Programme Area: Smart Systems and Heat

Project: WP1 Appliance Disaggregation

Title: Dynamic Modelling

Abstract:

In this deliverable Bournemouth Uni presents an original approach for mining utility data usage patterns relying on a novel Deep Hierarchical Dynamic model which consists of three modules, a Deep Belief network (DBN), a hierarchical mixture model which is based on Latent Dirichlet Allocation (LDA) and a Dynamic Bayesian Network based on Hidden Markov Model (HMM), called DBN-LDA-HMM. This architecture aims at extracting topics from data while taking into account the temporal structure of the data to model the inter-topic sequential dependency. While the mathematical details of the proposed algorithm are described elsewhere, a full empirical evaluation of this pattern mining algorithm using the ETI data is discussed, highlighting its performance on various mining tasks.

Context:

The High Frequency Appliance Disaggregation Analysis (HFADA) project builds upon work undertaken in the Smart Systems and Heat (SSH) programme delivered by the Energy Systems Catapult for the ETI, to refine intelligence and gain detailed smart home energy data. The project analysed in depth data from five homes that trialed the SSH programme's Home Energy Management System (HEMS) to identify which appliances are present within a building and when they are in operation. The main goal of the HFADA project was to detect human behaviour patterns in order to forecast the home energy needs of people in the future. In particular the project delivered a detailed set of data mining algorithms to help identify patterns of building occupancy and energy use within domestic homes from water, gas and electricity data.

Project: HFADA

HIGH FREQUENCY APPLIANCE DISAGGREGATION ANALYSIS

Dynamic Modelling

Contents

1. History

Date	Issue	Details of Change
	Version 0.0	Initial Version. Authors: Dr Saad Mohamad Professor Hamid Bouchachia

2. Documents Referenced

Ref	Document	Title
1	Word document that describes the HEMS data.	Data collection and data format – ELECTRIC, WATER and HEMS-V1 MONITORING
2	Word document that describes the HEMS V1 Mongo data base structure.	HEMS V1 Mongo Data Base Structure
3	Deliverable 1	HFADA_Deliverable_Ver2
4	Deliverable 2	HFADA_Deliverable2_V0
5	Paper	Deep Online Hierarchical Unsupervised Learning for Pattern Mining from Utility Usage Data
6	Paper	Deep Online Hierarchical Dynamic Unsupervised Learning for Pattern Mining from Utility Usage Data

3. Glossary of Terms

Ref	Description
ETI	Energy Technologies Institute
HEMS	Home Energy Management System (also referred to as HEMS V1)
HFAD	High Frequency Appliance Detection
LDA	Latent Dirichlet Allocation
GLDA	Gaussian Latent Dirichlet Allocation
DBN	Deep Belief network
HMM	Hidden Markov Model
DBN-LDA-HMM	Deep Belief Network-Latent Dirichlet Allocation – Hidden Markov Model (Deep Hierarchical-Dynamic model)
NILM	Non-intrusive load monitoring

4. Executive Summary

This deliverable describes the third task which is related to dynamic modelling. Specifically, it presents an original approach for mining utility data usage patterns relying on a novel Deep Hierarchical Dynamic model which consists of three modules, a Deep Belief network (DBN), a hierarchical mixture model which is based on Latent Dirichlet Allocation (LDA) and a Dynamic Bayesian Network based on Hidden Markov Model (HMM), called DBN-LDA-HMM. This architecture aims at extracting topics from data while taking into account the temporal structure of the data to model the inter-topic sequential dependency. While the mathematical details of the proposed algorithm are described elsewhere[1], a full empirical evaluation of this pattern mining algorithm using the ETI data is discussed, highlighting its performance on various mining tasks.

5. Introduction

In this report, we propose a fully unsupervised novel non-intrusive load monitoring (NILM) solution that combines a dynamic Bayesian hierarchical mixture model and a deep belief network (DBN). The processing flow in this s architecture consists of two stages. First, DBN learns, in unsupervised fashion, low-level generic features from the raw signals of the house utilities usage. Then, the hierarchical Bayesian model (Latent Dirichlet Allocation) learns high-level features that capture correlations among the low-level ones; whereas the temporal ordering of the high-level features is captured by the Dynamic Bayesian Model (Hidden Markov Model). Thus, in contrast to Deliverable 2, the proposed solution harnesses the ability of DBN to learn distributed hierarchies of features in order to construct sophisticated appliances-specific features without the need to rely on precise human-crafted input representations. On the other hand, the clustering capability of the hierarchical Bayesian models (LDA) helps summarise the input data by extracting higher-level information that represents the residents' consumption patterns (appliance patterns). In the meantime, the dynamic Bayesian network correlates these patterns over time. Hence, the proposed architecture models the temporal ordering in the data extending the approach of online Gaussian latent Dirichlet allocation (GLDA) developed in Deliverable 2.

Using this DBN-LDA-HMM architecture, we aim at overcoming the computational complexity that would occur if temporal modelling was directly applied to the raw data or even to the manually engineered and constructed features. The computational efficiency is crucial as our application involves massive data from different utilities usage. Moreover, we develop a novel online inference algorithm to cope with this big data. Finally, we propose different evaluation methods to analyse the results showing that DBN-LDA-HMM finds useful patterns and improves on the results of Deliverable 2.

6. Dynamic Modelling

In the following we introduce the DBN-LDA-HMM approach after providing some background related to human activity recognition and NILM techniques.

[1] This work was submitted to *IEEE transactions on pattern analysis and machine intelligence*.

6.1 Context and Motivations

This work is a continuation of the work in Deliverable 2 [1] where online Gaussian Latent Dirichlet Allocation (GLDA) is proposed to extract global components that summarise the energy signal. These components provide a representation of the consumption patterns. The algorithm is applied on the same data-set as in this report. However, in contrast to [1] temporal dependency is considered here. We also employ deep learning to construct features rather than engineering them using signal processing technique.

Recently, the field of deep learning (DL) has made a huge impact and achieved remarkable results in computer vision, natural language processing, and speech recognition. Yet it has not been exploited in the field of NILM. DL provides an effective tool for extracting multiple layers of distributed features representations from high-dimensional data. Each layer of the deep architecture performs a non-linear transformation of the outputs stemming from the previous layer. Thus, through DL, the data is represented in the form of a hierarchy of features, from low-level to high-level [2, 3] Instead of relying on heuristic hand-crafted features, DL learns to extract features that allow for more discriminative power. Supported by the sheer size of the available data and its high sampling rate (205 KHZ) which results in a very high-dimensional data, we are the first to use unsupervised DL model in NILM. In contrast to existing electrical engineering and signal processing approaches adopted in NILM, ours relies fully on the data to construct informative features.

In this deliverable, we pre-train a DBN [4] to learn generic features from unlabelled raw electrical signal with 1 second granularity. The extracted features are fed to the LDA-like part of the model with 30 minutes granularity to build clusters which correspond to appliance usages. Ideally, each cluster refers to an appliance, but that's a very demanding task since, we rely only on the utility data. Moreover, although, the bag-of-words assumption adopted here is a major simplification, it breaks down unnecessary low-level hard-to-model complexity leading to computationally efficient inference with no much loss as shown in GLDA [1] and [5]. The second aim of this study to capture the temporal dynamics hidden in the data in order to understand the relationship between the clusters/topics/components (implicitly representing appliances usage). By doing so, sequences of appliance usages are captured, which could hopefully and ideally correspond, in turn, to the daily human activity. This modelling is obtained through the application of a dynamic Bayesian network (HMM layer) in the proposed model.

6.2 The DBN-LDA-HMM Model

The proposed layered three-module architecture is motivated by the ultimate goal of capturing an abstraction of human activities. As explained earlier, each module in this architecture has a well-defined purpose: deep network to extract high-level granular features, LDA to generate topics (patterns corresponding to appliance usage), and HMM to model the temporal dependencies between the topics.

Such multi-module design has been inspired from existing work [6,7], though not in the same context as ours. In particular, authors in [6] plugged a hierarchical Dirichlet process (HDP) prior on top of a Deep Boltzmann Machine (DBM) network which allows learning multiple layers of abstractions. The low-level abstraction represents generic domain-specific features that are hierarchically clustered and shared to yield high-level abstraction representing patterns. However, this work does not consider the temporal ordering of the high-level representations (patterns). On the other hand, the study in [7] proposed an LDA-HMM hybrid model to perform action recognition. The model was motivated by the success and the efficiency of the bag-of-features approach, adopted by topic modelling, in solving general high-level problems. The temporal ordering power of HMM is harnessed to correlate the activity at high-level. The paper uses collapsed Gibbs sampler for approximate inference and learning.

In our work, we use an unsupervised version of LDA-HMM [7] which is more similar to the approach taken in [8]. But on the contrary, instead of using Gibbs sampling, we propose a stochastic variational inference (SVI) [9] algorithm that allows to do large-scale inference in order to cope with the massive amount of energy consumption data (around 8 TB).

Moreover, we employ DBN to construct appliance-specific features which are used as input to the hierarchical Bayesian mixture model to construct topic-specific features. The mixture of these components forms the residents' energy consumption patterns. The dynamic part of the three-module architecture (HMM) exploits the temporal regularity in the human behaviour leading to better performance and allowing forecasting energy demand.

In this work, we demonstrate that this approach can capture significant statistical structure in a specified window of data over a period of time. This structure provides understanding of regular patterns in the human behaviour that can be harnessed to provide various services including energy efficiency. For example, understanding of the usage and energy consumption patterns of residents to help them acquire insight into their usage of utilities and provide them with recommendation concerning their lifestyle to improve their consumption behaviour [10]. Also, but not high relevance to this work, energy consumption patterns could be used to predict the power demand (load forecasting), to apply management policies, and to avoid overloading the energy network.

As already mentioned, the DBN-LDA-HMM algorithm is going to be trained on a very huge amount of data resulting from the high sampling rate around 205 kHz of the electricity signal which gives us an advantage compared to the data used in other research studies except for [11]–[13]. Besides the advantage the data availability, its diversity (energy, water and gas usage data) and high rate go beyond what similar studies use [14,15]. Moreover, measurements provided by additional sensors are also exploited to refine the performance of the pattern recognition algorithm. More details on the data can be found in [1] and [16].

The mathematical formulation of the proposed DBN-LDA-HMM algorithm can be found in [16].

7 Experiments

In this section, we will first introduce the experimental setting of DBN-LDA-HMM. Then, data pre-processing is described followed by results and discussion.

7.1 Experimental Settings

Initially, the utility usage data is pre-processed in 3 steps: 1) synchronisation of same utility data, 2) alignment of data coming from different utilities. and 3) feature extraction. Details about the pre-processing steps and data description are given in [5] and Deliverable 1. In this section, we focus on the experiments performed on the pre-processed data, where the online LDA-HMM is applied on the features extracted by DBN. DBN-LDA-HMM comes with a number of parameters, we refer the reader to [16] (Section 4) for an exhaustive coverage of the experimental setting of such parameters which concern: the granularity of the processing windows of the signals, the feature windows after DBN, the number components/clusters, the learning rate, the number of iterations and the hyper-parameters of the three modules.

7.2 Global Components

The learned components by DBN-LDA-HMM correspond to clusters of data produced by DBN (output space of DBN). They represent patterns of energy consumption that underlie the human activities. Such clusters are represented by multinomial distribution over the discrete features obtained by means of DBN. A pattern in this context is considered as a mixture of the clusters (each cluster contribute proportionally to the emergence of a pattern). For the sake of visualisation, we plot gray-scale images of these components where black colour indicates zeros-probability and white colour indicates one-probability. Figure 1 shows the aforementioned clusters where x-axis corresponds to different clusters and y-axis represents the discrete DBN's outputs (input in the new feature space generated by DBN).

It can be seen from Fig.1 that only around 350 dimensions of DBN's output represent the appliance related components that summarise the data. Hence, different combination in these 350 dimensions form the clusters whose mixtures represents patterns. In analogy to topics model, these are the words composing the topics forming the documents. This observation shows that DBN has managed to reduce the high-dimensional input space (raw signal) to discrete lower dimensional output space (See Appendix in [5]) where countably small number of points represents most of the input signal over 1 second granularity. Hence, we expect

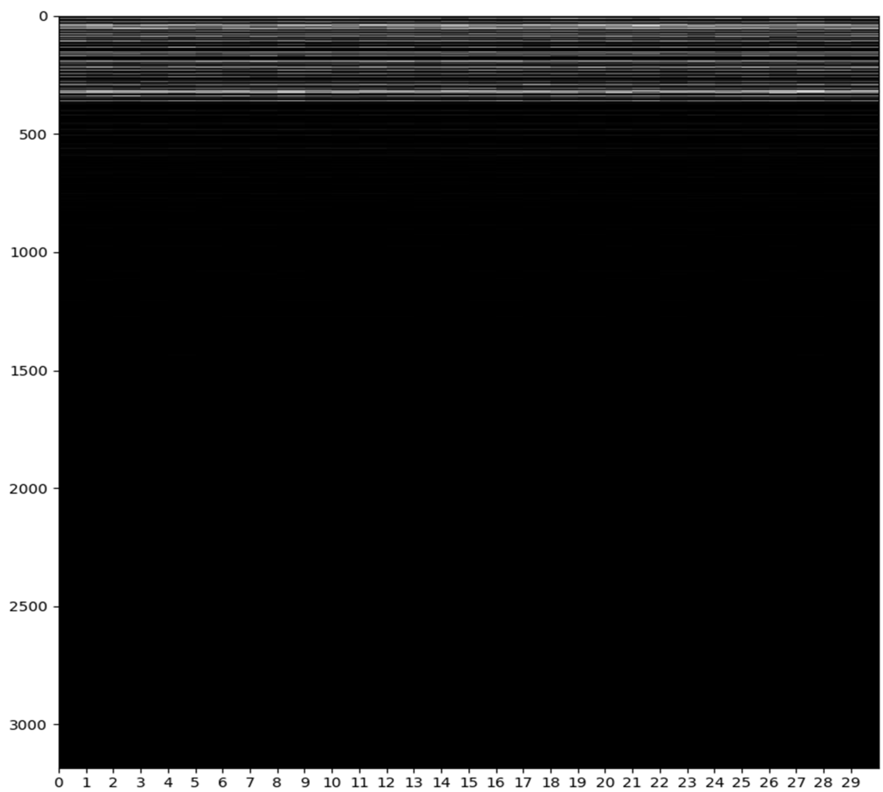

Figure 1: Clusters in DBN’s output space

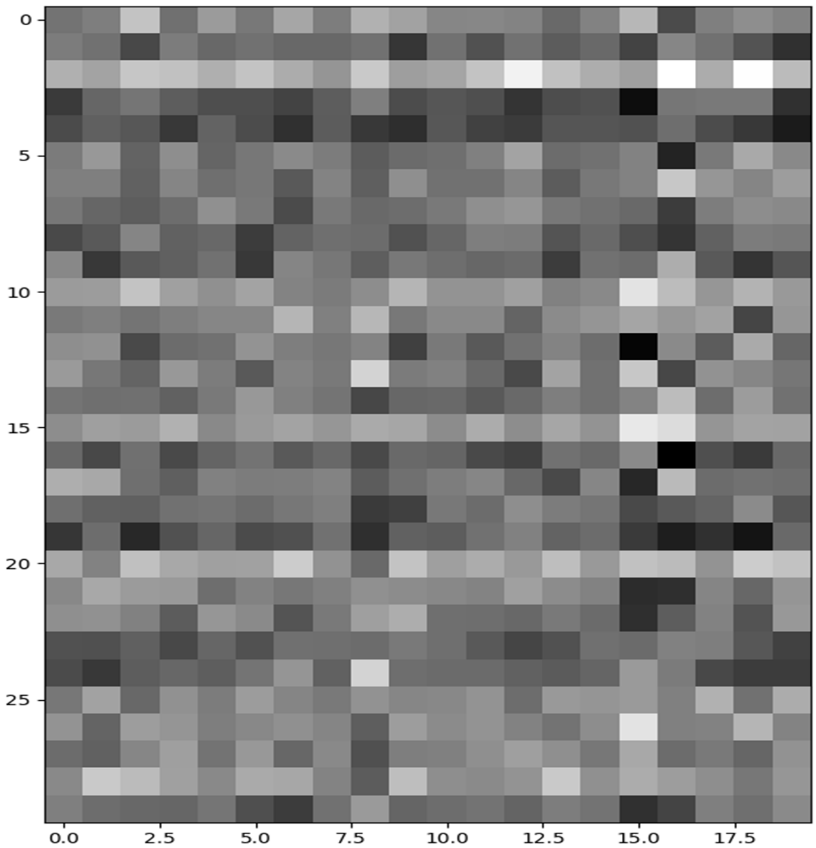

Figure 2: Patterns of energy consumption activities

these points to have strong relation with appliances usage. This is supported by the few light-coloured points appearing with each component which indicate that different clusters (appliance-related components) are mainly composed of these points.

To highlight the emergence of patters, Figure 2 shows the relationship between clusters (y-axis) and their proportion. The sparse light-coloured rectangles in Fig. 2 indicate that the majority of patterns consist mainly of few appliance-related components. For example, the main components for breakfast patterns will relate to cooking and heating appliances such as the hob and the oven. Figure 2 also shows few horizontal strips of light colour located at component index 2, 10, 15 and 20. This can be explained by the fact that these components appear in most patterns meaning that they may belong to appliances which are used in different activities (e.g. lighting - lamps).

In the following, we propose two evaluation methods to support our claims about the relation between clusters and appliances, patterns and activities.

7.3 Evaluation and Analysis

In order to investigate the quality of the results, we study the regularity of the mined patterns by matching them across similar periods of time. For instance, it is expected that similar patterns will emerge in specific time slots like breakfast in every morning, watching TV in the evening, etc. This regularity can also be seen across days, for instance, consumption behaviour during working days is different from that during the weekend. Hence, it is interesting to understand how such patterns emerge as regular events. We also provide a quantitative evaluation of the algorithm by proposing a mapping method that reveals the specific energy consumed from the inferred patterns within the patterns' granularity (fixed to 30 minutes). By doing so, we can evaluate numerically the consistency between energy consumption and the extracted patterns. This is achieved by fitting a regression model to the energy consumption over the K components (clusters in DBN's output space). This technique will also allow numerically checking the predicted consumption against the real consumption.

7.4 Pattern Regularity

Using the optimal parameters' setting as explained in [16], in the following we examine the regularity of the mined patterns. To do that, we use the first two weeks of the data (from 11-05-2017 10:10:10 to 25-05-2017 10:10:10) for analysis. To study the regularity of the energy consumption behaviour of the residents, we compare the mined patterns across different days of the testing period. The similarity of the patterns across the two weeks are computed (details can be found in [16]).

Table 1 shows the per-day dissimilarity. It can be clearly seen from the table that there are regular patterns across the same days from two different weeks. That is, similar energy consumption patterns appear across these days. This dissimilarity is a bit higher for the weekend during which more irregular activities could take place. Computing the dissimilarity measure between week and weekend days confirms this observation. For instance, the dissimilarity between first week's Monday and second week's Sunday is equal to 0.0091 which is much higher than that between Mondays of the two weeks. In contrast, the dissimilarity among working days is generally low.

Table 1: Patterns dissimilarity matrix

Week 1 / Week 2	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Mon	0.0052	0.0058	0.0068	0.0061	0.0066	0.0090	0.0093
Tue	0.0048	0.0049	0.0057	0.0058	0.0069	0.0080	0.0095
Wed	0.0064	0.0055	0.0049	0.0048	0.0071	0.0083	0.0089
Thu	0.0069	0.0060	0.0059	0.0058	0.0074	0.0085	0.0090
Fri	0.0070	0.0064	0.0069	0.0073	0.0068	0.0080	0.0097
Sat	0.0089	0.0086	0.0081	0.0086	0.0079	0.0082	0.0080
Sun	0.0091	0.0087	0.0091	0.0093	0.0096	0.0080	0.0088

This regularity may be caused by regular user lifestyle leading to similar energy consumption behaviour within and across the weeks. Such regularity is violated in the weekend time, as more irregular activities could take place. Having shown that there is some regularity in the mined patterns, it is more likely that specific energy consumption can be associated with each component.

For the sake of completeness, in the next section, we apply a regression method to map the patterns within the patterns' granularity (fixed to 30 minutes) to energy consumption. Thus, the parameters of interest are the energy consumption associated with the components. By attaching energy consumption with each component, we can help validate the coherence of the extracted patterns and evaluate numerically the consistency between energy consumption and the extracted patterns which can be exploited to predict the load demand.

7.5 Energy Mapping

As shown in the previous section, DBN-LDA-HMM can express the energy consumption patterns by mixing multinomial distributions over mixture of components (clusters) that summarise the data. Each component is a distribution over a high-dimensional feature space and understanding what it represents is not easy. Hence, we propose to associate consumption quantities to each component. Such association is motivated by the fact that an energy consumption pattern is normally governed by the usage of different appliances in the house. There should be a strong relation between components and appliances usage. Hence, a relation between components and energy consumption is plausible. Note that the best-case scenario occurs if each component is associated with the usage of a specific appliance (cluster ≈ appliance usage). Apart from the coherence study, associating energy consumption with each component can be used to predict the energy consumption demand.

As explained in [16], we apply a simple least-square regression method to map the inferred patterns within the patterns' granularity (fixed to 30 minutes) to energy consumption. We train the regression model on the week from 18-05-2017 23:45:22 to 25-05-2017 23:45:22 and run the model on the following week from 25-05-2017 23:45:22 to 01-06-2017 23:45:22. Figure 3 and 4 show the energy consumption (in joules) along with the estimated consumption computed using the learned per-component consumption parameters.

The similarity between the estimated and computed energy consumption demonstrates that the inferred components express distinct usages of energy. Such distinction can be the result of the usage of different appliances likely having distinct energy consumption signatures. Thus, the proposed approach produces coherent and regular patterns that reflect the energy consumption behaviour and implicitly human activities. Note that it is possible that different patterns (or appliance usages) may have the same energy consumption and that might be one reason why both estimated and computed energy consumption is not fully the same.

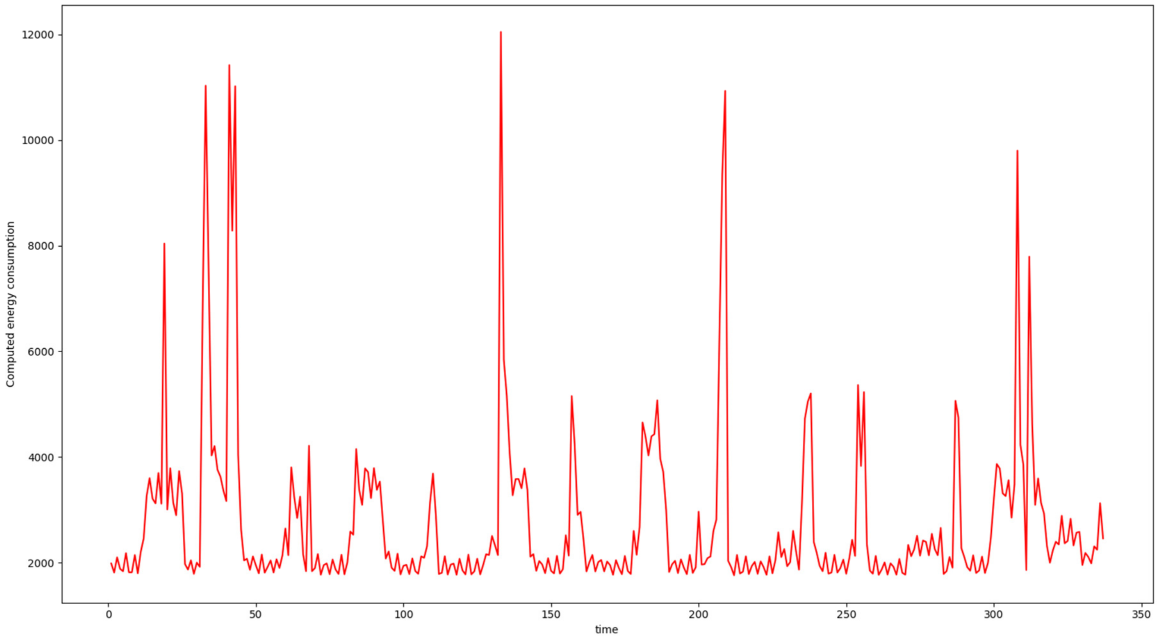

Figure 3: Computed energy consumption

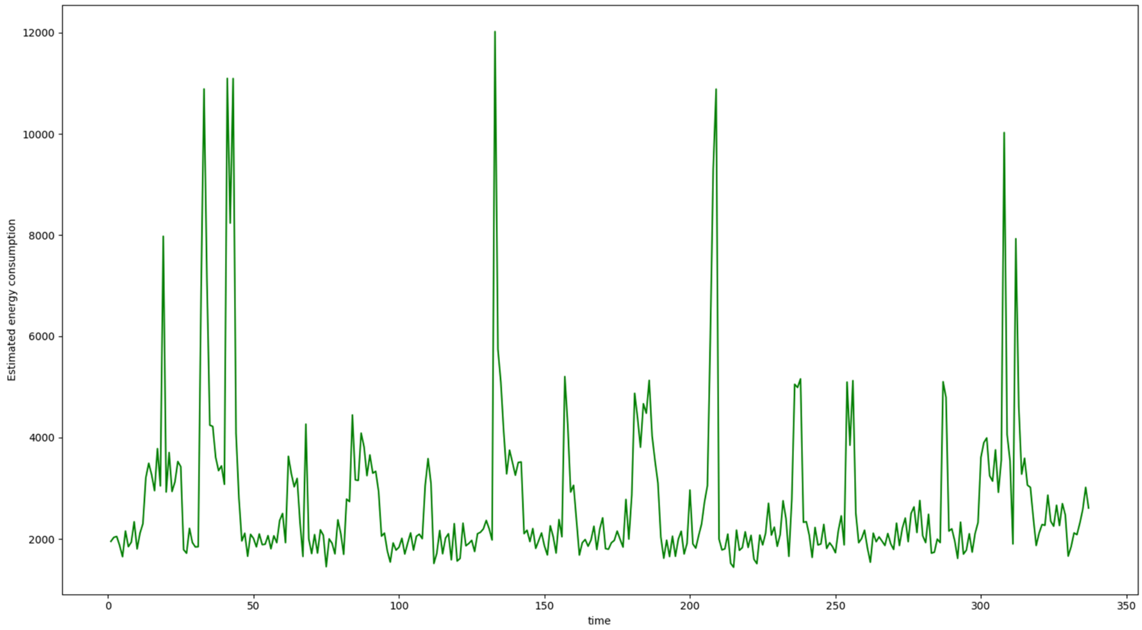

Figure 4: Estimated energy consumption

8. Conclusion

In this report, we presented a novel approach to extract patterns of the users' consumption behaviour from data involving different utilities (e.g., electricity, water and gas) as well as some sensors measurements. DBN-LDA-HMM is fully unsupervised and the LDA-HMM component' training is done online which made it efficient for big data. To analyse the performance, we proposed a two-step evaluation that covers: patterns regularity and coherency. The experiments show that the proposed method is capable of extracting regular and coherent patterns that highlight energy consumption over time.

9. References

1. M. Saad and B. Abdelhamid, “Online Gaussian LDA for unsupervised pattern mining from utility usage data, ”submitted to ECML-PKDD, 2018

2. L. Deng, “A tutorial survey of architectures, algorithms, and applications for deep learning,” APSIPA Transactions on Signal and Information Processing, vol. 3, 2014.

3. Y. Bengio et al., “Learning deep architectures for AI,” Foundations and trends in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.

4. G. E. Hinton, S. Osindero, and Y.-W. Teh, “A fast learning algorithm for deep belief nets,” Neural computation, vol. 18, no. 7, pp. 1527–1554, 2006.

5. M. Saad, A. Damla, M. Chemseddine, and B. Abdelhamid, “Deep online hierarchical unsupervised learning for pattern mining from utility usage data,” submitted to UKCI, 2018.

6. R. Salakhutdinov, J. B. Tenenbaum, and A. Torralba, “Learning with hierarchical-deep models,” IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 8, pp. 1958–1971, 2013.

7. M. R. Malgireddy, I. Nwogu, and V. Govindaraju, “Language motivated approaches to action recognition,” The Journal of Machine Learning Research, vol. 14, no. 1, pp. 2189–2212, 2013.

8. T. Hospedales, S. Gong, and T. Xiang, “A markov clustering topic model for mining behaviour in video,” in Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009, pp. 1165–1172

9. M. D. Hoffman, D. M. Blei, C. Wang, and J. Paisley, “Stochastic variational inference,” The Journal of Machine Learning Research, vol. 14, no. 1, pp. 1303–1347, 2013.

10. C. Fischer, “Feedback on household electricity consumption: a tool for saving energy?” Energy efficiency, vol. 1, no. 1, pp. 79–104, 2008.

11. J. Kelly and W. Knottenbelt, “The uk-dale dataset, domestic appliance-level electricity demand and whole-house demand from five uk homes,” Scientific data, vol. 2, p. 150007, 2015.

12. A. Filip, “Blued: A fully labeled public dataset for event-based non-intrusive load monitoring research,” in 2nd Workshop on Data Mining Applications in Sustainability (SustKDD), 2011, p. 2012.

13. J. Z. Kolter and M. J. Johnson, “Redd: A public data set for energy disaggregation research,” 2011.

14. S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic, “Ampds: A public dataset for load disaggregation and eco-feedback research,” in Electrical Power & Energy Conference (EPEC), 2013 IEEE. IEEE, 2013, pp. 1–6.

15.S. Makonin, B. Ellert, I. V. Bajic, and F. Popowich, “Electricity, water, and natural gas consumption of a residential house in canada from 2012 to 2014,” Scientific data, vol. 3, 2016.

16. M. Saad, A. Damla, M. Chemseddine, and B. Abdelhamid, “Deep Online Hierarchical Dynamic Unsupervised Learning for Pattern Mining from Utility Usage Data,” submitted to IEEE transactions on pattern analysis and machine intelligence, 2018