Forecasting Multi-Appliance Usage for Smart Home Energy Management

Ngoc Cuong Truong, James McInerney, Long Tran-Thanh,
Enrico Costanza and Sarvapali D. Ramchurn
University of Southampton, UK.
{nct1g10, jem1c10, ltt08r, ec, sdr}@ecs.soton.ac.uk

Abstract

We address the problem of forecasting the usage of multiple electrical appliances by domestic users, with the aim of providing suggestions about the best time to run appliances in order to reduce carbon emissions and save money (assuming time-of-use pricing), while minimising the impact on the users' daily habits. An important challenge related to this problem is the modelling the everyday routine of the consumers and of the inter–dependencies between the use of different appliances. Given this, we develop an important building block of future home energy management systems: a prediction algorithm, based on a graphical model, that captures the everyday habits and the inter–dependency between appliances by exploiting their periodic features. We demonstrate through extensive empirical evaluations on real–world data from a prominent database that our approach outperforms existing methods by up to 47%.

1 Introduction

Energy security is recognised as one of the most important challenges of this century (Department of Energy & Climate Change, 2009a). Indeed, as countries move to a low–carbon economy and ageing power stations are decommissioned, it is becoming increasingly important to reduce energy usage and the associated CO_2 emissions at all levels: domestic, industrial, and commercial (Department of Energy & Climate Change, 2009b). At the domestic level, a set of agent–based demand–side management techniques have recently been proposed to optimise the schedule of loads in order to minimise peak demand and hence reduce the need to operate carbon-intensive power plants (Ramchurn *et al.*, 2011; Gil-Quijano and Sabouret, 2010; Kashif *et al.*, 2011). In particular, these approaches take into account the real time carbon content/cost of electricity in order to optimise the schedule of specific loads. However, they typically do not take into account the homeowner's preferences in their optimisation. Thus, such scheduling methods may eventually not be acceptable to homeowners as they are not compatible with their everyday routine. For example, suppose that a homeowner prefers to use the washing machine on weekends when he has time to take the clothes out to dry and iron them. Thus, he would not accept a suggestion to use the washing machine on weekly day, even though it may be cheaper to do so.

Moreover, demand–side management algorithms generally ignore inter-dependencies between the usage of different appliances. In particular, the homeowner might use the dishwasher and the oven on the same day, or prefers to turn on the TV whenever he starts cooking. Suggested schedules that do not take these inter–dependencies into account may not meet the homeowner's preferences, and thus, not be accepted.

To produce rescheduling suggestions that meet the homeowners' preferences and are therefore acceptable, it is crucial to forecast their energy consumption activities. Taking such forecasts into account, an agent would be able to provide more informed advice about how to plan the usage of appliances a day ahead to reduce cost and CO_2 emissions. Here, we focus on the prediction aspect of this problem. To date, human activity prediction models have typically been designed for location prediction (González *et al.*, 2008; McInerney *et al.*, 2012), and thus, may not be adaptable for modelling complex inter–dependencies between the usage of different appliances within a typical home (a key difference is that while location prediction has to deal with only one data stream, in our domain of application we have multiple concurrent data streams, one per appliance). In contrast, a number of efficient methods for tackling complex prediction problems with multiple inter–dependent data streams have been developed (Gunawardana *et al.*, 2011). However, as we will show in that since these are not designed for human activities, they do not perform well for our application.

Against this background, we propose a novel approach to predicting the energy consumption of different home appliances, that takes into account both the human routine activities and the inter–dependency between appliances, relying on the assumption that human behaviour follows certain cyclic patterns (González *et al.*, 2008). Through empirical evaluation on real-world data, we demonstrate that our approach outperforms the state–of–the–art. Thus, this paper advances the state–of–the–art as follows:

- We propose the first, graphical model based, algorithm that can address both human behaviour prediction and inter–dependency pattern identification to efficiently predict the usage of electrical appliances in the home.
- We demonstrate through extensive empirical evaluation,

using real–world data, that our algorithm outperforms the state–of–the–art by up to 47%.

The remainder of this paper is structured as follows. In Section 2, we review existing models that could potentially be applicable to our scenario. We then formalise our problem scenario in Section 3. In Section 4 we experimentally evaluate the algorithm and analyse the results, and in Section 5 we discuss the further steps that need to be made in intelligent home energy management systems. Finally we present concluding remarks in Section 6.

2 Related work

To date, research in the home energy management domain typically has neither addressed user behaviour prediction nor the inter–dependency between the usage of different appliances (Gil-Quijano and Sabouret, 2010; Kashif *et al.*, 2011; Kolter and Ferreira, 2011). In particular, Gil-Quijano and Sabouret (2010) use reinforcement learning mechanisms to predict user behaviour. More recently, Kashif *et al.* (2011) provide an agent-based framework to analyse the user behaviour. However, these do not address the challenges of inter–dependencies between different sequences of data, and thus, are not suitable for our settings. On the other hand, Kolter and Ferreira (2011) aim to predict the energy usage of a whole building, but do not take into account the periodic nature of user behaviour.

To describe inter–dependencies within time series of events of different types, graphical models have been shown to produce promising results. In particular, graphical models have been used to represent the structure of conditional independence among random variables (Didelez, 2008), while Bayesian networks (Aguilera *et al.*, 2011; Heckerman *et al.*, 2001) are widely used for cases when missing data entries occur. Combining different approaches, Gunawardana et al. (2011) proposed PCIM, a technique for modelling inter–dependencies of Web interaction data streams. These algorithms, however, are not designed to exploit the cyclic behaviour of human users, and thus thye fail in predicting human related data sequences (see Section 4 for more details).

On the other hand, prior work on human behaviour prediction has mainly been in the specific context of predicting the position of mobile phone users in space and time. These approaches include, but are not limited to, prediction tasks with eigenvalue decomposition (Eagle and Pentland, 2009), non-linear time series analysis of arrival times (Scellato *et al.*, 2011), and variable order Markov models (Bapierre *et al.*, 2011). A number of projects also relied on the use of the Dirichlet Process to detect whether users are away from home (Tominaga *et al.*, 2012; Gao *et al.*, 2012). In addition, McInerney et al. (2012) addressed the problem of predicting human behaviour with sparse data. Although these techniques are efficient at predicting a single user's behaviour, they do not address the challenges of the inter–dependency between different sequences of data (i.e., history of appliance activity). In addition, these algorithms can only predict within the temporal scope of one day, given initial observations of the same day, while in our scenario we need to forecast electricity consumption at least one day ahead. In the next section, we present a graphical model based approach to address the aforementioned shortcomings.

3 Appliance Usage Prediction

In this section we propose a model for the prediction of appliance usage. Our main goal is to generate time–specific predictions of appliance usage based on historical behaviour. In more detail, given a time context indicating the day of the week, and a set of training data of past behaviour, we wish to predict *which* appliances are likely to be used, and *when* they are likely to be used during the day. Therefore, we are concerned with modelling discrete binary information, $x_{n,l,t}$, indicating whether appliance l was used on day n at time t. In probabilistic terms, this problem requires us to find the conditional probability $p(x_{n,l,t}|\mathbf{X}, n, l, t)$, where $\mathbf{X}$ represents history appliance use behaviour.

In what follows, we present our approach to this modelling problem. In particular, we present our model based on appliance interdependency in Section 3.1. We then give the algorithm for model inference (based on training data) in Section 3.2, and finish with the equations required for performing prediction with this model in Section 3.3.

3.1 The Inter–Dependency Clustering Model

Complex human behaviour involving interdependent streams of (appliance) activity is a difficult problem to address. A key assumption we make here is that such behaviour comes in blocks of fixed size, where each block represents a single day of activity. This approach has been effective in related areas of human presence prediction (Tominaga *et al.*, 2012) and derives from the periodic features of human behaviour that have been widely observed in empirical data (González *et al.*, 2008). We now consider the conditional dependencies between day blocks of behaviour for a single household (multiple day dependencies) and within day blocks (intra–day dependencies).

In general, we eschew complicated dependencies between day blocks in favour of the assumption that each day of behaviour is independent of any other day given the assignment of days to discrete classes of behaviour. Since supervised labels of these assignments are unavailable, we consider them to be latent random variables in our model. These latent classes compactly represent sets of behaviours that we call *day types*. As we will show, it is possible to infer the nature of these day types in an unsupervised way. Intuitively, day types can be understood as representing e.g., working days, weekend days, or family visiting days. At a basic level, day types can be captured by a mixture model with non-standard likelihood structure that we deal with later in more depth.

We now discuss the random variables controlling observations within each day block (i.e., intra–day dependencies). Since we are interested in predicting far ahead in time, (i.e., the next day or next several days of appliance usage), there is little advantage in making behaviour at one time of the day dependent on behaviour at another time, because we have already made the assumption that each day's behaviour is generated from a hidden day type class. In more detail, because day types are never directly observed (i.e., the random variable indicating latent assignment is never instantiated), we

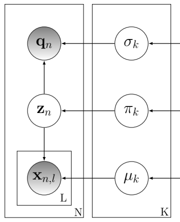

Figure 1: *Graphical model for the usage of multi appliances in the home. Shared nodes indicate observed information.*

already achieve dependencies between appliance usage at different times of the day. This can be intuitively understood as a flow of information between the random variables indicating appliance usage at different times of the day, all flowing *via* the latent assignment of that day's behaviour. We may make a similar argument for dependencies between appliances (e.g., between the oven and the kettle, or between the television and lighting). In summary, we achieve the desired dependencies between appliance usages whilst simultaneously using the fast and well–established machinery of mixture modelling, by taking advantage of the periodicities of routine behaviour and the fact that appliance use is explained by a set of uninstantiated day types (e.g., weekend, holiday).

We now formalise these assumptions into a full Bayesian model of the appliance usage in a single home, that avoids problems of over–fitting which can affect models involving large numbers of parameters relative to the number of observations. This requires the specification of two main components to the model (along with their respective parameters): the likelihood function of appliance usage and day of the week observations, and, the prior distribution over latent day types. The observations and parameters to this model are summarised in Figure 1. The dependencies between parameters is represented by directed arrows. In this graph, the observation $\mathbf{x}_{n,l}$ ($n \in N$) depends on the day types μ_k ($k \in K$), the indicator parameter $\mathbf{z}_n$ (generated from the prior distribution π_k) that indicates which day type the observation $\mathbf{x}_n$ belongs to, and the day of the week $\mathbf{q}_n$ ($\mathbf{q}_n \in W$) is parameterised by multinomial distribution σ_k with prior Dirichlet distribution γ. In what follows, we elaborate our graphical model by showing how we form the likelihood functions of the appliance behaviour and how we employ the Dirichlet Process Mixture to estimate day type classes.

Likelihood Functions

Starting with the likelihood functions, the behaviour for each appliance throughout a day can be represented by a Bernoulli distribution:

$$p(\mathbf{x}_{n,l}|\mu) = \mu^{\mathbf{x}_{n,l,t}}(1-\mu)^{1-\mathbf{x}_{n,l,t}} \qquad (1)$$

where $\mathbf{x}_{n,l}$ be the observation of appliance l on n^{th} day; and $\mathbf{x}_{n,l,t} = \{0,1\}$ be the observation of appliance l on n^{th} day at time slot $t \in T$ of the day (0: not being used, 1: not being used). For example, Figure 2 shows the use of the washing machine on a specific day (i.e., $\mu_{n,l,t}$ for $0 \leq t \leq 24$). In this graph, we can see the washing machine is likely to be used around 18:00 with the probability is 0.9. The day type can be used to describe the behaviour of the appliance on a specific day. Let $k = 1, \dots, K$ be the ID of the day types. Each

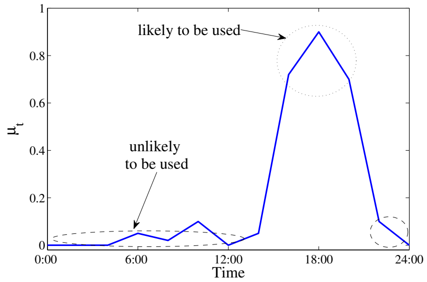

Figure 2: *An example of the washing machine usage throughout a day.*

day type has a sequence of parameter $\mu_k = (\mu_{k,l,1} \dots \mu_{k,l,T})$, where $\mu_{k,l,t} \in [0,1]$ represents for the probability of the appliance l belongs to the day type class k in the time of the day $t \in T$. Given this, the likelihood of the observation $\mathbf{x}_n$ is:

$$p(\mathbf{x}_n|\mu_k) = \prod_{l=1}^{L}\prod_{t=1}^{T}\mu_{k,l,t}^{x_{n,l,t}}(1-\mu_{k,l,t})^{1-x_{n,l,t}} \qquad (2)$$

The beta distribution is used as a conjugate prior for the parameter $\mu_{k,l,t}$. That is:

$$\mu_{k,l,t} \sim \mathcal{B}(\mu_{k,l,t}|\beta_1,\beta_2) \propto \mu_{k,l,t}^{\beta_1-1}(1-\mu_{k,l,t})^{\beta_2-1} \qquad (3)$$

where β_1 and β_2 are preset hyperparameters. Furthermore, we exploit the cyclic features of human everyday routine. In particular, we assume that human behaviour in home energy usage follows a weekly cycle. Thus, to increase the accuracy of the prediction, we condition the indicator with the day of the week. More precisely, if the goal is to predict the activity usage profile on the next day $\mathbf{x}_{n+1}$, where $\mathbf{q}_{n+1}$ can be found as the day of the week, then we only consider the same day of the week $\mathbf{q}_{n+1}$ from the past to predict the activity profile on the $(n+1)^{th}$ day.[1] As we show later, doing so can significantly improve prediction accuracy. Given $\sigma_{k,w}$ as the probability of the day type $k \in K$ belongs to the day of the week $w \in W$, the day of the week $\mathbf{q}_n$ is modelled using a multinomial distribution as $\mathbf{q}_n \sim \mathcal{M}(\mathbf{Z}, \sigma_{k,w})$ (see the dependency in Figure 1). In particular, we use a conjugate Dirichlet distribution to model $\sigma_{k,w}$, which can be expanded as $\sigma_{k,w} \sim \mathrm{Dir}(\omega)$, where ω is a preset hyper parameter.

Dirichlet Process Mixture

To denote that an appliance usage activity fits a given day type, we define an indicator $\{\mathbf{Z}\}_{n,k}$. In particular, $\{\mathbf{Z}\}_{n,k}$ indicates the observation on day $n \in N$ belongs to the day type ID $k \in K$ as is defined as follows:.

$$\{\mathbf{Z}\}_{n,k} = \{\mathbf{z}_{n,k}|\mathbf{z}_{n,k} = \{0,1\}, \text{and} \sum_k \mathbf{z}_{n,k} = 1\} \qquad (4)$$

As shown from the graphical model (see Figure 1), z_n is a random variable following multinomial distribution $\mathcal{M}(z_n|\pi)$. Thus, we employ the Dirichlet Process Mixture

[1]However, if more information is available, such as weather forecasts for the home location, or calendar information, then the model may treat these additional observations in a similar way to the day of the week observations, and they get an even more accurate prediction of appliance usage.

(DPM) to describe the infinite Dirichlet distribution with unknown component coefficient parameters as a prior distribution of day types class. Literature has shown that DPM parameterises the distribution of the size of the day type classes, and effectively estimates the number of day types as well as the parameters of the day types (Tominaga *et al.*, 2012). The number of day types can in principle be infinite. However, in practice an upper bound K is set to a suitably large value (e.g., 50, 100). The method estimates the number of day types as $k \ll K$ that is guaranteed to be bigger than what would be expected in any given case. In particular, we apply DPM truncated stick–breaking process (Sinica, 1994) to approximate the infinite–dimensional Dirichlet distribution, which represents the beta distribution as a prior of each coefficient of a multinomial distribution. We get the coefficients π_k as:

$$\pi_k = \mathbf{v}_k \prod_{i=1}^{k-1}(1-\mathbf{v}_i), \mathbf{v}_k \sim \mathcal{B}(\mathbf{v}_k|1,\alpha) \tag{5}$$

where α is a preset hyperparameter. Given this, we can now define the conjugate prior of Dirichlet distribution for π as:

$$\pi \sim \mathrm{Dir}(\pi|\alpha) \propto \prod_k \pi_k^{\alpha-1} \tag{6}$$

where α is a preset hyperparameter. If $\mathbf{M} = (\mu_1, \dots, \mu_K)$, and $\mathbf{Z} = \{\mathbf{z}_1, \dots, \mathbf{z}_N\}$ be the parameter sequences of the day types, then we can define the likelihood of the activity usage profile given all parameters is:

$$p(\mathbf{x}_n|\mathbf{Z},\mathbf{M}) = \prod_{k=1}^{K}\prod_{l=1}^{L}\prod_{t=1}^{T}[\mu_{k,l,t}^{\mathbf{x}_{n,l,t}}(1-\mu_{k,l,t})^{1-\mathbf{x}_{n,l,t}}]^{\mathbf{z}_{n,k}} \tag{7}$$

In the next section, we describe how these parameters are estimated.

3.2 Inference of Parameters and Cluster Number

It has been shown that among existing inference methods for implementing DPM, such as blocked Gibbs sampler (Ishwaran and James, 2001), variational Bayes (Blei and Jordan, 2005), collapsed Gibbs sampler (Maceachern, 1994), the blocked Gibbs sampler has higher probability to reach global optima than the others (Tominaga *et al.*, 2012). Thus, we use the blocked Gibbs sampling within our paper. The Gibbs sampling process firstly initialises the parameters randomly. Then, it iteratively alternates resampling from the posterior distributions of the unknown random variables as follows:

$$\mathbf{V},\mathbf{M},\sigma \sim p(\mathbf{V},\mathbf{M},\sigma|\mathbf{X},\mathbf{Z}) \tag{8}$$

$$\mathbf{Z} \sim p(\mathbf{Z}|\mathbf{X},\mathbf{V},\mathbf{M},\sigma) \tag{9}$$

where $\mathbf{V} = \{v_1, \dots, v_K\}$; $\sigma = \{\sigma_1, \dots, \sigma_K\}$. Based on the aforementioned discussion, the posterior distributions can be calculated as follows:

$$\mathbf{v}_k \sim \mathcal{B}(\mathbf{v}_k|1+\sum_n z_{n,k}, \alpha + \sum_{i=k+1}^{K}\sum_n z_{n,i}) \tag{10}$$

$$\begin{aligned}\mu_{k,l,t} \sim{}& \mathcal{B}(\mu_{k,l,t}|\beta_1 + \sum_n \mathbf{x}_{n,l,t}\mathbf{z}_{n,k},\\ & \beta_2 + \sum_n (1-\mathbf{x}_{n,l,t})\mathbf{z}_{n,k})\end{aligned} \tag{11}$$

$$\sigma_{k,w} \sim \mathrm{Dir}(\sigma_{k,w}|\omega + \sum_n \mathbf{q_n} z_{n,k}) \tag{12}$$

$$\mathbf{z}_n \sim \mathcal{M}(\mathbf{z}_n|\pi^*), \pi^* := \frac{\pi_k p(\mathbf{x}_n|\mu_k)p(q_n|\sigma_k)}{\sum_k \pi_k p(\mathbf{x}_n|\mu_k)p(q_n|\sigma_k)} \tag{13}$$

In Equation 11, we sample the weights $\mathbf{v_k}$, which are the beta random variables of the stick breaking construction of the DP, as per the standard stick-breaking construction (Sinica, 1994). The posterior distribution for weights can be directly calculated from the total counts of latent day type assignments and are added to the hyper parameters to find the current pseudo count of each beta distribution (Bishop, 2006). In Equation 12, a similar process is applied to the beta random variables μ indicating the probability of using appliances at all times of the day. Equation 12 defines how the day of the week probabilities for each day type are sampled from their posterior Dirichlet distribution. Finally, Equation 13 uses Bayes' theorem to incorporate the likelihood of appliance use observations and the prior distribution of day types to randomly sample the day assignments for each day block, $\mathbf{z_n}$. As per normal Gibbs sampling, we iterate through these sampling steps until convergence, then take a set of samples. Next, we show how to use this model to perform prediction.

3.3 The Prediction

Given the historical observation for all appliances $\mathbf{X}$ and the parameter estimated from the training process, we now consider the prediction of the probability of all appliance usages for the next day $\mathbf{x}_{n+1,l} = \mathbf{x}_{n+1,l,1} \dots \mathbf{x}_{n+1,l,T}$ (where $l \in L$). In this scenario, the day of week on the $(n+1)^{th}$ day is a known value, denoted as $\mathbf{q}_{n+1} \in W$. The prediction algorithm uses marginalization over unknown random variables. To obtain the model parameters, we use Gibbs sampling, which for this model has time complexity $O(NTK)$ per iteration, where N is the number of training observations, T is the number of time slots in the day (e.g., 12 or 24), and K is the number of day types. We found that 60 iterations was enough to ensure convergence, after which we retained every 3^{rd} sample (to obtain independent and identically distributed samples). Finally, we obtain the mean probability for each appliance on the prediction day $(n+1)$. The likelihood for each appliance $l \in L$ can be expanded as:

$$\begin{aligned}& p(\mathbf{x}_{n+1,l}|\mathbf{q}_{n+1},\mathbf{h}_N) = \\ & = \frac{1}{R}\sum_{r=1}^{R} p(\mathbf{x}_{n+1}|\mathbf{q}_{n+1}^{(r)}, \pi^{(r)}, \sigma^{(r)}, \mu^{(r)})\end{aligned} \tag{14}$$

where R is the number of samples obtained in the Gibbs sampling process. Thus, once the parameters have been inferred using sampling, the prediction calculation runs in $O(1)$ time. We marginalise over unknown day types for the day $(n+1)$, and take out normalising constant $p(q_n)$, then $p(\mathbf{x}_{n+1}|\mathbf{q}_{n+1},\pi,\sigma,\mu)$ can be computed as follows:

$$p(\mathbf{x}_{n+1}|\mathbf{q}_{n+1},\pi,\sigma,\mu) \tag{15}$$

$$\propto \sum_{z_n} p(x_n|\mu,z_n)p(z_n|\pi)p(q_n|\sigma,z_n) \tag{16}$$

$$= \sum_{k=1}^{K} (\sigma_{k,\mathbf{q}_{n+1}} \pi_k \mu_{k,l}) \quad (17)$$

The algorithm determines the occurrence of the target activity at time step $t \in T$ based on the value of the threshold. In particular, if the probability of the target activity at the specific time of the day is greater than or equal to the prediction threshold, the algorithm predicts that the target activity will occur at the time of the day $t \in T$ at the next day $n + 1$. Otherwise, if the probability of the target activity at the specific time of the day is less than the threshold value, the algorithm predicts that the target activity will not occur at that time of the day. We vary the threshold in range of [0,1] in order to evaluate the performance.

4 Empirical Evaluation

Given the prediction model, we now turn to demonstrating how our algorithm outperforms the existing prediction algorithms in predicting the next day usage of electrical appliances in the home. To do so, we first introduce a set of benchmark algorithms against which we compare our method (Section 4.1). We also describe two real–world datasets that we use in our experiments in Section 4.2. Finally, we describe and analyse the results in Section 4.3.

4.1 Benchmark Algorithms

As mentioned in Section 1, related work has typically focused on single user behaviour prediction and dependency model prediction for non–human data. Given this, we choose a number of state–of–the–art methods from these domains to benchmark against. In particular, we compare our method against the following approaches:

- The piece–wise constant conditional intensity model (PCIM): a state–of–the–art approach in predicting multiple–source web data where data from different sources might depend on each other. In particular, it uses a set of piece–wise constant dependency functions to capture the correlation between labels (i.e., data from different sources). Based on this model, it then estimates the probability of event occurrence in the future by using forward and importance sampling (Gunawardana et al. 2011).
- Dirichlet Process based (DP): This algorithm is designed for predicting the presence at locations of a single appliance (Gao et al. 2012, Tominaga et al. 2012). To adapt this algorithm to our settings, we run it on each appliance, as it was independent from the others.
- An extension of the Dirichlet Process based (DP–Ext): We extend the model of Tominaga et al. (2012) to capture inter–dependency by assuming that if two appliances share the same day type class, they are highly correlated.

We refer to our algorithm as GM–PMA (for graphical model–based prediction of multi–appliance usage). Recall that both DP and DP–Ext can only predict within a day (and not day or days ahead). Thus, in order to be able to compare the performance of these algorithms with that of GM–PMA, we only consider the second part of the day (although GM–PMA can predict the whole day or days ahead).

4.2 Real–World Datasets

In this section, we describe two datasets that are collected from a field trial of energy feedback systems and are used in our experiments to evaluate our algorithm and the benchmark approaches. In particular, we use the REDD dataset (Kolter and Johnson, 2011) and real–world data, collected by using the FigureEnergy system (Costanza et al. 2012). For all houses, we use the first 75% as a training data set, and the remaining 25% as a test set (i.e., comparing the appliance usage prediction against the ground–truth dataset).

The REDD dataset

The REDD dataset includes six different houses. These houses have been monitored for approximately 35 days with sub-meters installed on multiple relevant electrical home appliances. The raw data in the REDD set is the power consumption for the specific devices every 3 seconds. We converted the raw data of power consumption into a list of cyclic on–off events (i.e., a list of tuple ⟨ appliance name, starting time, end time ⟩), and use these lists to test our prediction performance. We observed that there were 3 houses which do not have enough data to judge the performance of the prediction. Hence, we only carry out our tests on data from the other 3 houses.

Data Collected from FigureEnergy

In addition to the REDD dataset, we also use another dataset collected from homeowners in the UK. In particular, this included 13 participating homes. Each household was given an off–the–shelf energy monitoring device, which integrated into the user's home and transferred data into the application's server over the internet. Users then could observe their aggregated energy consumption from their web browser using FigureEnergy (FE) (Costanza et al. 2012), a web–based application designed for appliance usage labelling, which allows users to identify and label the activities. The FE data varies between 25 and 45 days per user.

4.3 Experimental Results

In this section, we first use the REDD dataset to evaluate the performance of the algorithms and then continue with the data collected from FigureEnergy.

Performance on REDD Data

Here, we run our algorithms to predict all the labels of the REDD dataset. We depict the Receiver Operating Characteristic (ROC) curve of the algorithms for each home in Figure 3. From this figure, we can see that our algorithm dominates all the others. In particular, the area under the curve (AUC) of GM–PMA in home 3 is 0.85, while the AUC value for DP, DP–Ext, and PCIM is 0.55, 0.61, and 0.38, respectively. In other words, our algorithm (GM–PMA) outperforms DP, DP–Ext, and PCIM by 30%, 24%, and 47% respectively in home 3. Similarly, our algorithm dominates the nearest best algorithm DP-Ext up to 17% in home 1. Note that since home 1 and home 3 have the most detailed data, all the algorithms typically performed best on this home. Note that the data from homes 4 is less detailed, and thus, all the algorithms perform worse, compared to themselves in home 3 and home 1. However, our algorithm still dominates the benchmarks. An exception is the PCIM method, which performs by far the

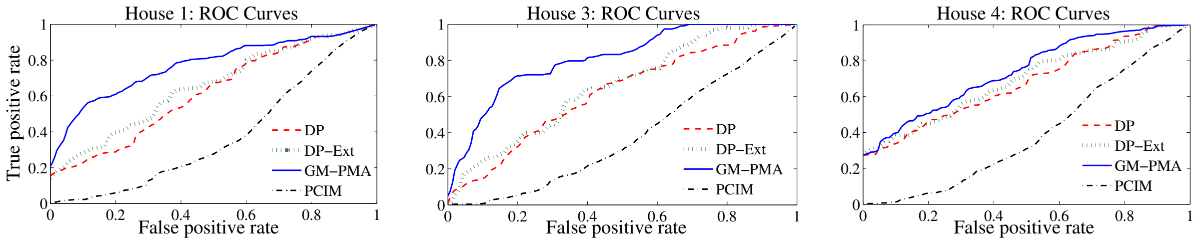

Figure 3: *ROC curve of the algorithms run on three homes from REDD.*

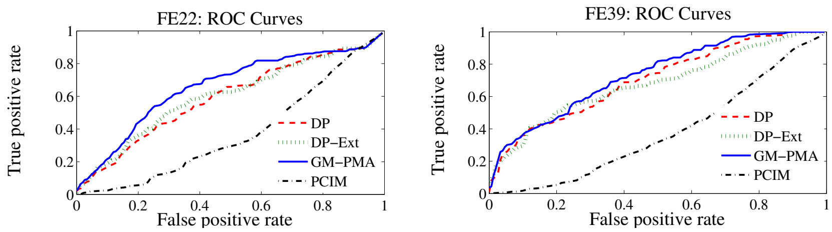

Figure 4: *ROC curve of the algorithms run on three homes from FigureEnergy.*

worst. The reason here is that due to the large size of available data, the PCIM overfits the inter–dependency model (since it does not take into account the cyclic feature of human routine). Given this, it fails to correctly detect the occurrence of activities.

Performance on Data from FigureEnergy
In this section, we test the performance on two selected homes from the FE dataset. In particular, the other homes did not provide sufficient data. Thus, we were not able to set up a proper training dataset for those homes. Similar to the previous section, we also consider the overall performance of the algorithms. Note that within the FE dataset, the labels of energy usage activities were mainly annotated by consumers. There are chances that users might mistakenly give wrong information such as incorrect type and duration of activities. Thus, the uncertainty of the labels is high and this uncertainty in labels could cause the learning structure of dependencies to behave incorrectly, and hence worsen the prediction performance.[2] Therefore we selected labels that occurred sufficiently in both training and test datasets. we also plot the ROC curve of the algorithms for these homes in Figure 4. From this figure, we can observe that, due to the uncertainty of the homeowners' manual labelling process, the performance of the algorithms are much worse, compared to the case of the REDD dataset. However, GM–PMA still provides the highest accuracy in predicting future activities. For example, GM–PMA outperforms DP-Ext, DP, and PCIM by up to 10%, 12%, and 30% respectively. In addition, note that since GM–PMA can predict a day (or multiple days) ahead, it clearly outperforms DP and DP–Ext in this aspect, as the latter two can only perform intra–day prediction.

5 Discussion

Having predicted the consumer's activities, the agent can optimise the schedule of activities such that the carbon emissions and the savings can be optimal. However, user behaviour prediction itself does not fully solve the problem, as it also has to deal with the challenge of efficient feedback. In particular, feedback goes beyond accounting for agent errors. Since an efficient energy management system aims to meet the user's comfort level, the agent may need to negotiate with the user to trade-off personal comfort with carbon emissions and electricity costs. Moreover, users may get quickly annoyed if the feedback/advice does not meet their personal comfort many times. Thus, given that GM–PMA be a powerful solution as it can be used to support the agent's advice. In particular, GM–PMA can prevent the agent from giving advice that does not fit the user's preferences.

6 Conclusions and Future Work

We investigated the problem of predicting the usage of electrical appliances in the home. To solve this problem, we proposed a graphical model based algorithm that addresses human behaviour prediction with respect to energy consumption. In particular, our algorithm models the inter–dependencies between the individual appliance usage activities and the cyclic features of homeowners' everyday routine. We also demonstrated through extensive evaluations, using real–world data taken from the REDD and FigureEnergy datasets, that our algorithm outperforms state–of–the–art methods by up to 47% in prediction accuracy.

Note that in our experiments on data from the FigureEnergy, all the algorithms (including ours) suffer from uncertainty within the labelling process of homeowners, as well as from limited training data. Since our current model does not take into account this source of uncertainty, it is not trivial to extend our approach to such settings. Given this, we aim to further study prediction with noisy or uncertain labels as future work. In addition, we intend to improve the quality of prediction by allowing interactive feedback from users, where the agent can use these feedback to learn and refine its prediction in real–time.

7 Acknowledgments

This work is supported by the EPSRC-Funded ORCHID Project EP/I011587/1.

[2]This aspect will be further investigated as future work

References

P. A. Aguilera, A. Fernández, R. Fernández, R. Rumí, and A. Salmerón. Bayesian networks in environmental modelling. *Environmental Modelling & Software*, 26(12):1376–1388, December 2011.

H. Bapierre, G. Groh, and S. Theiner. A variable order markov model approach for mobility prediction. *Pervasive Computing*, pages 8–16, 2011.

C. E. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

D. M. Blei and M. I. Jordan. Variational inference for dirichlet process mixtures. *Bayesian Analysis*, 1:121–144, 2005.

E. Costanza, S. D. Ramchurn, and N. R. Jennings. Understanding domestic energy consumption through interactive visualisation: a field study. *International Conference on Ubiquitous Computing (Ubicomp 2012)*, pages 216–225, 2012.

Department of Energy & Climate Change. Smarter Grids : The Opportunity. (December), 2009.

Department of Energy & Climate Change. UK National Strategy for Climate and Energy : Transition to a Low Carbon Society. Technical report, 2009.

V. Didelez. Graphical models for marked point processes based on local independence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):245–264, January 2008.

N. Eagle and A. S. Pentland. Eigenbehaviors: identifying structure in routin. *Behavioral Ecology and Sociobiology*, 63(7):1057–1066, 2009.

H. Gao, J. Tang, and H. Liu. Exploring social-historical ties on location-based social networks. *Sixth International AAAI Conference on Weblogs and Social Media*, pages 114–121, 2012.

J. Gil-Quijano and N. Sabouret. Prediction of humans' activity for learning the behaviors of electrical appliances in an intelligent ambient environment. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pages 283–286, 2010.

M. C. González, C. A. Hidalgo, and A. Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–82, June 2008.

A. Gunawardana, C. Meek, and P. Xu. A Model for Temporal Dependencies in Event Streams. *NIPS*, pages 1962 – 1970, 2011.

D. Heckerman, D. M. Chickering, C. Meek, R. Rounthwaite, and C. Kadie. Dependency Networks for Inference, Collaborative Filtering, and Data Visualization. *Journal of Machine Learning Research*, 1:49–75, 2001.

H. Ishwaran and L. F. James. Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association*, 96:161–173, 2001.

A. Kashif, X. H. B. Le, J. Dugdale, and S. Ploix. Agent based framework to simulate inhabitants' behaviour in domestic settings for energy management. *3rd International Conference on Agents and Artificial Intelligence*, pages 190–199, 2011.

J. Z. Kolter and J. Ferreira. A large-scale study on predicting and contextualizing building energy usage. *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, pages 1349–1356, 2011.

J. Z. Kolter and M. J. Johnson. REDD : A Public Data Set for Energy Disaggregation Research. In *In proceedings of the SustKDD Workshop on Data Mining Applications in Sustainability*, number 1, pages 1–6, 2011.

S. N. Maceachern. Estimating normal means with a conjugate style dirichlet process prior. *Communications in Statistics - Simulation and Computation*, 23(3):727–741, 1994.

J. McInerney, A. Rogers, and N. R. Jennings. Improving location prediction services for new users with probabilistic latent semantic analysis. *4th International Workshop on Location-Based Social Networks*, 2012.

S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. Jennings. Agent-Based Control for Decentralised Demand Side Management in the Smart Grid. In *The Tenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pages 5 – 12, 2011.

S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell. A spatio-temporal prediction framework for pervasive systems. *STAMI Workshop at IJCAI*, 6696:152–169, 2011.

S. Sinica. A constructive definition of dirichlet priors. *Journal of the American Statistical Association*, 4:639–650, 1994.

S. Tominaga, M. Shimosaka, R. Fukui, and T. Sato. A unified framework for modeling and predicting going-out behavior. *Proceedings of the 10th international conference on Pervasive Computing*, pages 73 – 90, 2012.