Energy maps for Small and Medium Enterprises

Danica Vukadinović Greetham
Open University, UK
danica.vukadinovic-greetham@open.ac.uk

Georgios Giasemidis
CountingLab, Ltd., UK
georgios@countinglab.ac.uk

ABSTRACT

Recurrence Quantification Analysis (RQA) is becoming a popular technique to analyse time-series obtained from complex dynamical systems. In this work, a recently presented RQA-based method to analyse and manage energy demand at low sampling rate (5 min and 30 min) is tested using data-sets from three small enterprises. From recurrence plots, different RQA variables are obtained and analysed, following parameter optimisation that depends on a system observed. Based on RQA variables, energy maps of 'normal' behaviour are created. Here, preliminary robustness tests concerning the training phase length, missing data and noise are presented. Our test results show that this approach has great potential in energy management of small and medium enterprises.

CCS CONCEPTS

• **Theory of computation → Pattern matching**.

KEYWORDS

recurrence quantification analysis(RQA), SME energy profiles

ACM Reference Format:
Danica Vukadinović Greetham and Georgios Giasemidis. 2019. Energy maps for Small and Medium Enterprises. In *Proceedings of The Tenth ACM International Conference on Future Energy Systems (e-Energy '19).* ACM, New York, NY, USA, 2 pages. https://doi.org/https://doi.org/10.1145/3307772.3330164

1 INTRODUCTION

In this work, we investigate robustness of the recently proposed Recurrence Quantitive Analysis (RQA) based method (for further details see [2], [1]) for energy management of small and medium enterprises (SMEs). SMEs energy profiles consist of stochastic, complex and non-linear components. Recurrence, a fundamental property of dynamical processes can help analyse it. The method consists of two phases, the *training phase* where the map(s) of 'usual' behaviour is obtained, and the *operational phase* where the new data is tested against the existing map. To investigate robustness, we quantify the impact on the output of missing and noisy data in the input, and training phase length.

2 METHOD

Given a time-series of n observations $X = \{x_1, \ldots, x_n\}$, (e.g. energy usage readings, usually at 5 min or 30 min resolution) the phase space is defined by transforming the readings into time-delayed vectors at each time-step, i.e. $\mathbf{Y} = \{Y_1, \ldots, Y_m\}$, where $Y_i = (x_i, x_{i-\tau}, \ldots, x_{i-(D-1)\tau}) \in \mathbb{R}^D$, τ is the delay and D is the embedding dimension of the phase-space, and $m = n - (D-1)\tau$. The distance matrix $DM(i,j) = ||Y_i - Y_j||_2$, $i,j = 1, \ldots, m$, is the Euclidean distance between vectors Y_i and Y_j in the phase-space. The recurrence plot follows from the distance matrix, defined as

$$R(i,j) = H(\epsilon - DM(i,j)), \tag{1}$$

where $H(x)$ is the Heaviside step function.

The entry $R(i,j)$ equals one and the states Y_i and Y_j are considered recurrent, when the distance between Y_i and Y_j is within an ϵ-radius in phase-space. As a state Y_i in phase-space corresponds to a time-step of the original time-series X, recurrence plots inform us of recurring patterns within our current time-series.

The recurrence plot $R(i,j)$ is created from the input data - energy usage readings over a period of time, and 5 RQA variables are computed ([2] REC, DET, LAM, ENT, TT), using a sliding window over the diagonal of the recurrence plot. The RQA variables are then projected onto 2-d space using Principal Component Analysis, and a training map is created using density of the points in different areas to define its boundary at a given quantile (*sensitivity level*). New data is tested against the map (computing RQA variables and projecting them onto the 2-d space) and the number of values outside the map is recorded as alerts.

$R(i,j)$ depends on three parameters: the time-delay τ, the embedding dimension D and the radius ϵ, to capture the correct dynamics of a system with noise. The parameters need to be optimised according to the system's characteristics and the application. In [3] several potential issues are recognised when choosing parameters. This is especially true with emerging areas of applications, e.g. 30 min resolution electric energy data. The optimal selection of parameters for energy usage data was discussed in [1].

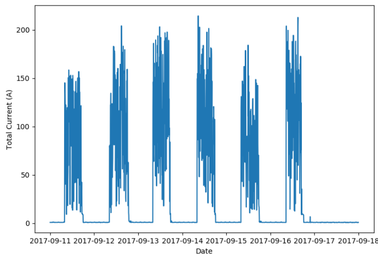
3 DATA

Dry-cleaners. The total demand (measured in total current (A)) in the 5 min resolution from 11/09/17 until 09/10/17 was used for the training phase to create the map (a typical weekly profile is on Fig 1a, incl. Sundays closure) . The following week (from Tuesday 10/10/17 to Monday 16/10/17) was used as the operational phase.

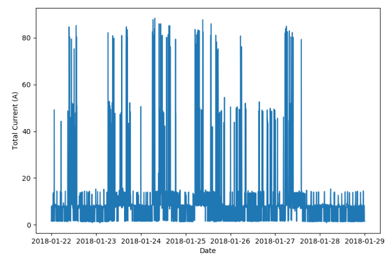
Butchers. Due to constantly switched on fridges, high demand is observed during closure times, so we treat all time-steps as 'operational' from the algorithmic point of view. For the butchers, the total demand (in total current (A)) at the 5 min resolution from 16/01/18 until 09/02/18 was used for the training phase to create the map. The data from Saturday 17/02/18 until Sunday 23/02/18 is used for the operational phase.

Office. The data covers the period from 12th of May 2018 to 2nd of August 2018 and we have "Power" readings (in Watts), which

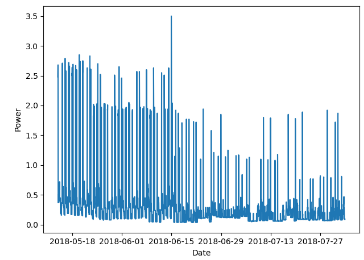
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
e-Energy '19, June 2019, Phoenix, Arizona, US
© 2019 Copyright held by the owner/author(s).
ACM ISBN https://doi.org/10.1145/3307772.3330164.
https://doi.org/https://doi.org/10.1145/3307772.3330164



(a) A weekly profile of the total current from a dry-cleaners business.



(b) A weekly profile of the total current from a butchers business.



(c) Timeseries of power data for the office SME.

Figure 1: Energy/power profiles for three SMEs

are collected via a smart meter. In Figure 1c, the whole time-series for the aforementioned period of available data is given. Regular spikes and a transition to lower consumption at about the sixth week can be observed.

4 RESULTS

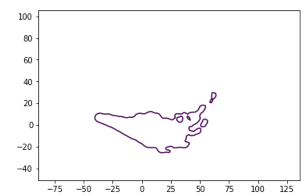
The impact of missing and noisy data and length of training period was tested. For the missing data, we set a percentage $p\%$, where $p = 5, 10, 20, 40$, of data input to zero. For the input data of length l, we calculate $n = \lceil pl \rceil$. Then uniformly at random, a sample of n entries out of l is chosen and those entries are set to zero. For the noisy data, Gaussian noise with mean 0 and standard deviation of $\frac{1}{2}$ std of the original was added to time-series. For the training phase length, percentage of regular points (inside a map) was recorded for varied lengths.

5 CONCLUSIONS

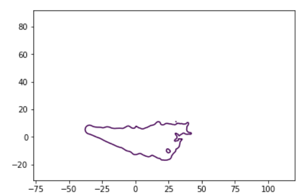
SME's energy management is challenging: the wide range of different appliances; no data available (automatic metre readers instead of smart meters etc.); very low sampling rates (e.g. once per month); lack of engagement from owners, etc. For this reason, RQA based method that uses Principle Component Analysis is helpful. It produces 2-d maps - easily visualising normal behaviour and it is robust enough to be used across different businesses.

We ran several experiments to assess the robustness in missing data, noise, training phase length of the methods developed in [1]. In the majority of the cases, the method is robust to the aforementioned tests. The number of alerts remains within the sensitivity levels. We found that the output does not change drastically with noise, missing data or duration of the training period. When optimal parameters vary with changes in the input, the PCA and maps co-adjust so that the final percentage of alerts is within the expected levels. We also showed, that when there is a change in behaviour, e.g. a transition to a lower demand, this can be identified by the proposed method.

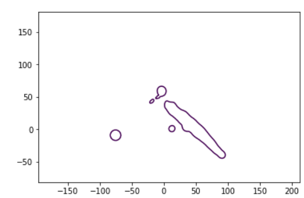
A further testing of the method on similar and different businesses to the ones used here is needed. We also hope to look at seasonal trends and how they influence maps. Finally, we aim to apply these techniques for map creation of individual appliances and their combinations.



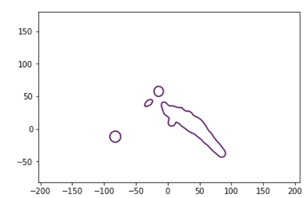
(a) Dry-cleaner



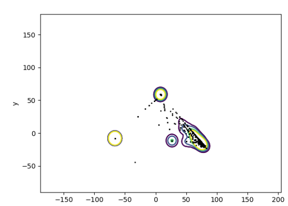
(b) Noisy data - dry-cleaner



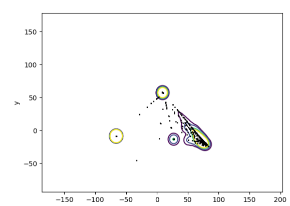
(c) Butcher



(d) Missing 20% data - butcher



(e) Office



(f) Training 5 wks - office

ACKNOWLEDGMENTS

The work was supported by AND Technology&Research through NSDEMIC (UK Government BEIS).

REFERENCES

[1] G. Giasemidis and D. Vukadinović Greetham. 2018. Optimising Parameters inRecurrence Quantification Analysis of Smart Energy Systems. In *The 9th International Conference on Information, Intelligence, Systems and Applications (IISA 2018) (IASC, 2nd Workshop on Intelligent Applications in Smart Cities).*

[2] Laura Hattam and Danica Vukadinović Greetham. 2018. Energy Disaggregation for SMEs Using Recurrence Quantification Analysis. In *Proceedings of the Ninth International Conference on Future Energy Systems (e-Energy '18)*. ACM, New York, NY, USA, 610–617. https://doi.org/10.1145/3208903.3210280

[3] Norbert Marwan. 2011. How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos* 21, 04 (2011), 1003–1017. https://doi.org/10.1142/S0218127411029008 arXiv:https://www.worldscientific.com/doi/pdf/10.1142/S0218127411029008