

Big IoT data mining for real-time energy disaggregation in buildings

Citation for published version (APA):

Mocanu, D. C., Mocanu, E., Nguyen, H. P., Gibescu, M., & Liotta, A. (2017). Big IoT data mining for real-time energy disaggregation in buildings. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC 2016), 9-12 October 2016, Budapest, Hungary* (pp. 3765-3769). [7844820] Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/SMC.2016.7844820>

DOI:

[10.1109/SMC.2016.7844820](https://doi.org/10.1109/SMC.2016.7844820)

Document status and date:

Published: 09/02/2017

Document Version:

Accepted manuscript including changes made at the peer-review stage

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Big IoT data mining for real-time energy disaggregation in buildings

Decebal Constantin Mocanu*, Elena Mocanu*, Phuong H. Nguyen*, Madeleine Gibescu* and Antonio Liotta*

*Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

Email: {d.c.mocanu, e.mocanu, p.nguyen.hong, m.gibescu, a.liotta}@tue.nl

Abstract—In the smart grid context, the identification and prediction of building energy flexibility is a challenging open question, thus paving the way for new optimized behaviors from the demand side. At the same time, the latest smart meters developments allow us to monitor in real-time the power consumption level of the home appliances, aiming at a very accurate energy disaggregation. However, due to practical constraints is infeasible in the near future to attach smart meter devices on all home appliances, which is the problem addressed herein. We propose a hybrid approach, which combines sparse smart meters with machine learning methods. Using a subset of buildings equipped with subset of smart meters we can create a database on which we train two deep learning models, i.e. Factored Four-Way Conditional Restricted Boltzmann Machines (FFW-CRBMs) and Disjunctive FFW-CRBM. We show how our method may be used to accurately predict and identify the energy flexibility of buildings unequipped with smart meters, starting from their aggregated energy values. The proposed approach was validated on a real database, namely the Reference Energy Disaggregation Dataset. The results show that for the flexibility prediction problem solved here, Disjunctive FFW-CRBM outperforms the FFW-CRBMs approach, where for classification task their capabilities are comparable.

I. INTRODUCTION

Unprecedented high volumes of data and information are available in the smart grid context, with the upward growth of the smart metering infrastructure [1], [2]. This recently developed network enables two-way communication between smart grid and individual energy consumers (i.e., the customers), with emerging needs to monitor, predict, schedule, learn and make decisions regarding local energy consumption and production, all in real-time. On the one hand, the grid mesh infrastructure could be seen as a collaborative communication network from which are expected benefits towards better planning and operation of the smart grid, helping the customers transition from a passive to an active role. On the other hand, previous studies have shown that customers naturally adopt energy conserving behaviors when presented with a breakdown of their energy usage [3]–[5]. Concomitantly, an ongoing research thread focuses on a new automatic and scalable solution [6] required to extract useful patterns from energy metered data [7], which constitutes a big data problem.

One possible way to detect building energy flexibility in real-time is by performing energy disaggregation. Disaggregation refers to the extraction of appliance level energy signals from an aggregated energy consumption signal, e.g. the whole-

building. Often only this aggregated signal is made available via the smart meter infrastructure to the grid operator, due to privacy concerns of the end-user. This new approach should open new paths towards better planning and operation of the smart grid, helping the customers transition to active roles. In addition, informing the end-user in real-time, or near real-time, about how much energy is used by each appliance can be a first step in voluntarily decreasing the overall energy consumption.

The energy disaggregation problem, also known as the Non-Intrusive Load Monitoring (NILM) problem, is an important task, introduced by W. Hart [8] in the early 1980s. Traditional approaches for NILM start by investigating if the device is turned on/off [9], and followed by many steady-state methods [10] and transient-state methods [10] aiming to identify more complex appliance patterns. At the same time, advance building energy management systems are looking beyond quantification of the energy consumption by considering fusion information [11] including for instance acoustic sensors to identify the operational state of the appliances [12], motion sensors, the frequency of the appliance used [13], as well as time and appliance usage duration [13], [14]. A more comprehensive discussion about these can be found in recent reviews, such as [15]–[17]. Moreover, new data analytic challenges arise in the context of an increasing number of smart meters, and consequently, a big volume of data, which highlights the need of more complex methods to analyze and take benefit of the fusion information [18]. More recent researches have explored a wide range of different machine learnings methods, using both supervised and unsupervised learning, such as sparse coding [14], clustering [19], [20] or different graphical models (e.g. Factorial Hidden Markov models (FHMM) [13], Factorial Hidden Semi-Markov Model (FHSM) [13], Conditional FHMM [13], Conditional Factorial Hidden Semi-Markov Model (CFHSM) [13], additive FHMM [21] or Bayesian Nonparametric Hidden Semi-Markov Models [22]) to perform energy disaggregation. Recently, in [23] it was shown that by infusing data mining techniques, such as Support Vector Machine or AdaBoost, the accuracy of flexibility detection can be improved significantly. Still, there is an evident challenge to develop an accurate solution that could perform well for every type of appliance.

The energy demand is dependent on the complexity of the buildings energy producing and consuming technologies and the uncertainty in the influencing factors, resulting in frequent fluctuations. The Internet of Things (IoT) [24] infrastructure

used at the neighborhood level contain useful information with high level of variation, acting as an aggregator in the smart grid. These fluctuations are due to the building architecture and thermal properties of the physical materials used, price, the occupants and their behavior, climate conditions, and sub-level system components. Inspired by the successful application of Deep Learning (DL) methods to the building energy prediction [25], [26] problem, in this paper, our proposed approach goes beyond the state-of-the-art in the ongoing research on demand flexibility methods. Firstly, two DL algorithms are investigated in order to perform the energy disaggregation, namely Factored Four-Way Conditional Restricted Boltzmann Machines (FFW-CRBM) [27] and Disjunctive Factored Four-Way Conditional Restricted Boltzmann Machines (DFFW-CRBM) [28]. Secondly, as a main theoretical contribution, a general framework is developed in order to incorporate in one unified architecture each of these algorithms, making them suitable for performing energy disaggregation, flexibility identification and flexibility prediction simultaneously. Thirdly, the proposed methods are validated using a real-world database, with a fine granularity (e.g. 3 seconds), specially constructed to assess the flexibility identification and prediction problems.

The remaining of this paper is organized as follows. Section II introduces the problem description. Section III describes our proposed approach for the energy disaggregation problem. In Section IV the experimental validation of the proposed methods is detailed and Section V concludes the paper.

II. PROBLEM FORMULATION

This section details the problem definition targeted in this paper. In one unified framework, we divide the problem in three subproblems, namely energy disaggregation, flexibility identification and flexibility prediction. Formally, let us consider $\mathcal{B} = \{\mathbf{B}^{(t)}\}_{t=1}^n$ to be the dataset of all the buildings energy consumption, where n is the number of buildings and $\mathbf{D}^{(t)} \in \mathbb{R}^{d \times n}$ to represent a $d \times (N - 1)$ dimensional input sequence, where d represents the number of electrical devices and appliances considered for each time frame (e.g. refrigerator, dishwasher, electric heater), and N is the number of history frames considered in a temporal window.

1) *Energy disaggregation*: **Given** a set of observation $\mathbf{D}^{(t)} \in \mathbb{R}^{d \times n}$ learn a model for every electrical device, \hat{d} .

2) *Flexibility identification*: **Given** the set of building demand energy profiles $\mathbf{B}^{(t)}$ and their corresponding sum of disaggregated electrical parts $\sum_{i=1}^d \hat{d}$ classified at every moment in time find how many devices are operating in the building.

3) *Flexibility prediction*: **Given** the set of building demand energy profiles, $\mathbf{B}^{(t)}$ learn the *time-of-use* (ToU) predictive function (or the power consumption) for every device such that the empirical loss is minimized,

$$\min \|ToU_{\hat{d}}(d|\hat{d}, \mathbf{B}) \wedge ToU_{empirical}(d|\mathbf{D})\| \quad (1)$$

III. PROPOSED METHODS

Recently, it has been proven that it is possible in an unified framework to perform both, classification and prediction, by

using deep learning techniques, such as in [27]–[29]. Consequently, in the context of flexibility detection and prediction, we are exploring the generalization capabilities of Factored Four-Way Conditional Restricted Boltzmann Machines (FFW-CRBM) [27] and Disjunctive Factored Four-Way Conditional Restricted Boltzmann Machines (DFFW-CRBM) [28]. Both models, FFW-CRBM and DFFW-CRBM, have shown to be successful on outperforming state-of-the-art techniques in both, classification (e.g. Support Vector Machines) and prediction (e.g. Conditional Restricted Boltzmann Machines), on time series classification and prediction in the context of human activity recognition, 3D trajectories estimation and so on. Below we are making a brief description of these models, while for a comprehensive discussion on their mathematical details the interested reader is referred to [27], [28].

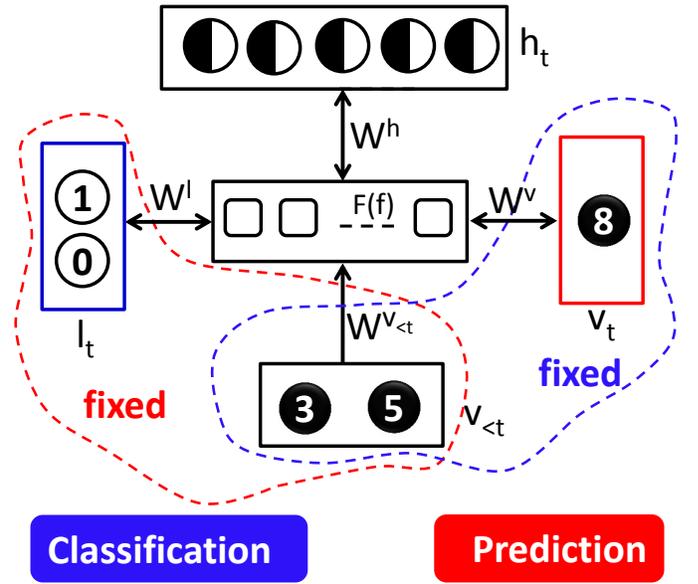


Fig. 1. Classification and prediction schemes for FFW-CRBM. To perform classification the value of each neuron from the dotted blue area has to be fixed (i.e. present and history layers) and to let the model to infer the values of the label neurons. To perform prediction the value of each neuron from the dotted red area has to be fixed (i.e. label and history layers) and to let the model to infer the values of the present neurons.

A. FFW-CRBM and DFFW-CRBM description

FFW-CRBM is an extension of standard Restricted Boltzmann Machines (RBMs) [30], [31] (the basic block of deep learning models [32]) and consists of four layers of neurons, the history layer $\mathbf{v}_{<t}$ which contains the last time series values, the present layer \mathbf{v}_t encoding the current time step (prediction), the hidden layer \mathbf{h}_t which assures that the model is complex enough to model the time series complexity, and the label layer \mathbf{l}_t which incorporates the various type of time series (classification). These layers are connected between them with a fourth order tensor, $\mathbf{W}_{ijkl} \in \mathbb{R}^{n_v \times n_h \times n_{v_{<t}} \times n_l}$, where n_v , n_h , $n_{v_{<t}}$, n_l represent the number of neurons from the present, hidden, history and label layers, respectively. Further on,

W_{ijk_o} is factorized to decrease the computational complexity to $\mathcal{O}(n^2)$, as follows:

$$W_{ijk_o} = \sum_{f=1}^{n_F} W_{if}^v W_{jf}^h W_{kf}^{v<t} W_{of}^l \quad (2)$$

where n_F represents the number of factors in the model. Overall, the model is defined by an energy function:

$$\begin{aligned} \mathbf{E}(\mathbf{v}_t, \mathbf{h}_t, \mathbf{l}_t | \mathbf{v}_{<t}, \Theta) = & \quad (3) \\ & - \sum_{i=1}^{n_v} \frac{(v_{i,t} - a_i)^2}{\sigma_i^2} - \sum_{j=1}^{n_h} h_{j,t} b_j - \sum_{o=1}^{n_l} l_{o,t} c_o \\ & - \sum_{f=1}^{n_F} \sum_{i=1}^{n_v} W_{if}^v \frac{v_{i,t}}{\sigma_i} - \sum_{j=1}^{n_h} W_{jf}^h h_{j,t} - \sum_{k=1}^{n_{v<t}} W_{kf}^{v<t} \frac{v_{k,<t}}{\sigma_k} - \sum_{o=1}^{n_l} W_{of}^l l_{o,t} \end{aligned}$$

The stochastic inference in FFW-CRBM can be done in parallel for the neurons from the same layer as there is no conditional dependence between them. In the case of the label and hidden neurons the inference is done by sampling from a sigmoid function, while for the present layer neurons is done by sampling from a Gaussian distribution. To train such model, Sequential Markov Chain Contrastive Divergence procedure can be used, as described in [27].

To improve the prediction capabilities of FFW-CRBM, in [28] the Disjunctive FFW-CRBM (DFFW-CRBM) was introduced. DFFW-CRBM has a similar architecture with FFW-CRBM with the main difference that it contains two factorized fourth order tensors, one specialized in classification and one in regression (prediction). The inference and training of the model can be done in a similar manner with FFW-CRBM. To clarify, the classification and prediction schemes for FFW-CRBMs¹ are briefly illustrated in Figure 1.

B. Flexibility identification and prediction procedure.

Thus, after the FFW-CRBM or DFFW-CRBM models are trained on data coming from buildings which have smart meters devices at the appliances level, the trained models can be used in real-time on other buildings (which do not have smart meters for appliances) to identify and predict their energy flexibility. Further on, the flexibility information can be used to take decisions in the smart grid or to provide real-time feedback to the buildings. Schematically, this flow of information is depicted in Figure 2 and evaluated in the next Section.

IV. NUMERICAL RESULTS

This section summarizes the experiments performed, flexibility identification (i.e. classification) and flexibility prediction (i.e. prediction), and the main results obtained. Besides that, it includes the dataset characteristics, implementation details of the proposed method and the metrics used to evaluate their performance.

¹Please note that for DFFW-CRBMs the classification and prediction schemes are similar.

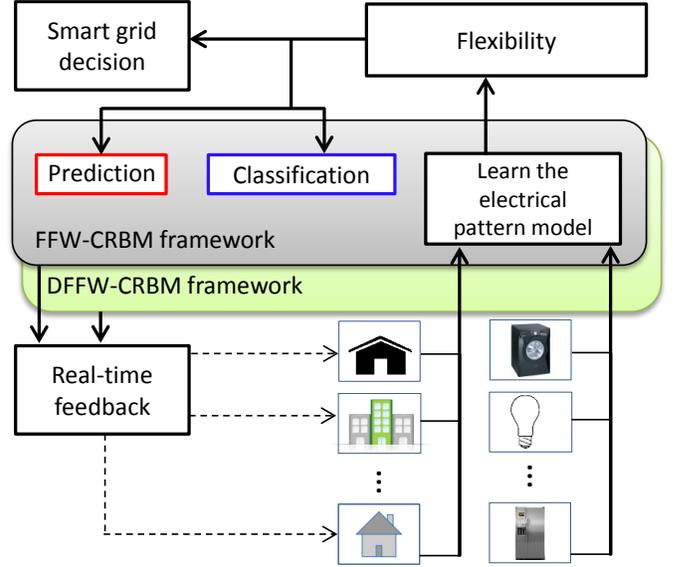


Fig. 2. A general schematic representation of an IoT architecture used for the real-time identification and prediction procedure of the buildings energy flexibility.

A. Dataset characteristics and processing

Thus, to validate our proposed approach we have used a real-world database, namely *The Reference Energy Disaggregation Dataset* (REDD), described by Kolter and Johnson in [33]. This data was chosen as it is an open dataset² collected specifically for evaluating energy disaggregation methods. It contains aggregated data recorded from six buildings over few weeks sampled at 1 second resolution, together with the specific data for all appliances of each building at 3 seconds resolution. In our experiments, we have used the data from the last 5 buildings (i.e. 2, 3, 4, 5, 6) to train the models and the data from a different one (i.e. the first one) to test them.

Similarly with the data processing step from [23] we have applied a median filter of 6 samples to make the power data smoother. At each discrete time step, to make the classification (i.e. detect each appliance activation) or the prediction (i.e. predict the power consumption and the time-of-use of any appliance during any activation) we have used a history window of 10 consecutive time steps of the whole building energy consumption. This procedure leads to 1750614 training data points and 745878 testing data points with an imbalanced number of classes for any device (as the number of activations is smaller than the number of non-active regions). To exemplify the energy disaggregation problem, in Figure 3 we plotted the inflexible and the flexible load from an arbitrary chosen building over one day.

B. Implementation details

The REDD data were processed in the MATLAB[®] environment and then passed on to Python in which FFW-CRBM and DFFW-CRBM were implemented. For each appliance we have

²<http://redd.csail.mit.edu/>, Last visit November 5th, 2015

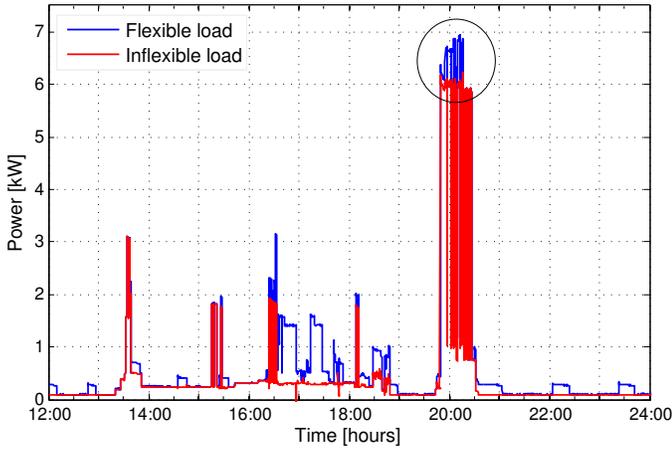


Fig. 3. An example of electrical energy consumption in buildings over one day for inflexible load and flexible load (refrigerator, electric heater, washer dryer and dishwasher).

build one model. In every experiment performed, the history layer had 10 neurons representing the window of 10 values for the building energy consumption, the label layer had 2 binary neurons (i.e. one representing the class of activations, while the other representing the non-active situations), and the present layer had 2 real-valued neurons representing the power consumption and the time-of-use for that specific device in that specific activation period. The choice of the other model parameters was carefully made, as discussed in more details in [28]. Thus, the number of factors was set to 20, the learning rate to 0.001, the weight decay to 0.0002, the momentum to 0.5, and the number of training epochs to 10.

C. Evaluation Metrics

To assess the models performance, three standard metrics were used. Firstly, in order to test the significance of the classification results we used the accuracy metric followed by the balanced accuracy metric. The balanced accuracy metric is imposed by the fact that there is an unbalanced number of data points per class. Secondly, the estimation (prediction) task was evaluated using the Normalized Root Mean Square Error (NRMSE) measure between the real measured flexibility values versus their corresponding predicted values. All the reported performance metrics are measured in percentage.

TABLE I
RESULTS SHOWING ACCURACY [%] AND BALANCED ACCURACY [%] FOR FFW-CRBM AND DFFW-CRBM, WHEN CLASSIFYING AN APPLIANCE VERSUS ALL DATA.

Appliance	Method	Accuracy [%]	Balanced accuracy [%]
refrigerator	FFW-CRBM	86.23	90.05
	DFFW-CRBM	83.10	91.27
dishwasher	FFW-CRBM	97.42	80.21
	DFFW-CRBM	97.26	87.06
washer dryer	FFW-CRBM	98.83	99.03
	DFFW-CRBM	99.06	92.16
electric heater	FFW-CRBM	99.10	90.58
	DFFW-CRBM	99.03	92.05

D. Results

Flexibility identification. In Table I the classification results for all the flexible devices are reported. Both models performed very well (without a clear winner), reaching at least 97.26% accuracy, with the exception of the refrigerator on which the accuracy of DFFW-CRBM is 83.10% and the one of FFW-CRBM is 86.23%. Similarly, the balanced accuracy for both models shows very good performance with a minimum of 80.21% in the case of FFW-CRBM on dishwasher and a maximum of 99.03% of FFW-CRBM on the washer dryer. Overall, these results are comparable with the state-of-the-art results on flexibility identification reported on the same database in [23].

Flexibility prediction. The advantage of using FFW-CRBM and DFFW-CRBM instead of the methods proposed in [23] consists in the fact that beside flexibility identification the same model can perform simultaneously also flexibility prediction. Thus, in Table II we report the prediction performance

TABLE II
RESULTS SHOWING THE NRMSE [%] OBTAINED TO ESTIMATE THE ELECTRICAL DEMAND AND THE TIME-OF-USE FOR FOUR BUILDING ELECTRICAL SUB-SYSTEMS USING FFW-CRBM AND DFFW-CRBM.

Appliance	Method	Power NRMSE [%]	Time-of-use NRMSE [%]
refrigerator	FFW-CRBM	9.36	8.79
	DFFW-CRBM	9.27	8.71
dishwasher	FFW-CRBM	5.49	5.89
	DFFW-CRBM	5.41	5.87
washer dryer	FFW-CRBM	2.70	2.43
	DFFW-CRBM	2.59	2.44
electric heater	FFW-CRBM	1.86	1.78
	DFFW-CRBM	1.85	1.77

on every appliance.

Both models performed very well obtaining a minimum prediction error on the power consumption of 1.85% and a maximum error of 9.36%, while for the time-of-use prediction the minimum error reached was 1.77% in the case of the electric heater and the maximum error obtained was 8.79% for the refrigerator. Even both models show good prediction capabilities, we may observe, same as in [28], that DFFW-CRBM has an easy advantage over FFW-CRBM.

V. CONCLUSION

In this paper, we have proposed a novel IoT framework to perform simultaneously and in real-time flexibility identification and prediction, by making use of Factored Four Way Conditional Restricted Boltzmann Machines and their Disjunctive version. The experimental validation performed on a real-world database (i.e. REDD) shows that both models perform very well, reaching a similar performance with state-of-the-art models on flexibility identification, while having the advantage of being capable to perform also flexibility prediction (i.e. real-time estimation of the power consumption and time-of-use of the flexible appliances). As further research direction, we intend to understand better how the various model parameters (e.g. learning rate, number of factors and

hidden neurons) would affect the performance of the proposed approach.

ACKNOWLEDGMENT

This research has been partly funded by the European Union's Horizon 2020 project INTER-IoT (grant number 687283), and by the NL Enterprise Agency under the TKI SG-BEMS project of Dutch Top Sector.

REFERENCES

- [1] A. Liotta, D. Geelen, G. van Kempen, and F. van Hoogstraten, "A survey on networks for smart-metering systems." *Int. J. Pervasive Computing and Communications*, vol. 8, no. 1, pp. 23–52, 2012.
- [2] D. Geelen, G. van Kempen, F. van Hoogstraten, and A. Liotta, "A wireless mesh communication protocol for smart-metering," in *International Conference on Computing, Networking and Communications, ICNC 2012, Maui, HI, USA, January 30 - February 2, 2012*, pp. 343–349.
- [3] S. Darby, "The effectiveness of feedback on energy consumption: a review for defra of the literature on metering, billing and direct displays." Environmental Change Institute, University of Oxford, Tech. Rep., 2006.
- [4] H.-M. Chen, C.-W. Lin, S.-H. Hsieh, H.-F. Chao, C.-S. Chen, R.-S. Shiu, S.-R. Ye, and Y.-C. Deng, "Persuasive feedback model for inducing energy conservation behaviors of building users based on interaction with a virtual object," *Energy and Buildings*, vol. 45, pp. 106 – 115, 2012.
- [5] S. Darby, "Smart metering: what potential for householder engagement?" *Building Research & Information*, vol. 38, no. 5, pp. 442–457, 2010.
- [6] D. C. Mocanu, "On the synergy of network science and artificial intelligence," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, USA, July 9-15, 2016*.
- [7] E. Mocanu, P. H. Nguyen, W. L. Kling, and M. Gibescu, "Unsupervised energy prediction in a smart grid context using reinforcement cross-building transfer learning," *Energy and Buildings*, vol. 116, pp. 646 – 655, 2016.
- [8] G. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec 1992.
- [9] F. Sultanem, "Using appliance signatures for monitoring residential loads at meter panel level," *IEEE Transactions on Power Delivery*, vol. 6, no. 4, pp. 1380–1385, Oct 1991.
- [10] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *IEEE Power and Energy Magazine*, vol. 1, no. 2, pp. 56–63, Mar 2003.
- [11] G. Fortino, A. Guerrieri, G. M. P. O'Hare, and A. Ruzzelli, "A flexible building management framework based on wireless sensor and actuator networks," *J. Netw. Comput. Appl.*, vol. 35, no. 6, pp. 1934–1952, Nov. 2012.
- [12] M. A. Guvensan, Z. C. Taysi, and T. Melodia, "Energy monitoring in residential spaces with audio sensor nodes: Tinyears," *Ad Hoc Networks*, vol. 11, no. 5, pp. 1539 – 1555, 2013.
- [13] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *SIAM International Conference on Data Mining*, 2011, pp. 747–758.
- [14] J. Z. Kolter, S. Batra, and A. Y. Ng, "Energy disaggregation via discriminative sparse coding," pp. 1153–1161, 2010.
- [15] Y. Du, L. Du, B. Lu, R. Harley, and T. Habetler, "A review of identification and monitoring methods for electric loads in commercial and residential buildings," in *IEEE Energy Conversion Congress and Exposition*, Sept 2010, pp. 4527–4533.
- [16] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, February 2011.
- [17] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, p. 16838, 2012.
- [18] J. Kelly and W. Knottenbelt, "Metadata for energy disaggregation," in *IEEE 38th International Computer Software and Applications Conference Workshops*, July 2014, pp. 578–583.
- [19] D. Bergman, D. Jin, J. Juen, N. Tanaka, C. Gunter, and A. Wright, "Distributed non-intrusive load monitoring," in *IEEE PES Innovative Smart Grid Technologies*, Jan 2011, pp. 1–8.
- [20] A. Iwayemi and C. Zhou, "Leveraging smart meters for residential energy disaggregation," in *IEEE PES General Meeting — Conference Exposition*, July 2014, pp. 1–5.
- [21] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," *Journal of Machine Learning Research - Workshop and Conference Proceedings*, vol. 22, pp. 1472–1482, 2012.
- [22] M. J. Johnson and A. S. Willsky, "Bayesian nonparametric hidden semi-markov models," *Journal of Machine Learning Research*, vol. 14, no. 1, pp. 673–701, Feb. 2013.
- [23] E. Mocanu, P. H. Nguyen, and M. Gibescu, "Energy disaggregation for real-time building flexibility detection," in *IEEE Power and Energy Society General Meeting (PES GM)*, July 2016.
- [24] A. Liotta, "The cognitive net is coming," *IEEE Spectrum*, vol. 50, no. 8, pp. 26–31, August 2013.
- [25] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, "Deep learning for estimating building energy consumption," *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91 – 99, 2016.
- [26] E. Mocanu, E. M. Larsen, P. H. Nguyen, P. Pinson, and M. Gibescu, "Demand forecasting at low aggregation levels using factored conditional restricted boltzmann machine," in *IEEE Power Systems Computation Conference (PSCC)*, June 2016.
- [27] D. C. Mocanu, H. B. Ammar, D. Lowet, K. Driessens, A. Liotta, G. Weiss, and K. Tuyls, "Factored four way conditional restricted boltzmann machines for activity recognition," *Pattern Recognition Letters*, vol. 66, pp. 100 – 108, 2015.
- [28] D. C. Mocanu, H. B. Ammar, L. Puig, E. Eaton, and A. Liotta, "Estimating 3d trajectories from 2d projections via disjunctive factored four-way conditional restricted boltzmann machines," *CoRR*, 2016.
- [29] E. Mocanu, D. C. Mocanu, H. B. Ammar, Z. Zivkovic, A. Liotta, and E. Smirnov, "Inexpensive user tracking using boltzmann machines," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2014, pp. 1–6.
- [30] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," in *Parallel Distributed Processing: Volume 1: Foundations*, 1987.
- [31] D. C. Mocanu, E. Mocanu, P. H. Nguyen, M. Gibescu, and A. Liotta, "A topological insight into restricted boltzmann machines," *Machine Learning*, 2016.
- [32] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1–127, 2009, also published as a book. Now Publishers, 2009.
- [33] J. Z. Kolter and M. J. Johnson, "REDD: A Public Data Set for Energy Disaggregation Research," 2011.