Energy and Buildings 125 (2016) 219–230

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

An occupant-differentiated, higher-order Markov Chain method for prediction of domestic occupancy

Graeme Flett*, Nick Kelly

Energy Systems Research Unit (ESRU), Department of Mechanical and Aerospace Engineering, University of Strathclyde, Glasgow, UK

ARTICLE INFO

Article history:
Received 2 September 2015
Received in revised form 26 April 2016
Accepted 6 May 2016
Available online 9 May 2016

Keywords:
Occupancy
Markov chain
Domestic
Energy demand
Microgeneration
Higher-order

ABSTRACT

Household energy demand is closely correlated with occupant and household types and their associated occupancy patterns. Existing occupancy model performance has been limited by a lack of occupant differentiation, poor occupancy duration estimation, and ignoring typical occupancy interactions between related individuals. A Markov-Chain based method for generating realistic occupancy profiles has been developed that aims to improve accuracy in each of these areas to provide a foundation for future energy demand modelling and to allow the occupancy-driven impact to be determined. Transition probability data has been compiled for multiple occupant, household, and day types from UK Time-Use Survey data to account for typical behavioural differences. A higher-order method incorporating ranges of occupancy state durations has been used to improve duration prediction. Typical occupant interactions have been captured by combining couples and parents as single entities and linking parent and child occupancy directly. Significant improvement in occupancy prediction is shown for the differentiated occupant and occupant interaction methods. The higher-order Markov method is shown to perform better than an equivalent higher-order 'event'-based approach. The benefit of the higher-order method compared to a first-order Markov model is less significant and would benefit from more comprehensive occupancy data for an objective comparison.

1. Introduction

Technical and commercial analysis of distributed generation energy projects, particularly for smaller schemes with typically fewer than 500 households, requires a detailed understanding of the likely demand profile at both short and long time-scales. Intra-day demand diversity estimation is required to assess the sizing of localised energy supply systems, the demand management potential, and the scope for grid import/export.

Demand prediction is of particular importance for small-scale low-carbon projects. Generation may either be seasonal, intermittent or benefit from stable demand [1], and as the scale reduces individual household demand behaviours become increasingly influential. Accurate matching of supply and demand and adequate storage sizing are therefore critical for ensuring that such projects perform as anticipated. The UK Government has identified a lack of energy demand data as a key barrier to growth in low-carbon community energy and demand management projects [5].

1.1. Relationship between household characteristics, occupancy, and demand

The relationship between household characteristics, occupancy and demand is complex.

A number of factors have been shown from analysis of measured data to influence household energy demand characteristics. Yohanis [20], Haldi and Robinson [7], and McLoughlin et al. [8] have determined that these include, but are not limited to; floor area, household size, bedroom number, occupant age, income, social class, children, employment status, and tenure.

The specific influence of occupancy probability has also been identified. Capasso [3] incorporated occupancy potential as a primary demand driver in a developed demand model that combined a variety of socio-economic and behavioural factors. Yao and Steemers [19] concluded that "both behavioral determinants and physical determinants related energy-consumption are more or less influenced by people's occupancy pattern", and that employment related daytime absences were the most significant occupancy effect. An extensive review of the literature linking time-use behaviour and electrical demand was performed by Torriti [15], stating that "residential electricity demand profiles are highly

* Corresponding author.
E-mail address: graeme.flett@strath.ac.uk (G. Flett).

http://dx.doi.org/10.1016/j.enbuild.2016.05.015

correlated with timing of active occupancy, i.e. when consumers are at home and awake".

The link between household characteristics and occupancy was analysed in detail by Wilke [18] using French Time Use Survey data. Specific variations were observed based on employment, gender and day type (weekend/weekday) with age ranges also identified as a key factor for the developed occupancy model.

Despite the existing work linking both household characteristics and occupancy with demand. There has been little work done that specifically quantifies the impact of occupancy on demand and the related influence of different types of occupants (e.g. full-time workers, stay-at-home parents, retired individuals etc.).

The need for research in this area becomes more critical when considering the changing demand characteristics of dwellings; as the thermal efficiency of dwellings improves, occupancy driven electrical and hot water demands will predominate as heating demand, which is less occupancy sensitive, falls. Moreover, the residual heating load in low-carbon houses may be more closely linked to active occupancy as pre-heat times reduce and heating times tend towards actively occupied periods for a proportion of potential heating system and building thermal design combinations. Consequently, realistic predictions of occupancy patterns will be crucial in determining the characteristics of future domestic energy consumption.

1.2. Occupancy data sources

There is currently no large UK dataset that specifically tracks the occupancy of individuals over a prolonged period of time. Assessing long-term occupancy patterns for individual households is therefore difficult. However, there is extensive single day time-use data, allowing assessment of occupancy patterns across identifiable sub-populations.

The UK Time-Use Survey (TUS) dataset compiled in 2000/2001 [11] was used for the initial analysis and final model development for the work reported in this paper. The dataset comprises approximately 20,000 diaries with a 10-min resolution, with one weekday and one weekend day diary per person. Additionally a smaller UK time-use dataset was compiled in 2005 [12]. This comprises approximately 5000 diaries and includes the same data as the earlier, larger TUS survey. This later dataset was used for verification of the model outputs.

Each individual diary includes detailed personal information (age, gender, relationships to other occupants, etc.), household information (size, type, age of youngest child, etc.), and a primary activity, secondary activity (e.g. watching TV while undertaking primary activity) and location for each of the 144 10-min timesteps. 146 standard TUS activities are defined that consolidate all potential occupant activities into appropriately linked groups. For example, the 'Food Prep' TUS activity comprises all cooking and meal preparation activities. The 2000/2001 survey also includes one-week work diaries from which typical working patterns can be derived.

Torriti [16] reviewed time-use datasets, identifying some inherent problems; (1) large datasets are required to provide sufficiently representative behavioural data, (2) typical days are captured, ignoring the potential for extreme weather or communal events, and (3) TUS surveys are rarely undertaken, consequently the use of older survey data for use in future projections could yield potentially misleading results. Further, the 24-h duration of TUS diaries prevents identification of occupancy and activity patterns for individuals occurring over time periods exceeding 24 h.

Despite these limitations, TUS datasets remains the sole source for occupancy and activity data with a sufficient breadth of respondents to be representative of the overall population and also smaller sub-populations. With a 10-min time resolution, they provide sufficient data to allow effective modelling of occupancy and behaviour that affects energy use.

It should be noted that a new UK survey is to be completed in 2015 [4]. This dataset may show significant changes in daily activities and the work reported in this paper will be updated when this data becomes publicly available.

State	1	2	..	i	..	n
1	$p_{1\to 1}$	$p_{1\to 2}$	..	$p_{1\to i}$	..	$p_{1\to n}$
2	$p_{2\to 1}$	$p_{2\to 2}$	..	$p_{2\to i}$	..	$p_{2\to n}$
..	..	..	..	..	..	..
i	$p_{i\to 1}$	$p_{i\to 2}$	..	$p_{i\to i}$	..	$p_{i\to n}$
..	..	..	..	..	..	..
n	$p_{n\to 1}$	$p_{n\to 2}$	..	$p_{n\to i}$	..	$p_{n\to n}$

Fig. 1. Transition probability matrix (TPM) structure.

1.3. Prediction of occupancy for demand modelling

Grandjean et al. [6] conducted a comprehensive review of demand modelling and concluded that bottom-up models featuring stochastic occupancy prediction represented the best current method. Richardson et al. [13] and Widen et al. [17] have developed such models. These authors use a first-order Markov-Chain approach to predict changes in occupancy.

Markov-Chain (MC) techniques allow the occupancy status at a time, t, to be determined based only on the status at the previous time, t-Δt. The basis for any MC model is transition matrices (see Fig. 1). These hold the probability of transition from one state a to another state b ($p_{a\to b}$). The size of this matrix is determined by the number of independent states to be modelled. For a model with n states, an $n \times n$ matrix is required. A row in this matrix therefore contains the probabilities of a transition from some state i to all n possible states (including no change from state i) and all entries per row should sum to 1.

To calculate a sequence of states over a number of time steps, a random number R between 0 and 1 is generated for each modelled time step and the new state is determined by systematically comparing the generated random number with the cumulative probabilities, $1 \ldots n$, in the appropriate row i of the matrix. For example, if a state i persists at time step $t - \Delta t$ then k, the next state at time t, is the first cumulative probability $\sum_{j=1}^{j=k} p_{i\to j}$ that exceeds R.

For a first-order MC model, only the state at the preceding time step is considered. A second-order model considers the two preceding states. Higher-order models consider the duration of the existing state at each modelled time step.

In the Richardson et al. [13] model, the states in the transition probability matrices (TPM) are the number of active occupants in a household, ranging from 0 to h, with h being the total number of occupants. Consequently, different sized matrices are required for different household sizes: 2×2 for a 1-person household (the occupant is out or in the dwelling), 3×3 for a 2-person household (both out, one person, or two people in dwelling), etc. Widen et al. [17] model each individual independently with three potential occupancy states (inactive (sleep), active, out) requiring a 3×3 matrix.

TPMs were generated for each timestep (10-min basis for [13] and 1-min for [17]) during the day to account for changing occupancy behaviour with time. Further differentiation is also made between weekdays and weekends. Therefore, depending on the current occupancy state, day type, time period, and, in the case of [13], household size, the corresponding TPM is selected to generate the next occupancy state.