

Modeling User Behavior through Electricity Consumption Patterns

Jorge Martinez-Gil, Bernhard Freudenthaler, Thomas Natschlaeger

Software Competence Center Hagenberg

Softwarepark 21, 4232 Hagenberg, Austria

{jorge.martinez-gil, bernhard.freudenthaler, thomas.natschlaeger}@scch.at

Abstract—Reducing energy consumption in buildings of all kinds is a key challenge for researchers since it can help to notably reduce the waste of energy and its associated costs. However, when dealing with residential environments, there is a major problem; people comfort should not be altered, so it is necessary to look for smart methods which take into account this circumstance. Traditional techniques have not considered the study of human behavior when providing solutions in this field, but new human-centric paradigms are emerging gradually. We present our research on user behavior concerning electricity consumption in office buildings and residential environments. Our goal consists of inspiring practitioners in this field for developing new human-aware solutions.

Keywords-user behavior; electricity consumption; smart metering;

I. INTRODUCTION

Modern housing has been changed by complexity, diversity, increase of energy consumption, and rise of costs. For these reasons, efficiently using energy is becoming more and more important. Great efforts have been made in order to build low-energy buildings, since it is supposed that an improved design can lead to significant cost reduction, but the problem is that this kind of buildings does not always guarantee low energy performance. The reason is that most of successful strategies require some kind of cooperation from the human occupants that, in varying degrees, are going to make use of the building.

Advances in this field are not only beneficial for users, but also for energy providers since the most costly consumers to serve are those who demand electricity during a daily peak period. This consumption peak usually occurs in the evening. If utility companies do not generate enough electric energy to meet peak demand, they are forced to buy short-term contracts through the energy market. Therefore, the overall costs of electricity are highest during this peak, so in some locations where flat-rate costs are charged to the customers, utility companies are going to make no profit during this period.

This work is intended for describing our research when modeling user behavior through electricity consumption patterns. Until now, most of research in this field has been based on two main approaches: sensor-based and camera-based. However, both approaches are intrusive, expensive and not very comfortable for home occupants (since special

clothes, cameras, motion sensors, etc. are needed). This novel approach is possible due to the recent development of a new technology based on smart meters. A smart meter is an artifact which records electricity consumption in regular time periods and communicates that information to the utility company for monitoring and billing purposes. Smart meters are becoming very popular in many countries from the world since both electricity distributors and people can access to very valuable statistics. However, research community has not focused intensively in this field yet despite an appropriate analysis of this huge amount of data could lead to great economic and environmental benefits.

Therefore, we can summarize the main contributions of this work as follows:

- 1) We propose to model user behavior using electricity consumption patterns (instead of traditional sensor-based or camera-based approaches).
- 2) We present the results from a number of experiments following this paradigm.

The rest of this work is organized as follows: Section 2 overviews the related research concerning user modeling. Section 3 describes our core contribution. Section 4 presents the results from a number of experiments we have performed to show the capabilities of this paradigm. Finally, we draw conclusions and put forward future lines of research.

II. RELATED RESEARCH

First of all, we are going to define user behavior; some authors have tried to do that in the past, for instance: Hoes el al. [7] define user behavior as the presence of people in the building, but also as the actions users take (or not) to influence the indoor environment. This is a quite interesting definition since it assumes that the user behavior can be modeled in two independent ways: the first consists of trying to model only the presence of the user in the building, and the second which tries to model the actions that the home occupants will perform. Yu et al. [14] say that user behavior is the occupant attitude toward energy consumption. This means that solutions should not care about other kinds of aspects when trying to model user behavior. Our idea is in line with these definitions. In fact, for us, user behavior could be considered as the set of activities displayed by the occupant in response to its environment. It should be noted

that we use the concept of activity which involves some kind of real-time actions.

In our opinion, there are many factors which can determine user behavior, but a brief analysis shows us that each of them can influence other factors. Therefore, it is quite difficult to try to understand what causes a specific behavior in order to predict it. Some of the most widely accepted factors could be:

- Knowledge of the social norms, acquired through methods of education, communication and social programs
- Socio-cultural factors which in turn depend on ecological, environmental, and demographic causes
- Past choices and the aptitudes acquired through popularization, technical assistance or training
- Economic factors including influence of marketing and incentives

Beyond the causes which influence behavior, our goal is to build human-aware systems by reflecting the activities of the users. This could lead to a number of advantages for the occupants that should facilitate them not only to save energy, safety, and a greater sense of comfort. Some of these advantages are:

- By knowing the behavior of the home occupants, human-aware systems can take a proactive approach in switching certain lights, opening windows, or programming the washer and dryer only to operate at certain times of the day, for example.
- Modeling user behavior can provide human-aware systems with the ability to design models for fine tune energy. This can be very useful for allowing users to choose pre-paid energy plans.
- Residential buildings can receive warning messages from the electric utility companies via the human-aware systems. Occupants do not have to care about how to respond to these events, because their systems are able to predict the response of the home occupants.
- Home occupants can have less hassle which means that they do not have to waste time while service technicians visit them. This is mainly due to the fact that human-aware systems usually provide better reliability, software updates, and preventative maintenance.

At this point we have explained the notion of user behavior, some of its original causes and the advantages of catching it. Now, we are going to focus how to model it in order to technological solutions can get benefit from it.

Research community has focused on new user centric paradigms as building companies are required to enforce the laws on energy efficiency and sustainability. The reason is that the usual way to address the problem of electricity consumption has consisted of studying in advance the fixed factors which can influence the total amount of energy consumption. These factors could be the house location and type, the floor area, the number and age of home occupants,

hot water sources, lighting systems, and so on. Therefore, during many years the usual way to proceed consisted of preparing a questionnaire for getting detailed information concerning the aforementioned fixed factors [12]. Then, this information will be processed (either by human experts or by expert systems) and used for optimizing electricity consumption. The reason is that by analyzing statistics of this kind, researchers can create models to predict energy consumption in advance, and therefore design plans for reducing the derived costs. For example, outdoor conditions can affect user behavior, so that if the house is located in a not very sunny place, then home occupants will use the lighting system very often. However, modeling predictions of this kind is frequently far from being a trivial task. For example, electric devices not only consume electricity, but produce heat. This leads to an automatically increase of loading on ventilation systems and, further increasing electricity consumption. Therefore the creation of good models has been always a challenging task.

One of the recent attempts to model user behavior, proposed by de Groot et al. [6], consists of a set of fixed profiles for classifying some kinds of home occupants. This classification is made on basis of the attitude of the home occupants in relation to four specific factors: Ease, consciousness, costs and environment. Ease is the profile for people who care about comfort without concerns of electrical energy use, costs or environment. Conscious is the profile for people who care about comfort but taking into account costs and environmental friendliness. Profile Costs is the profile for people who care about costs above all factors. Finally, Environment is the profile for people who act mainly taking into account the environment.

Although this classification represents clear improvement on traditional approaches and, it allows to human-aware systems working according the preferences of the occupants, the model is too rigid. This means that it does not allow faithfully representing the behavior of the users, but it is intended for trying to classify them in a set of monolithic categories. These categories reflect general trends of users, but in practice the human behavior is much more chaotic. For this reason, most of researchers think that the concept of activity is better to model this behavior [10], [3], [4].

III. ANALYSIS OF ELECTRICITY CONSUMPTION RECORDS

Current user modeling strategies which use cameras and sensors are not as powerful and flexible as they should be. One of the reasons is that these strategies require to take into account the characteristics of building structure, building equipment, and many other influence factors. This means that tuning the control precisely to the requirements and also preferences of home occupants is a task reserved to experts. Therefore, optimizations of current approaches are hardly ever realized in full due to the large effort encountered

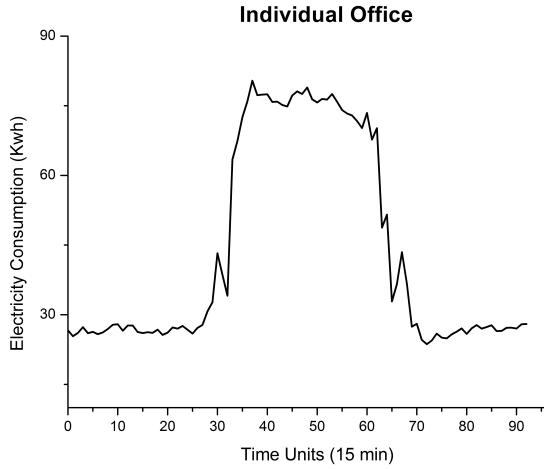


Figure 1. Typical electricity consumption curve in an individual office

[11]. As a result, there are two problems that still remain open: electricity consumption is still higher than actually necessary and users are unable to reach full comfort in their automated homes. On the other hand, traditional electrical meters only measure total electricity consumption, and so provide no information of when the electricity was spent, but the emergence of smart meters allows watching and recording how electrical energy is spent within a great level of detail. Therefore, it is now possible to work with data which could not be possible to use before by using some machine learning techniques.

Figure 1 shows us an example of electricity consumption record for an individual office. Activity begins early in the morning and reaches the maximum electricity consumption in the central hours of the day, when most of employees are in the office. During night, we can see that the electricity consumption is notably lower since only some electrical devices are supposed to be switched on.

Figure 2 shows us an example of electricity consumption record for a familiar detached house. There are two important consumption peaks; one of them happens early in the morning (when people wake up), and the other one is in the evening (when people come back from school, working place, etc). During the central hours of the day, activity is much lower since people are out of the house. However, electricity consumption is still higher than during the night.

Our working hypothesis is that it should be possible to analyze electricity consumption records in order to understand the behavior of occupants who live or work in a house or an office. This analysis can provide us with very useful information leading to identify activities from home occupants, and therefore, improve the processes concerning resource planning. In order to extract information from these records and take this kind of decisions, we have decided to use statistical classifiers. Statistical classifiers are mathematical

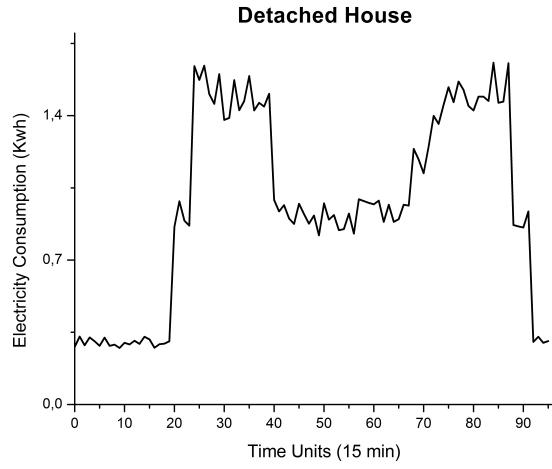


Figure 2. Typical electricity consumption curve for a detached house

tools to address the problem of identifying to which of a set of categories a new case belongs, on the basis of a training set of data containing past and annotated cases [2]. With an appropriate training, these mathematical tools may allow us to identify a lot of common situations what occur everyday in a house or an office. For our research, we have selected six statistical classifiers, one for each of the most popular paradigms for statistical classification, i.e. based on bayesian statistics, based on functions, lazy classifiers, meta-classifiers, based on decision trees, and based on rules. These are the representatives of each of them:

- *Naive Bayes* classifier which is a probabilistic classifier based on applying Bayes theorem with naive independence assumptions [8].
- *RBFNetwork* which is a classifier based on artificial neural networks that uses radial basis functions as activation functions [2].
- *KStar* which is an lazy classifier based on instances. It uses entropy as a distance measure [5].
- *AdaBoost* which is a meta-algorithm, and is going to be used in conjunction with a Decision Stump algorithm to improve the performance [9].
- *J48* which is based on decision tree algorithm trying to prune branches that reflect noise or outliers [13].
- *OneR* which is an algorithm that generates one rule for each predictor in the data, then selects the rule with the smallest total error as the only rule [1].

We propose to use these statistical classifiers to analyze electricity consumption records. One of the most important capabilities is that it is possible to analyze the electricity consumption curve in real time in order to know what is happening. For example, it is possible to analyze the electricity consumption early in the morning to know in advance how the rest of the day is expected to be.

IV. RESULTS

We present here the results concerning our research. But before to describe our major findings, it is necessary to provide some information concerning the context our experiments have been carried out. Firstly, all results presented have been obtained by analyzing sample electricity consumption records. This means that we have not used any kind of indoor or outdoor sensors or cameras. Additionally, the sample consumption records have been obtained from some consortiums of utility companies aiming to widespread the use of smart meters¹. These records have been obtained from the United States of America, so please take into account that consumption records in other parts of the world can vary. Finally, it is necessary to remark that our results are according to these sample records and there is no guarantee that can be replicated using other datasets. This is mainly due to the fact the nature of human habits is far from being deterministic since can be influenced by many unpredictable factors. In the following subsections, we describe our research results when working with individual offices and detached houses.

A. Offices

Studying electricity consumption patterns in offices can lead to a number of economic and environmental advantages. For example, it could be possible to determine the current kind of day, and therefore, take actions leading to electricity savings well in advance.

1) Distinguish between a working day or holiday early in the morning: This experiment consists of trying to distinguish between a working day and holiday or weekend early in the morning in an office. In this way, it is easier to define a strategy for generating accurate load profiles for better resource planning. For this case, we have considered two possible situations: a) it is a working day, b) it is a holiday (or weekend). Table 1 shows us the results for our pool of statistical classifiers which has been taken into account with a sample of 30 days. 15 random annotated days have been used for training, and the rest of samples have been used for determining the accuracy of the different classifiers. 45 minutes have been necessary to provide the predictions.

Classifier	#days	#training	Forecast accuracy
Naive Bayes	30	15	100.00%
RBFNetwork	30	15	93.33%
KStar	30	15	86.67%
AdaBoost	30	15	86.67%
J48	30	15	80.00%
OneR	30	15	100.00%

Table I

RESULTS CONCERNING ACCURACY WHEN DETERMINING KIND OF DAY FOR A OFFICE. KIND OF DAY CAN BE A WORKING DAY OR A HOLIDAY.
45 MINUTES HAVE BEEN NECESSARY TO MAKE A PREDICTION

2) Predict electricity consumption for a specific time frame: We have performed a number of experiments but we have not been able to achieve precise predictions concerning electricity consumption for a given timeframe. However, it is still possible to offer some predictions based on ranges. This means that for a given timeframe we are able to provide the minimum, the average, and the maximum amount of electricity that should be necessary. These predictions can be made on basis of aggregating past data. Therefore, it is supposed that the more historical data we aggregate, the better the prediction.

B. Detached Houses

Detached houses are residential buildings which do not share an inside wall with any other house. It has only outside walls and does not touch any other house. Being able to model the user behavior in this kind of residential buildings is interesting since it can lead to an important electricity savings and increase of occupant comfort.

1) Identifying user presence at home: This experiment consists of trying to determine if people are at home or not. In this way, it is possible to take decisions leading to electricity savings or comfort increase automatically. Our major problem is that the prediction should be made as soon as possible, but people can leave the house without switching off many electrical devices (like a washing machine or a dryer), so it is really difficult to provide information in a short period of time. For this case, we have considered three possible situations: a) people are at home, but they are sleeping, b) people are at home, and they are active, c) people are not at home. Table 2 shows the results for this experiment which has been taken into account with 30 samples. 15 random annotated samples have been used for training, and the rest of samples have been used for determining the accuracy of the different classifiers.

Classifier	#samples	#training	Forecast accuracy
Naive Bayes	30	15	68.75%
RBFNetwork	30	15	81.25%
KStar	30	15	75.00%
AdaBoost	30	15	62.50%
J48	30	15	62.50%
OneR	30	15	68.75%

Table II
RESULTS CONCERNING CORRECT IDENTIFICATION OF USER PRESENCE AT HOME. THERE ARE THREE POSSIBLE STATES: PEOPLE ARE SLEEPING, PEOPLE ARE ACTIVE OR PEOPLE ARE OUT OF HOUSE. 30 MINUTES HAVE BEEN NECESSARY TO MAKE A PREDICTION

It is necessary to take into account that 30 minutes have been necessary to provide a prediction. By increase the time for taking a decision, we can improve the accuracy of the classifiers. However, we think that this is not a good idea, since more time for taking decisions means less real-time capabilities.

¹<http://www.greenbuttondata.org/>

2) Identify a working day, holiday-in-house, holiday-out-house: This experiment consists of trying to determine the kind of day early in the morning. In this way, it is possible to perform a better resource planning for the rest of the day. We have considered three possible situations: a) it is a working day (adults are at work and children at school), b) it is a holiday (or weekend) and people are at home, c) it is a holiday but people are out of house. Table 3 shows the results for this experiment which has been taken into account with a sample of a month. 15 random days have been used for training, and the rest of days have been used for determining the accuracy of the different classifiers. 45 minutes have been necessary to provide a prediction.

Classifier	#days	#training	Forecast accuracy
Naive Bayes	30	15	93.33%
RBFNetwork	30	15	86.67%
KStar	30	15	100.00%
AdaBoost	30	15	100.00%
J48	30	15	93.33%
OneR	30	15	100.00%

Table III

RESULTS CONCERNING ACCURACY WHEN DETERMINING KIND OF DAY FOR A FAMILIAR DETACHED HOUSE. THERE ARE THREE KIND OF DAYS: WORKING DAY, HOLIDAY IN HOUSE OR HOLIDAY OUT OF HOUSE. 45 MINUTES HAVE BEEN NECESSARY TO MAKE A PREDICTION

3) Making recommendations for saving costs: Additionally, smart meters do not only provide data about the electricity consumption but also about the cost associated. In this way, it is going to be possible for us, to automatically recommend users ways to save costs. For example, we are able to predict when a consumption peak is going to occur, so we can automatically alert to home occupants about the amount of cost they will save in case they avoid this consumption peak.

V. CONCLUSION

We have described our current research for modeling user behavior through electricity consumption patterns. Our goal is to shed more light in a very interesting topic which can help practitioners to build systems which are not only interested in traditional aspects like outdoor and indoor data like weather forecasts, seasonal statistics, humidity indicators, and so on, but in the behavior of the occupants who live or work in a given house or office. In this way, when building a complete human aware system, it should be possible not to take into account only solutions using information from cameras (vision-based approach) or sensors (sensor-based approach), but from home occupant behavior.

As part of our future work, we have detected that current software simulators do not offer strong capabilities to include functionality for trying to simulate real-time user behavior. It could be a good idea to think on the requirements necessary to take into account the daily life of people when simulating the environment where they live in.

ACKNOWLEDGMENT

This work has been done in the research project mpcEnergy which is supported within the program Regionale Wettbewerbsfahigkeit OO 2007-2013 by the European Fund for Regional Development as well as the State of Upper Austria.

REFERENCES

- [1] Ali, S., Smith, K. A. (2006). On learning algorithm selection for classification. *Applied Soft Computing*, 6(2), 119-138.
- [2] Bouckaert, R. R., Frank, E., Hall, M. A., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H. (2010). WEKA - Experiences with a Java Open-Source Project. *Journal of Machine Learning Research* 11, 2533-2541.
- [3] Chen, L., Nugent, C. D. (2009). Ontology-based activity recognition in intelligent pervasive environments. *IJWIS* 5(4), 410-430.
- [4] Chen, L., Nugent, C. D., Wang, H. (2012). A Knowledge-Driven Approach to Activity Recognition in Smart Homes. *IEEE Trans. Knowl. Data Eng.* 24(6), 961-974.
- [5] Cleary, J. G., Trigg, L. E. (1995). K*: An Instance-based Learner Using an Entropic Distance Measure. In: 12th International Conference on Machine Learning, 108-114.
- [6] de Groot, E., Spiekman, M., Opstelten, I. (2008). Dutch Research into User Behavior in Relation to Energy Use of Residences. 25th Conference on Passive and Low Energy Architecture.
- [7] Hoes, P., Hensen, J. L. M., Loomans, M. G. L. C., de Vries, B., Bourgeois, D. (2009). User behavior in whole building simulation, *Energy and Buildings* 41(3), 295302.
- [8] Murphy, K. P. (2006). Naive Bayes classifiers.
- [9] Ratsch, G., Onoda, T., Mller, K. R. (2001). Soft margins for AdaBoost. *Machine learning*, 42(3), 287-320.
- [10] Reinhart, C. F. (2004). Lightswitch-2002: a model for manual and automated control of electric lighting and blinds, *Solar Energy* 77(1), 1528.
- [11] Reinisch, C., Kofler, M. J., Iglesias, F., Kastner, W. (2011). Thinkhome energy efficiency in future smart homes. *EURASIP Journal on Embedded Systems*, 2011, 1.
- [12] Rijal, H. B., Tuohy, P., Humphreys, M. A., Nicol, J. F., Samuel, A., Clarke, J. (2007). Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings, *Energy and Buildings* 39(7), 823836.
- [13] Rodriguez, J. J., Kuncheva, L. I., Alonso, C. J. (2006). Rotation forest: A new classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10), 1619-1630.
- [14] Yu, Z., Fung, B. C. M., Haghhighat, F., Yoshino, H., Morofsky, E. (2011). A systematic procedure to study the influence of occupant behavior on building energy consumption, *Energy and Buildings* 43(6), 1409-1417.