

Using Bayesian Networks for the Prediction of Domestic Appliances Power Consumptions

Khawla Ghribi and Dhafer Malouche
U2S-ENIT
1002 Tunis, TUNISIE
Email: ghribi.khawla2@gmail.com
Email: dhafer.malouche@me.com

Sylvie Sevestre
CEA-LinkLab
2083 Ariana, TUNISIE
Email: sylvie.ghalila@cea.fr

Zahia Guessoum
LIP6-UPMC
75015 Paris, FRANCE
Email: zahia.guessoum@lip6.fr

Javier Gil-Quijano
CEA-LIST-LADIS
F-91191 Gif-sur-Yvette, FRANCE
Email: Jesus-Javier.GILQUIJANO@cea.fr

Abstract—In this paper, we deal with the short term prediction of the domestic electrical appliances power consumption. We go further than the prediction of on/off activities of appliances to the prediction of a bounded interval of consumed power. The short term prediction of appliances consumptions is an unavoidable problem to solve when constructing optimal schedules of appliance consumptions. For reliable predictions, the appliances inter-dependencies and the cyclic characteristic of appliances activities should definitely be considered. Short term prediction is performed in the current work by graphical modeling dependencies between appliances use in a particular period of time. Obtained dependencies and short time history of appliances states are then used to infer future appliances states. We demonstrate through testing on real-world data that our model provides a promising result, nearly 90% of correct rate of classification.

Index Terms—Prediction, Home appliances power consumption, Bayesian network, Inference, Inter-dependencies.

I. INTRODUCTION

One of the greatest challenges of this century is to confront the problem of green gasses emissions and the climate change. The slowdown of the continuous growth of the electricity demand and the large-scale introduction of clean sources of electricity can significantly reduce carbon emissions. We are interested in the current paper in the efficiency of electricity consumption in the residential sector. The increase of the total domestic consumption efficiency returns to increase the efficiency of domestic appliances use.

Appliances use can be controlled automatically by proposing schedules of use. The construction of schedules is based on shifting the operation time of appliances to periods with low electricity prices, while maintaining the homeowner comfort intact.

Finding optimal schedules of appliances electricity consumptions has been treated in various prior work [1-2]. The prediction of future consumption and future prices are key variables to take into consideration when building schedules of consumptions. We are here interested in the prediction

of domestic appliances power consumption by the means of probabilistic graphical modeling. The issue of devices consumption prediction has not been widely tackled directly: generally, studies aim to predict firstly the devices on/off activities and to approximate the power consumption by its mean value.

Clearly, the prediction of appliances use is strongly related to the prediction of homeowner activities. The prediction of homeowner activities has been, in particular, addressed by predicting his mobile location. It has been done also by coupling reinforcement learning mechanism to human displacement statistics [3].

Using only homeowner activities to predict appliances activities can be inefficient if there are more than one homeowner. The include of inter-dependencies between appliances use can brings significant information.

Interdependencies between domestic appliances use can be addressed by different classes of methods. One of them is the Rule-based approach [4]. It consists in establishing dependencies relating the use of one appliance or a set of appliances to another appliance or a set of appliances. Various metrics are used to quantify the strength of the rules. Established rules have a temporal validity i.e. the use of a set of appliances can be inter-dependent in one temporal window of a day and independent in a different time window of the same day. [5-7]. Another approach that naturally models dependencies between temporal data is graphical modeling. Bayesian networks is a family of graphical models that has been previously used in [8] to model electrical appliances dependencies, authors of [8] seek the disaggregation of total energy consumption recorded at circuit level into a set of appliances consumptions. In addition to Bayesian network, another family of graphical models has been previously used in the context of home load prediction: the Markov chain which has been used in [9] to predict the total electricity demand of homes and in [10-12] to predict the appliances activities. The latter study exploits in the same time the dependencies between appliances and

the cyclic characteristic of homeowner behavior to predict the activities.

In the current paper, we aim in a first step to model domestic electric consumers and environmental features as a Bayesian network and to exploit then exact probabilistic methods to predict the interval in which ranges each individual consumption within a time horizon of 6 hours. The modelling should take into consideration, at the same time, the appliances inter-dependencies and the cyclic characteristic of homeowner behavior.

The remainder of the paper is organized as follows: Section 2 details the pre-processing performed on data. Section 3 formalizes our problem and introduces the model. A theoretical description of the prediction algorithm is given in Section 4. Our results are presented in Section 5. Section 6 concludes.

II. DATA PRE-PROCESSING

In this paper, we are using a real-world database: the Smart* database [13] which is a set of daily time series data columns. Each column represents an appliance electricity consumption collected at the electrical outlet level using a plug meter system, and environmental data recorded by a weather station attached to the monitored home. The number of electrical consumers is around thirty. We use here only temperature and humidity as environmental data.

Home appliances can be grouped in classes according to their frequencies of use or according to the number of modalities taken by each consumer. They can be classified also by ranges of the consumed power. Examples of modalities of consumption for a set of appliances are presented in Fig 1. We can distinguish as an example the consumption of dehumidifier with a big variety in taken modalities and an important mean power, in opposition to the fan that consumes, on average, 1% of the dehumidifier consumption and takes few modalities.

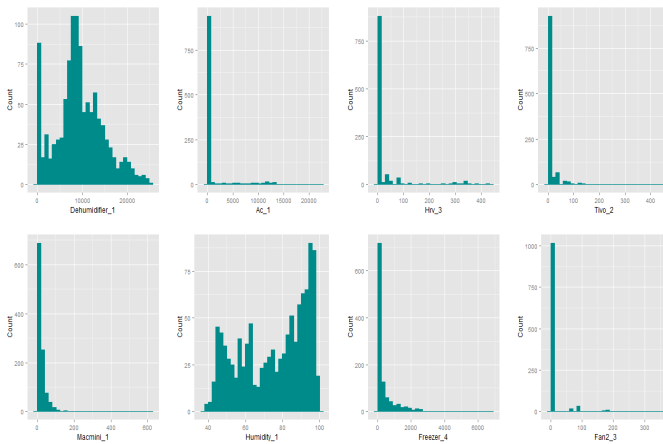


Figure 1. Example of consumers' histograms

To unify the representation of consumption of the different appliances and for methodological reasons we will discretise each appliance consumption into quin-tile ranges. Appliances with few modalities are discretised into less than 5 classes.

Thereby, we have transformed our initial database into categorical data. We are preparing the data for performing prediction of consumption within a time horizon of 6 hours. So it seems plausible to divide the daily consumption of each appliance into four sets of consumption. By this way, each appliance consumption within each set is a separate variable. For instance, Freezer (day_d) = (Freezer_1, Freezer_2, Freezer_3, Freezer_4) Freezer_n (n =1..4) is the time series consumption of the freezer during the nth 6 hours of the day_d.

III. FORMALISM OF THE PROBLEM

The prediction of individual appliances consumption is the core of the domestic load management.

In this paper, we seek to find more reliable prediction by exploiting exact probabilistic calculations and incorporating as much as possible knowledge about past consumption.

History of consumption is incorporated in our system of prediction at two levels. Knowledge about consumption of the week that precedes the current week is used to establish a model of correlation between variables which are in our case appliances consumption in different temporal windows. The establishment of the correlation model consists in the construction of a Direct Acyclic Graph (DAG) linking dependent variables and in the computation of probabilities of all possible configurations. To construct the DAG we can adopt a constraint-based method which searches conditional independence between variables in order to eliminate edges from an initially complete graph [14], as we can use a score-based method that adds edges to a initially empty graph to maximize a scoring criterion [15].

Each variable in the graph (see Fig 3) has a number of configurations proportional to the number of its parents in the graph. One possible configuration is to assign to one variable and to its parents the lowest range of consumption.

Now, having the parametrized Bayesian network (Graph + probabilities of each configuration), the prediction of the consumption of an appliance X during the first temporal window of June'2, for example, consists in entering as evidence the knowledge about the three last temporal windows in the established model and updating probabilities of each configuration known in statistical terminology as belief update [16] (See Fig 2).



Figure 2. : Required past knowledge to predict an appliance state

The process of entering evidence and probabilities update will be explained in the next Section. As mentioned above, an optimal prediction system should, in addition to capture appliances inter-dependencies, take into consideration the characteristics of the homeowners' behavior. This behavior has been proven to be weekly cyclic [17]. So, it can be

split into repetitive blocks. The assumption of weekly cycles of the homeowner behavior is exploited here by using the correlation model over a previous week to predict current state of appliances.

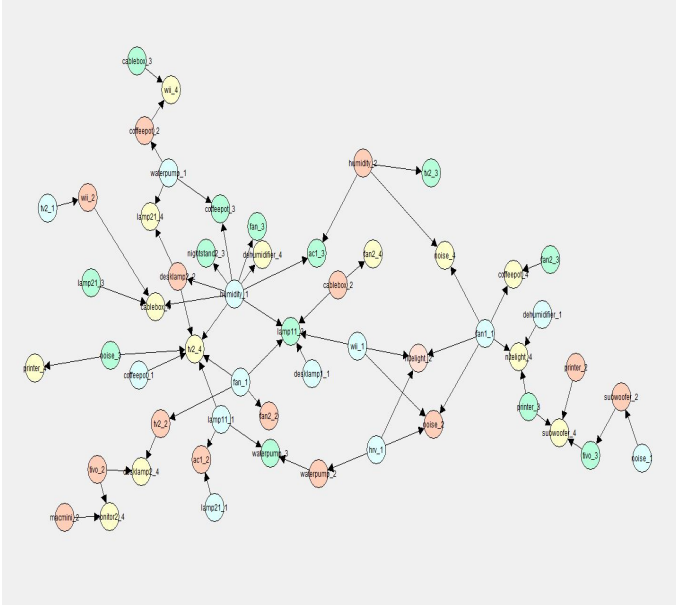


Figure 3. Example of a Bayesian network of appliances use of one week in June

IV. THEATRICAL TOOLS: THE UPDATE OF BELIEF

Bayesian network over a finite set of random variables X is a graphical model (traditionally understood to be a directed acyclic graph (DAG)) characterized by a property that connects the multivariate distribution and the uni-variate (conditional) distributions.

The representation of conditional distribution for each variable differs according to its type and the type of its parents in the graph.

In this paper, we are considering only categorical variables. In this case, the conditional distribution of each variable is specified by a table containing the probabilities of possible configurations of the variable and its parents.

The term Bayesian refers to the use of Bayesian inference. Inference is to update probabilities of a set of unobserved variables after fixing states of another set of variables as evidence. Bayesian inference is to exploit conditional independences modeled by the graphical model to realize efficient calculations of updated probabilities [18].

Technically, the term prediction means to perform Bayesian inference from a given Bayesian network to deduce the state of a response variable from explanatory variables states.

Prediction runs in two steps: setting evidence (states of the explanatory variables) and propagating the effect of the evidence on the rest of variables and particularly the response variable. The propagation algorithm consists in exchanging messages between adjacent nodes to update beliefs. It is an exact method based on local calculations. Here's below the description of

this algorithm according to Cowell in [20].

The efficiency of the propagation algorithm depends mainly on the structure in which the algorithm is executed. Intuitively, the structure needed must be modular.

Authors in [20] argue that any exact method based on local calculation and performed in a structure "S" other than the structure of junction tree¹ is either less efficient in terms of complexity of calculation than the same method performed in a junction tree structure or has an optimality problem equivalent to that of transforming "S" into a junction tree.

A. Compilation

Given a DAG, the compilation consists in the moralization and the triangulation. It ends by a new higher-level graphical model -called chordal graph- more amenable than its precedent to the application of inferences. The moralization transforms the DAG into an undirected graph first by joining all parents of each node by a line and second dropping the directions on the arrows. A new dependence graph is obtained after the moralization. Each clique of the new graph is associated to a function with values in \mathbb{R} called potential. The potential is the product of the conditional distributions of all nodes forming the clique.

The triangulation consists in adding edges between not joined variables in order to obtain a chordal² structure. It is known by the fill-in step.

Finding the optimal triangulation is an NP-hard problem. Hence, the only working way to triangulate graphs is using heuristics.

After triangulation we are no more interested by the nodes of the graph but by the cliques of the graph. The triangulation may lead to cliques without an associated potential. For each such clique we associate a potential equal to 1. The clique potential representation for the triangulated graph is:

$$p(X_{i_V}) = \prod_{j=1}^T \psi_{C_j}(X_{i_{C_j}}) \quad (1)$$

with X_{i_V} is the set of random variables, $\psi_{C_j}(X_{i_{C_j}})$ is the clique potential of the clique C_j and T is the number of cliques in the triangulated graph.

B. Junction tree

The junction tree constructed over the cliques of the triangulated graph is a tree with the junction property: for each pair of cliques A and B with intersection S , All paths between A and B contain S [20].

The construction of the junction tree over the cliques consists in:

- Constructing a spanning tree of maximal weight over the cliques (the weight is the number of variables in the intersection of the adjacent cliques).

¹1: A junction tree is a sub graph of a clique graph that satisfies the junction tree property [20].

²2: A chordal graph is Bayesian network with no V-structures.

- Giving labels to the edges of the obtained tree: edge between two cliques is labeled by its intersection (the separating set).
- Associating to each separating set a potential equal to 1.
- Associating to each clique its potential that has been calculated in the compilation step.

C. Propagation algorithm

The propagation algorithm consists in letting adjacent nodes communicate by exchanging messages in order to update their potentials and to reach the equilibrium point where the potential of each node holds the marginal of the joint probability distribution for the entire set of variables. The existence of this equilibrium is proved by Cowell and al. in [21]. The number of sent messages before reaching that point is a parameter to be optimized.

Between two adjacent nodes the update of potentials proceeds as follows: having two nodes A and B with their attached potentials $\psi(A)$ and $\psi(B)$. A and B are separated by their intersection S which has a potential $\psi(S)$ (see Fig 4).

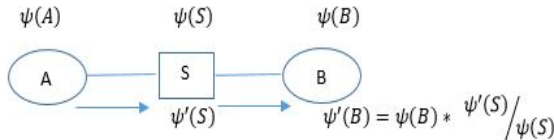


Figure 4. The message passing between adjacent nodes

The passage of a message from a node source to an adjacent node is called flow.

During the flow from A to B, the $\psi(A)$ is marginalized down to S resulting in $\psi'(S)$. $\psi'(S)$ is placed in the separator. $\psi(B)$ is multiplied by the ratio $\psi'(S) / \psi(S)$.

Each time, we have a modification of the state of one variable of the initial DAG, we localize this variable into the junction tree (to which clique it belongs) and perform an ordered message passing. The order of message passing is given by a list containing ordered directed edges of the junction tree and called schedule of flows.

V. RESULTS

Applying the propagation algorithm to our Bayesian network gives a significant coincidence between predicted values and real values.

We introduce few contingency tables of the predicted ranges and real ranges of appliances consumption.

Prediction 0 (0,66] 1 0 47 [1] "lamp11_1" [1] 0.9791667	prediction 0 (0,2] 4 (2,3] 7 (3,4] 1 (4,26] 12 0 24 [1] "nightstand2_3" [1] 0.5
prediction (1.3e+03,6.48e+03] (6.48e+03,8.06e+03]	
(1.3e+03,6.48e+03]	5 4
(6.48e+03,8.06e+03]	6 7
(8.06e+03,9.1e+03]	4 8
(9.1e+03,1.67e+04]	5 9
[0,1.3e+03]	0 0
[1] "dehumidifier_4"	
[1] 0.25	

Figure 5. Contingency tables

We notice that the rate of correct classification for variables having few modalities is nearly 100 %. This rate is significantly lower when variables take various modalities. The total rate of correct classification is near 90% which allows us to describe our model by promising.

The discretization of variables simplifies much the message passing process when updating beliefs in the network but it contributes, in parallel, to a loss of accuracy. We plan in next work to skip the discretization step and infer knowledge from graphs containing both continuous and discrete variables.

VI. CONCLUSIONS AND FUTURE WORK

This paper presents a novel approach for predicting the ranges of domestic appliances power consumption. Prediction is performed on temporal window of 6 hours. Predictions are mainly influenced by the short term history knowledge of appliances states. This approach satisfies previously established requirements for a reliable prediction of appliances activities: the consideration of inter-dependencies between appliances use and the cyclic characteristic of homeowner behavior.

We demonstrate using a real-wold data that this approach reaches high level of precision: nearly 90 % of good classification rate.

This approach is running over categorical data. The discretization of initial data facilitates the message passing when updating the beliefs but it causes also a loss of precision.

We are planning in next work to predict values instead of ranges of the power consumption.

ACKNOWLEDGMENT

We are grateful to Telnet Innovation Labs team.

REFERENCES

- [1] SmartCharge: cutting the electricity bill in smart homes with energy storage in e-Energy '12 Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet, 2012.

- [2] P. Du, N. Lu, Appliance Commitment for Household Load Scheduling in Smart Grid, *IEEE Transactions on* (Volume:2 , Issue: 2), May 2011, pp 411 – 419.
- [3] J. Gil-Quijano, 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*, pp 283- 286.
- [4] R. Agrawal, Rule base approach: Mining association rules between sets of items in large databases in *SIGMOD '93 Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, pp 207-216.
- [5] S. Rollins, N. Banerjee, Using Rule mining to understand appliance energy consumption patterns in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, 2014, pp 29 – 37.
- [6] Y.C Chen , W.C Peng ; W.C Lee, A novel system for extracting useful correlation in smart home environment in *IEEE 13th International Conference on Data Mining Workshops (ICDMW)*, 2013, pp 357 - 364
- [7] J.Palmer, N. Terry, T. Kane, S.Firth, M. Hughes P. Pope, J. Young, D. Knight, D. Godoy-Shimizu, *Electrical appliances at home: tuning in to energy saving*, Cambridge Architectural Research, Loughborough University and Element Energy under , November 2013.
- [8] Academia Sinica, Patent US20110307200, Recognizing multiple appliance operating states using circuit level electrical information, December 2011.
- [9] F.J McLoughlin, A. Duffy, M. Conlon, The generation of Domestic Electricity Profiles through Markov Chain Modelling in *Euro-Asian Journal of Sustainable Energy Development Policy Vol. 3*, 2010. [10] N.C. Truaong, L. Tran-Thanh, E. Costanza, S.D. Ramchurn, Activity Prediction for Agent-based Home Energy Management in *Agent Technologies for Energy Systems (ATES)*, May 2013.
- [10] N.C. Truong, J. McInerney, L. Tran-Thanh, E. Costanza, S.D. Ramchurn, Forecasting Multi-Appliance Usage for Smart Home Energy Management in *23rd International Joint Conference on Artificial Intelligence (IJCAI)*, Aug 2013.
- [11] N.C. Truaong, L. Tran-Thanh, E. Costanza, S.D. Ramchurn, Toward Appliance Usage Prediction for Home Energy Management in *Proceedings of the fourth international conference on Future energy systems*, 2013, pp 287-288.
- [12] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart*: An Open Data Set and Tools for Enabling Research in Sustainable Homes. In *SustKDD*, August 2012.
- [13] P. Spirtes, c. Glymour, R. Scheines, *Causation, Prediction and Search* in the MIT Press, 2000.
- [14] D. Chikering, *Optimal Structure Identification with Greedy Search* in *Journal of Machine Learning Search*, 2002.
- [15] S. L. Lauritzen, D.J. Spiegelhalter, Local Computations with Probabilities on Graphical Structures and their application to Expert Systems in *Journal of the Royal Statistical Society*, 1988, pp 157-224.
- [16] M.C. Gonzalez, C.A. Hidalgo, A.L. Barabasi, Understanding individual human mobility patterns in *Nature*, June 2008, pp 779-782.
- [17] S. Højsgaard, D. Edwards , S. Lauritzen, *Graphical Model with R*, Springer, 2012.
- [18] R.G. Cowell RG, A.P Dawid, S.L Lauritzen , D.J Spiegelhalter, *Probabilistic networks and expert systems*. Springer, 1999.
- [19] F.V. Jensen , F. Jensen, Optimal Junction tree in *UAI'94 Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*, 1994, pp 360-366.
- [20] R.G Cowell, P. Dawid, S.L Lauritzen, D.J. Spiegelhalter, *Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks*, Springer, 1999.