Greenwood DM, Davison PJ, Duby S, Wade NS.
[Methods and Applications for Electricity Demand Disaggregation in Developing Countries](#).
In: IEEE PES PowerAfrica.
27 June – 1 July 2016, Livingstone, Zambia: IEEE.

Copyright:

DOI link to article:

Date deposited:

05/07/2016

Methods and Applications for Electricity Demand Disaggregation in Developing Countries

David Greenwood, Neal Wade, Peter Davison

School of Electrical and Electronic Engineering
Newcastle University
Newcastle upon Tyne, UK
David.Greenwood@ncl.ac.uk

Sam Duby
SteamaCo Ltd
Kenya
sam@steama.co

Abstract— **Electricity demand disaggregation describes the identification of appliance level loads from aggregate power measurements. This topic is widely researched, but the majority of projects focus on grid connected customers with strong connections and many loads; however, there are numerous applications for disaggregation in the intelligent, off-grid systems emerging in developing countries. In this paper, we describe an established electricity demand disaggregation method, the Hidden Semi-Markov Model (HSMM), present results using this method to dissagregate demand data from a real microgrid in Kenya, and discuss applications for disaggregation in developing countries.**

Index Terms—**Microgrids, Power Demand, Hidden Markov models**

I. INTRODUCTION

In many areas of the world, a significant number of people and communities do not have access to electricity. Connecting these communities to conventional, large scale power networks may not represent the best way to provide swift, affordable, and reliable access to low carbon electricity; smaller, independent micro-grids, based around local – often renewable – generation and energy storage, may be in a better positioned to fulfil these requirements [1, 2]. These systems, however, present new challenges in terms of planning and operation, which have resulted difficulties for many projects seeking to install such a system [3].

The systems in question are usually limited in terms of power, energy or both, and, due to the relatively small number of customers, demand is often more variable and unpredictable than in a conventional large power system. This is compounded by the fact that the renewable energy sources – commonly solar photovoltaic panels – are intermittent and non-dispatchable, and that storage elements, whose performance depends enormously on how they are used, are overburdened [4].

The research presented in this paper has been carried out as part of the ESCoBox project. This is an EPSRC funded project which aims to develop smart microgrids for the developing world, with the aim of improving reliability, efficiency, flexibility, responsiveness, and cost-effectiveness.

The demand disaggregation methods described represent a key input to the project, enabling microgrid operators to understand their demand base, and thus make informed technical and commercial decisions.

Demand disaggregation is a valuable tool within smart micro grids; by understanding which appliances make up an aggregate load, it is possible to make better informed control and policy decisions to ensure system stability and reduce operational costs. It is suggested that demand disaggregation could be the 'holy grail of energy efficiency' [5], and could provide benefits for end users, research and development, policy makers, and utilities.

Demand disaggregation allows engineers and operators to understand demand to a greater extent, and can be applied to improve reliability and reduce energy costs. Some example applications of disaggregation, in the context of off-grid systems, are:

- Identifying problematic usage patterns that could be detrimental to system reliability.
- Understanding the impact of demand on battery state of health in off-grid systems.
- Creating pricing structures, which incentivise a change in problematic usage patterns.
- Predicting the effectiveness of demand side response or demand side management.

II. METHOD

This section describes the methods used for demand disaggregation. Sections *A* and *B* describe how the problem is represented mathematically, sections *C* and *D* describe the algorithms used, and section *E* describes a validation study.

A. Hidden Markov Models

A Hidden Markov Model (HMM) is a model that can occupy a number of states, *Z*, but the state of the system cannot be directly observed; however, an output from the state, known as an emission, *X*, can be observed. This model structure is illustrated in Figure 1. In the demand disaggregation application, the state is the combination of appliances currently in use, and the emission is the aggregate electricity demand. At each time step, the model will

transition from one state, *i*, to another, *j*; the probability of transitioning from each state to each other state is represented by element element *i,j*, in transition matrix, *T*:

$$T(i,j) = Pr(i|j) \quad (1)$$

The emission values are a function of the current state, and can take the form of probability distribution functions or discrete functions.
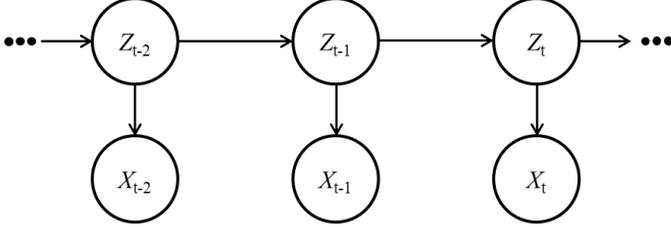

Figure 1: A Hidden Markov Model, showing a sequence of states and their corresponding emissions at each timestep.

An HMM obeys the Markov property: the probability of the next state is dependent only on the current state. This is not wholly appropriate for electricity demand, which is heavily dependent on external factors, such as time of day and weather patterns.

B. Hidden Semi-Markov Models

A Hidden Semi-Markov Model (HSMM) allows time dependencies to be represented, while maintaining the other properties of an HMM [6]. While an emission occurs every time step, transitions take place after an explicit duration, calculated when the model enters the state. This can be calculated using a probability function, such as a Poisson process, with a non-parametric vector of probabilities, or even based on an external factor such as time of day. An illustration of an HSMM models is shown in Figure 2.
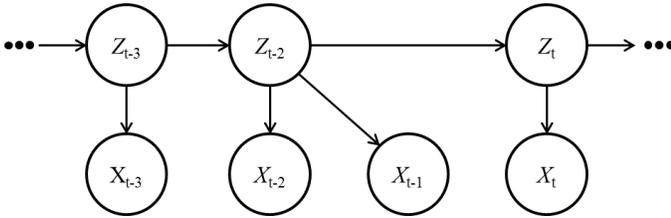

Figure 2: A semi-hidden Markov Model, showing states with varying durations, and corresponding emissions

An HSMM is used in this paper, and its emissions are represented by a vector of possible values and a vector of probabilities for each state, such that

$$\sum_{i=1}^{N} Pr_i = 1. \quad (2)$$

State durations are represented in a similar way, with a vector of probabilities for each state; the length of this vector represents the greatest number of time steps that can occur between transitions.

C. Inference Algorithms

The HSMM provides a representation of the problem, but not a means to solve it entirely. An inference algorithm is required to map an observed sequence of observations to the most likely sequence of states. Various algorithms have been demonstrated to work for demand disaggregation including Forward Backward algorithms [7], Viterbi algorithms [8] and Bayesian inference [9, 10]. In this paper, we used a modified form of the forward backward algorithm proposed by Yu [11]. The algorithm fulfills two functions:

- State estimation – calculating the most likely sequence of states based on a sequence of emissions
- Parameter re-estimation – updating the model parameters, based on the estimated state sequence and observed data

These functions are carried out iteratively, resulting in an improved model and state sequence estimate after each iteration. Convergence is tested using the log-likelihood of the state sequence.

D. Appliance Identification

The forward-backwards algorithm identifies a number of states within the data; each state represents a combination of appliance on/off states. For the disaggregation results to provide a useful input to planning or control, each state must be mapped onto its corresponding combination of these appliance states. A procedure to automate this mapping has been developed at Newcastle University, based on either a survey of microgrid customers or a list of likely appliances, and either measurements or estimates of their power consumption. The mapping uses the expected power of each state, EP_i, calculated as the sum of power draw, P_j, weighted by the probability of that power draw being observed, Pr_j,

$$EP_i = \sum_{j=1}^{M} Pr_j . P_j \quad (3)$$

and compares it with the power value of each appliance combination. The state is then mapped to that appliance combination, and the procedure is repeated for each state. The sum of the differences between EP_i and the mapped power, PM_i is then calculated:

$$SP = \sum_{i=1}^{N} EP_i . PM_i \quad (4)$$

Estimated power draws of the appliances are then updated, and the process repeated to minimize *SP*. The disaggregation and appliance identification process is shown in Figure 3.

E. Validation Study

One of the difficulties in performing the disaggregation in the remote communities that are likely to use an islanded microgrid is the lack of ground truth data. Consequently, a validation study was carried out in the Newcastle University Smart Grid Lab. Four loads – of power draw 3.8 W, 3.7 W, 36.9 W and 31.4 W – were connected to a monitored socket, and switched on and off in a controlled manner, while the overall power consumption was monitored.

One of the challenges of the validation was to ensure that the switching patterns were representative of appliance use in the types of systems targeted by ESCoBox. To achieve this, the disaggregation code was run on four different customers, and one appliance from each customer was extracted; these four independent appliance on/off sequences were then used

to switch the loads in the experiment, resulting in a new set of demand data. This approach allowed appliance usage patterns to be maintained, without using the exact output from the algorithm to be validated as in input.

The demand data available on a micro-grid scale is typically sampled at a 10 minute resolution. In the experiment, the sampling rate was 1 second, and the switching signals were issued every 2 seconds – ensuring full observability by Nyquist's sampling theorem. This allowed an equivalent of 40 days of data –some 6000 data points – to be recorded in around 3 hours, as shown in Figure 4.

The model's performance is shown in TABLE I; the algorithms had an overall accuracy of 92.4%. The algorithms were unable to differentiate between the two low energy lamps, but estimated their combined state with an accuracy of 93.9%. The model was slightly worse at estimating the state when one or more appliances were turned on – particularly a combination of the low energy lamps.
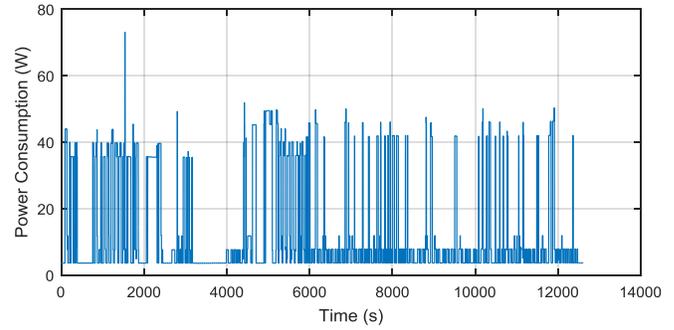

Figure 4: power consumption data for the validation study

III. EXAMPLE RESULTS

A. Case study data

For the work described in this paper, demand data was available from a microgrid on Mageta Island, Kenya, operated by SteamaCo [12]. The microgrid in question uses PV panels and lead-acid batteries to provide power to customers; data was available for the entire microgrid, and for each of 43 individual customers – some residential and some commercial. The system data was available on an hourly basis, and the customer level data was available at a variety of sampling rates, the fastest of which was 10 minutes, but the majority of the data was 30 minutes. Customers on the island use a pre-payment system, which leads to periods of zero consumption between payments.

B. Disaggregation results

The disaggregation and appliance identification results in a time series of states, each of which is labelled according to the combination of appliances that it comprises. Examples for commercial and residential customers are shown in figures 5 and 6 respectively. The algorithm has identified appliances with a power consumption as low as 2W.

The results illustrate the difference in complexity between commercial and residential customers – the algorithm is able to condense a residential customer's appliance usage into 5 states, while the commercial customer used 30 different combinations of appliances. Both of these numbers are significantly smaller than the possible state space for each, which would have yielded $2^3=8$ combinations for the residential customer and $2^8=256$ combinations for the commercial customer.

Figure 3: A flow chart showing the state inference and identification algorithm

TABLE I
Validation study results

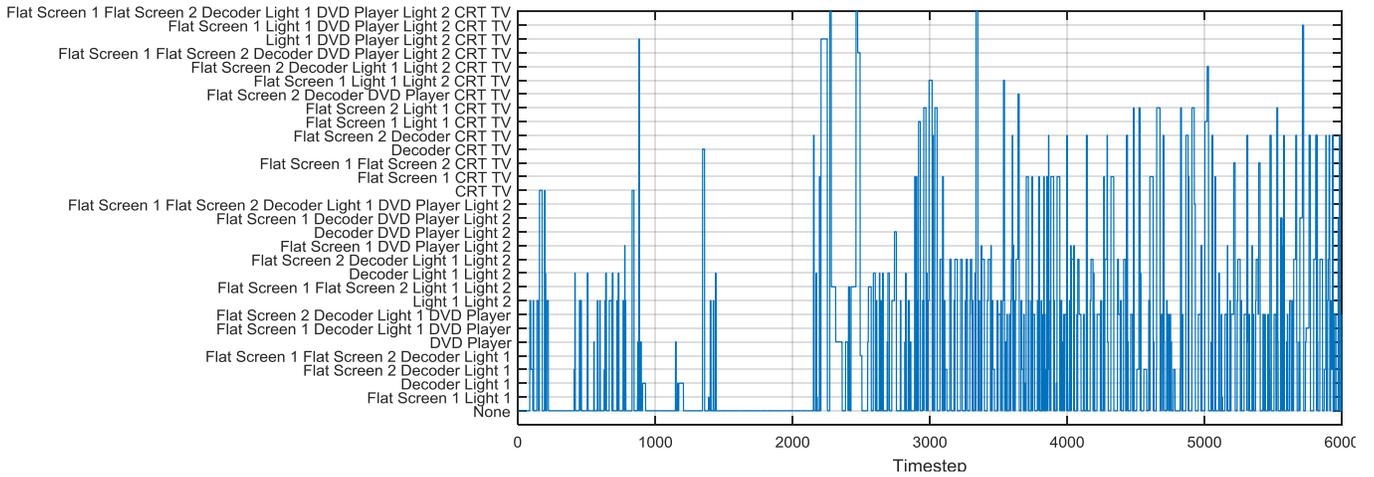| Appliance	On-State Accuracy	Overall Accuracy
Low Energy Lamps	85.18%	93.92%
Small Halogen Lamp	91.03%	98.90%
Large Halogen Lamp	84.13%	98.61%
State of System	88.02%	92.40%

Figure 5: Disaggregated demand states for a commercial customer Applications

IV. APPLICATIONS

A. Appliance usage by hour

One of the goals of the ESCoBox project is to allow microgrid managers to reduce peak demand by changing the usage patterns of particular – ideally high power – appliances.

Disaggregation results can be used to investigate the likelihood of appliances being used at different times of day, and therefore to ascertain whether time-shifting the appliance – perhaps through timeslot selling, or variable pricing – is likely to be impactful on peak demand. Figure 7 shows the probability of appliance use for the commercial customer shown in Figure 5. The probability of use for each appliance in each hour was calculated using equation (5). $Pr_{A,H}$ represents the probability of an appliance, A being in use during an hour, H.

$$Pr_{A,H} = \frac{1}{N}.\sum_{i=1}^{N} S_{A,H} \begin{cases} \text{On state}, S = 1 \\ \text{Off state}, S = 0 \end{cases} \quad (5)$$

These results suggest that the majority of appliance use takes place during the day, with some use in the evening. The appliances used throughout the night could be for security (the light), or may have been left on inadvertently (the DSTV decoder).

B. Synthetic demand data creation

Another goal of the ESCoBox project is to allow microgrid managers to understand the demand of a future microgrid, so that its power and energy resources can be built appropriately. The HSMMs described in this paper can be run backwards to produce synthetic demand data for a particular type of customer.
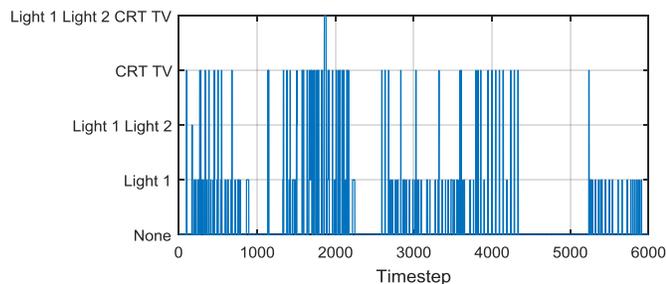

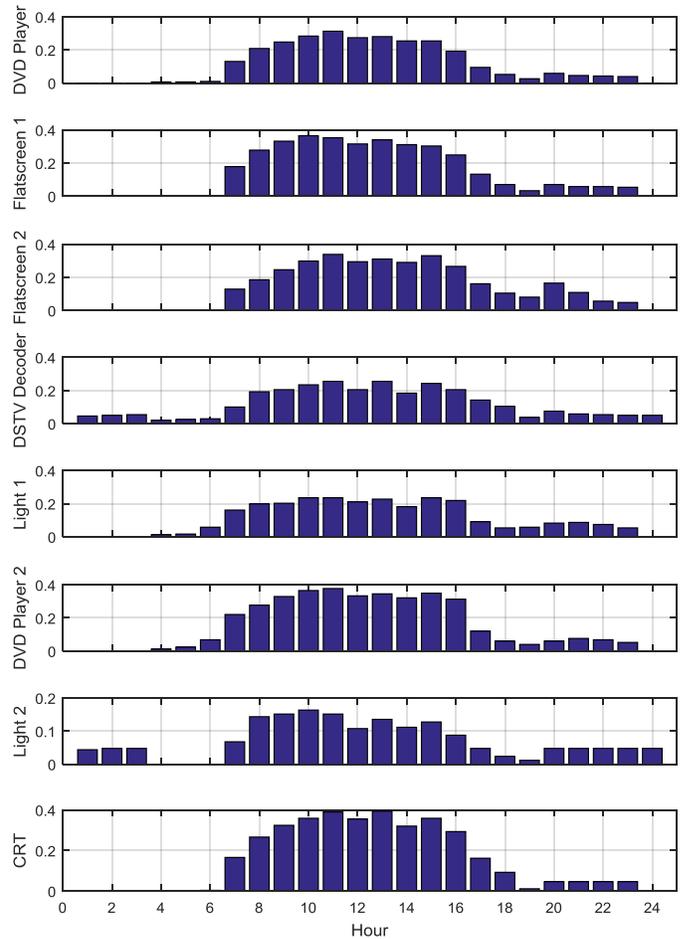Figure 6: Disaggregated demand states for a residential customer

Figure 7: Disaggregated appliance usage for a commercial customer with 8 appliances.

If the approximate number and types of customers to be served by a new microgrid are known, then that mixture of customers can be simulated – from a database of disaggregated models – and can illustrate the likely peak demand, and general electricity usage patterns. Figure 8 and equation (6) show how the models would be combined to produce this data.

This simulation approach involves using random numbers to trigger the emissions, durations and transitions. By applying a forced correlation between the random numbers used to trigger each customer model, it is possible to alter the correlation between electricity demand for those customers [13], and therefore modify the after-diversity maximum demand within the network.
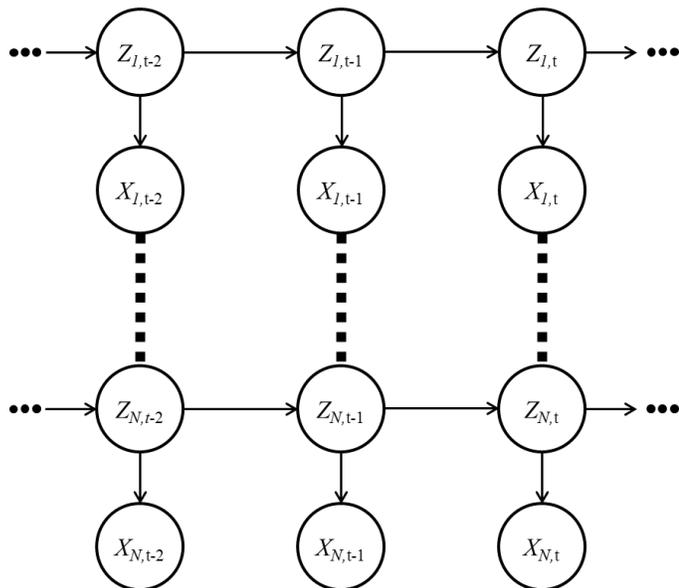

Figure 8: The structure used to generate synthetic microgrid demand data from customer disaggregation models

$$X_t = \sum_{i=1}^{N} X_{i,t} \qquad (6)$$

V. FUTURE WORK

The work presented in this paper represents the first steps for demand disaggregation in this type of system, and one of the building blocks of the ESCoBox project. Consequently, there is scope for substantial further work, including:

- Investigating the influence of different data sampling rates.
- Creating synthetic demand data from a range of customers.
- Combining demand data with battery degradation studies for a given system.
- Combining demand disaggregation with a demand management scheme, such as time of use tariffs.

VI. CONCLUSIONS

This paper has presented methods and applications for demand disaggregation in off-grid microgrids in developing countries. The disaggregation and appliance identification algorithm used was described in detail, and validated experimentally. Example results were shown using demand data from a real microgrid in Kenya. The next step is to make use of these methods to make off grid microgrids more reliable and more economical, by combining them with

analysis of battery degradation and demand management schemes.

REFERENCES

[1] A. Chaurey, M. Ranganathan, and P. Mohanty, "Electricity access for geographically disadvantaged rural communities—technology and policy insights," *Energy Policy,* vol. 32, pp. 1693-1705, 10// 2004.

[2] N. El Bassam, "Renewable energy for rural communities," *Renewable Energy,* vol. 24, pp. 401-408, 11// 2001.

[3] D. Yamegueu, Y. Azoumah, X. Py, and N. Zongo, "Experimental study of electricity generation by Solar PV/diesel hybrid systems without battery storage for off-grid areas," *Renewable Energy,* vol. 36, pp. 1780-1787, 6// 2011.

[4] A. F. Crossland, O. H. Anuta, and N. S. Wade, "A socio-technical approach to increasing the battery lifetime of off-grid photovoltaic systems applied to a case study in Rwanda," *Renewable Energy,* vol. 83, pp. 30-40, 11// 2015.

[5] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert, "Is disaggregation the holy grail of energy efficiency? The case of electricity," *Energy Policy,* vol. 52, pp. 213-234, 1// 2013.

[6] S.-Z. Yu, "Hidden semi-Markov models," *Artificial Intelligence,* vol. 174, pp. 215-243, 2// 2010.

[7] G. Zhenyu, Z. J. Wang, and A. Kashani, "Home Appliance Load Modeling From Aggregated Smart Meter Data," *Power Systems, IEEE Transactions on,* vol. 30, pp. 254-262, 2015.

[8] M. Zeifman, "Disaggregation of home energy display data using probabilistic approach," *Consumer Electronics, IEEE Transactions on,* vol. 58, pp. 23-31, 2012.

[9] A. Marchiori, D. Hakkarinen, H. Qi, and L. Earle, "Circuit-Level Load Monitoring for Household Energy Management," *Pervasive Computing, IEEE,* vol. 10, pp. 40-48, 2011.

[10] M. J. Johnson and A. S. Willsky, "Bayesian nonparametric hidden semi-Markov models," *J. Mach. Learn. Res.,* vol. 14, pp. 673-701, 2013.

[11] Y. Shun-zheng and H. Kobayashi, "An efficient forward-backward algorithm for an explicit-duration hidden Markov model," *Signal Processing Letters, IEEE,* vol. 10, pp. 11-14, 2003.

[12] J. Walton, "Ashden Award winners are leading a green energy revolution," *Renewable Energy Focus,* vol. 16, pp. 150-152, 12// 2015.

[13] D. M. Greenwood and P. C. Taylor, "Investigating the Impact of Real-Time Thermal Ratings on Power Network Reliability," *Power Systems, IEEE Transactions on,* vol. 29, pp. 2460-2468, 2014.