. Oliver, D. Haimonti, and K. William, “NILMTK v0.2: A non-intrusive load monitoring toolkit for large scale data sets,” in BuildSys 2014 - Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, Nov. 2014, pp. 182–183. doi: 10.1145/2674061.2675024.

[4] R. G. Pablo Salvador, “Implementación del algoritmo de análisis de demanda eléctrica por usos finales en las curvas de UVECASE,” 2020. doi: DOI: 10.13140/RG.2.2.27860.22405.

[5] V. Eduardo, A. Alfredo, M. Francisco G., B. Raul, and A.-C. Francisco M., “OpenZmeter: An efficient low-cost energy smart meter and power quality analyzer,” Sustainability (Switzerland), vol. 10, no. 11, Nov. 2018, doi: 10.3390/su10114038.

[6] B. Nipun, K. Rithwik, P. Ayush, M. Raktim, and K. Rajat, “A demonstration of reproducible state-of-the-art energy disaggregation using NILMTK,” in BuildSys 2019 - Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, Nov. 2019, pp. 358–359. doi: 10.1145/3360322.3360999.

21th International Conference on Renewable Energies and Power Quality (ICREPQ'23)
Madrid (Spain), 24th to 26th May 2023
Renewable Energy and Power Quality Journal (RE&PQJ)
ISSN 2172-038 X, Volume No.21, July 2023