Energy Consumption Patterns and Inter-Appliance Associations using Data Mining Techniques

ABASSI Abdelfattah^{1,3,*}, ARID Ahmed^{2,**}, LARAKI Maha^{3,***}, and BEN-AZZA Hussain^{4,****}

Abstract. In this paper, we propose to model the behaviors of Moroccan consumers in terms of energy consumption in different Moroccan buildings using the open MORED (A Moroccan Building Electricity Dataset) dataset as a data warehouse. The techniques used are Machine Learning Algorithms and Data Mining Techniques. The results obtained in this paper allow us to understand the behavior of a Moroccan consumer in terms of energy consumption and the use of appliances in the home. Inter-Appliance Association and Peak Hours detected in this study will be used later to develop an Energy Management System specifically for a Moroccan building. This can lay the foundation for efficient Energy Demand Management while improving end-user participation.

1 Introduction

Morocco's primary energy consumption increased by 32% between 2007 and 2017 (reaching 20.5 Mtoe that year). The growth in demand for electricity is even stronger and could continue at an average rate of 5% per year by 2021 according to the International Energy Agency (IEA) [1]. There is a growing interest in applying recent technologies in relevant fields such as smart homes in smart cities. Via Casablanca, Morocco aspires to become a benchmark in terms of smart cities in Africa. Recent technologies are trending to change from the fundamental constituents of a city, smart homes, buildings, and factories in the Fourth Industrial Revolution (Industry 4.0), to smart cities. A smart city is a city that uses Information and Communication Technologies (ICT) to improve the quality of urban services or reduce their costs. The technical combination of Internet of Things (IoT) technologies with Artificial Intelligence (AI) present novel knowledge into home automation, home healthcare, home surveillance, and home energy management in smart homes. The increase in the share of renewable energies in the total electricity production, the development of decentralized production within the framework of Law 13.09, the evolution of the uses of electricity in the full expansion (air conditioning, heating, production of electric vehicles, etc...) and the demands of customers who are more and more attentive to their consumption are all factors which make the development of smart grids now a necessity to meet all these new expectations.

¹Mouly Ismail Univeristy, ENSAM, MOROCCO MEKNES

^{*}e-mail: abassi.mri@gmail.com

^{**}e-mail: ahmed0arid@gmail.com

^{***}e-mail: m.laraki@edu.umi.ac.ma

^{****}e-mail: h.benazza@umi.ac.ma

The rest of this paper is organized as follows. The previous works are explained in Section 2. The proposed methodology is introduced in Section 3. Results and analysis are proposed in Section 4. The paper is concluded in Section 5.

2 Previous works

The futures building was equipped with smart meters, sensors and photovoltaic production, and many other appliances to facilitating the controls in the grid. The goal of sensors is collecting the data relative to the building such as the inside temperature, the outside temperature, the tank temperature heat pump, and the thermal storage system. The values of this variables could be collected from other sensors that are not installed in the building but in other local in the grid. The weather file related to the study zone could be collected also from the satellite. This dataset uses by researchers to develop a stochastic models [2][3]. In [4] the authors use genetic algorithms to evaluate a set of data relating to each type of building. He develops software created in java to collect the necessary information from a SqlLite database. At each stage, the software collect data and creating an initial population and IDF file for each individual in the population. The evaluation of each individual is done using Energy Plus (EP) simulation software (fitness function). The results obtained are then used to select the best individuals and regenerate a new population. The previous steps are repeated until an approximate convergence is obtained in the algorithm. Work in [5], have developed an EMS (Energy Management System) to predict and optimize the energy consumption in the building. The building is in Ukraine. The research uses a framework to collect and store all building information and EP was using to simulate the building energy's consumption. The house is equipped with a talk thermal pomp and other smart meters to calculate other variables measure. The principal objective is the minimize the energy consumption in the building according to the user's comfort. In [6], the authors develop a DDC (Direct Costing Control) using machine learning algorithms. Firstly the data was collected from sensors that are installed in each building. The sensors detect the movement in the building and calculating the occupancy. The data is daily collected and stored in the local database. The data was divided in the tow dataset: the local dataset and the global dataset. The K-means neighbor clustering algorithm is used to clustering the data set to different occupancy patterns. The model developed was to predict occupancy for each room based on the historical dataset. The prediction allows the control of cooling in the room and takes into account the occupancy prediction. The methodology illustrated in this paper shows that a 7%-52% energy saving was obtained by an adaptative cooling system depending on occupancy rates. In [7] the authors develop a model that is capable to predict the load cooling in the building using deep learning algorithms. The data has been collected 24h time ahead. The output of the model data was extracted using statistical methods and machine learning methods. The accuracy of the model has been compared with many machine learning algorithms such as ELN, GBM, SVR n XGB. The results demonstrate that the deep learning algorithm can give a sophisticated prediction accuracy. In [8] the authors modeling occupancy and activity patterns using machine learning algorithms such as non-linear multi-class Support Vector Machines (SVMs) and other methodologies such as Hidden Markov chains and K-nearest neighbor. In [2] the authors propose stochastic processes for modeling occupancy by using Markov chains to calculate transition probabilities(between absent and present) and to represent the occupancy probability density at that day and time. The dataset was obtained by the use of motion sensors. In [9] A genetic programming algorithm was implemented to learn the behavior of an occupant in a single-person office based on motion sensor data. The learned rules predict the presence and absence of the occupant with 80%-83% accuracy on testing data from 5 different offices. Multiple techniques have been implemented to collect data from building for

understanding the relationship between occupant behavior patterns and energy consumption [10][11]. In [3] a Hidden Markov Model-based algorithm for Activity Daily Living (ADL) detection from data acquired employing a combination of motion, temperature, and analog sensors for water and cooking hobs usage monitoring was implemented. In [12] the authors propose a model for occupancy detection to build demand-driven heating. In [13], a KNN with the inclusion of a kernel function to study the energy consumption profile of a building that might be affected by external and contextual factors. The authors in [14] applied a multiagent system based on Support Vectors appliances (SVM), to optimize the HVAc system by detecting the number of occupants in a room getting results of over 90% accuracy detection.

3 Proposed Methodology

This paper aims to model the behaviors of Moroccan consumers in terms of energy consumption. The community Moroccan is divided into different states socio-economic. Each neighborhood has its characteristics. For example, in affluent neighborhoods, the energy consumption is relatively higher if we compare it with the energy in disadvantaged neighborhoods. This section presents the proposed approaches to extract the frequent patterns in the MORED dataset based on the information recorded in the "event" file associated with each premise. The data extraction is done in 4 phases: data preparation, exploitation, generation of frequent itemset and association rules, and visualizations.

The analysis and discussions in this paper focuses on two types of building for 11 buildings in total. For us, we are interested in the approaches used during this study to extract important information for the given building. These approaches could be used for any building in the grid to capture and study energy consumption and user behavior. The outline of our approach is depicted in Figure 1.

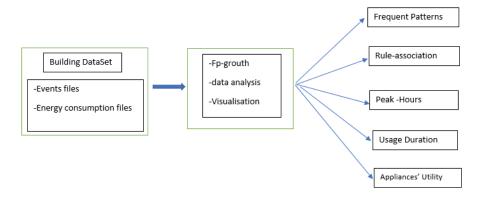


Figure 1: The Proposed Approach.

3.1 Data description

Many works in the literature [7],[8],[2] use wireless ambient sensors to get information from his surrounding simply and efficiently with no intrusive issues. Different sensors can be installed to capture various indoor parameters such as motion or inside temperature [10]. All this information is organized and stored in datasets. Some of these datasets are free and open to allows researchers to study and develop here owner works. In our work, we use the datasets described in [15]. The MORED datasets contain the energy consumption of different appliances in different neighborhoods. The energy consumption could use it for capturing different Energy Consumption Patterns based on each appliance that is active at each time in the day. MORED is a new datasets that comprises electricity consumption data of various Moroccan premises. MORED provides three main data components: whole premises (wp) electricity consumption, individual load (IL) ground-truth consumption, and fully labeled II-signature, from affluent and disadvantaged neighborhoods. Figure 2 show the premises targeted for MORED data acquisition.

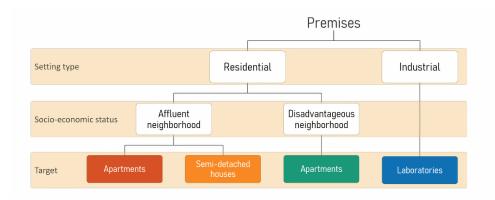


Figure 2: The premises targeted for MORED data acquisition.

3.2 Data Pre-processing

One of the major tasks is data preparation. MORED data set contains three important data set: Wall premises electricity consumption(WP), Individual Load energy consumption(ILs), and Events Datasets. For our study, we use Event Datasets to capturing the different frequents itemsets and Inter-appliances Association. For our study, we chose the important appliance in each building as presented in Table 1.

The time resolution data is extracted a 1/5[S/s].

3.3 Frequent patterns modeling and recognition

Frequent patterns are repeated patterns or itemsets, which often appear in a dataset. Considering smart meter data, an itemset could comprise, for example, of laptop and washing machine, which often present themselves together in a frequent pattern. Hence, frequent pattern mining can help discover association and/or correlation among appliances, which defines the relationship among data interpreting consumer energy consumption behavior [16].

Table 1: Appliances short code

Appliance	Shortcut
TV Monitor	TV
Hair Dryer	HD
Laptop ASUS	ASUS
Laptop MSI	MSI
Panini Grill	PG
Washing Machine	WM
Hair Straightner	HS
Oven	OV
Water Heater	WH
Blender	BL

3.4 Frequent Itemsets and Association Rules

$$R: X \to Y$$
 (1)

where X and Y are two disjoint itemsets.

The support of an $AR: X \to Y$, is the fraction of all transactions in D that support both X and Y: support($X \to Y$) = support($X \to Y$). The Support of AR is a symmetric measure: support($X \to Y$) = support($Y \to X$). The support of an AR : $X \to Y$ indicates the frequency with which this rule is applicable to a data set.

The confidence (C) of an $AR: X \to Y$, measures the strength of the co-occurrence between the antecedent X and the consequent Y. It measures the conditional probability of occurrence of the item set Y knowing Association rules the occurrence of the itemset X: confidence $(X \to Y) = \text{support}(XUY) / \text{support}(X)$. The confidence of an $AR: X \to Y$ determines how often the elements of Y appear in transactions containing X.

For measuring the interest of an association rule, we use the most used method is improvement (lift). The lift of an $AR: X \to Y$ is given by the following formula:

$$lift(X \to Y) = confidence(X \to Y)/support(Y).$$
 (2)

The lift of an AR is the ratio between the observed support and expected support, if the two itemsets X and Y are independent. If the antecedent X and the consequent Y are independent, then lift = 1 otherwise lift> 1.

3.5 Discovering Frequent patterns and Inter-Appliance Associations

To extract all association rules between the devices, an algorithm must be implemented. The input of the algorithm is the database with all transactions as well as S_{min} and C_{min} . The output of the algorithm will then be all the association rules of support $S \ge S_{min}$ and confidence $C \ge C_{min}$. To optimize the algorithm it is common practice to compose the problem into two sub-problems:

First sub-problem:

• Find all frequent itemsets. Let F be their set.

Second sub-problem:

• Determine, from F, all the association rules of associations of sufficient confidence.

3.5.1 finding frequent items

The Algorithm 1 (Brute Force Algorithm) is a naive algorithm to find all frequent items.

Algorithm 1 Find all frequent items

```
Require: TDB, I: Set of items, s_{min}.

1: F \leftarrow All subset of I.

2: F_{frequent} \leftarrow \emptyset.

3: for each itemset \in \mathcal{F} do

4: Support_{itemset} \leftarrow ComputeSuport(itemset).

5: if Support_{itmset} \geq s_{min} then

6: add itemset to F_{frequent}.

7: end if

8: end for
```

The previous algorithm is inefficient in practice because of its high complexity which is $O(2^{|I|})$. To resolve this problem, Agrawal and his co-authors in 1994 [17] propose Apriori algorithm for finding all frequent itemset in the database. It is a basic algorithm that allows extracting all frequent itemsets in a large TDB (having several thousands transaction for searching frequent patterns of attributes and several million records). It uses an intelligent enumeration of itemsets based on the decay of the support function. The algorithm pseudo-code is presented in Algorithm 2.

Algorithm 2 Apriori algorithm

```
Require: TDB, F_i: Set of items, s_{min}.

1: i \leftarrow 1.

2: C_1 \leftarrow set of 1-itemsets.

3: while C_i \neq \emptyset do

4: Compute the support of each itemset X \in C_i in D.

5: F_i \leftarrow \{X \in C_i | support(X) \ge s_{min}\}.

6: C_{i+1} \leftarrow all possible combinations of (i+1)- itemsets of F_i.

7: i \leftarrow i + 1.

8: end while

9: Return the union of sets F_i.
```

Apriori algorithm [17] with candidate generation can suffer from the following problems:

- Breadth-first approach; i.e., level-wise search.
- Generate a large number of candidate sets.
- Repeatedly search through the entire database to find support for an itemset.

To solve this problem, the work in [18] proposes a pattern growth or FP-growth approach. The FP growth algorithm finds frequent patterns without candidates generation. The main advantages of FP Growth Algorithm are the following:

- This algorithm needs to scan the database only twice when compared to Apriori which scans the transactions for each iteration.
- The pairing of items is not done in this algorithm and this makes it faster.
- The database is stored in a compact version in memory.
- It is efficient and scalable for mining both long and short frequent patterns.

The first step of the algorithm is FP-tree construction as presented in Algorithm 3. The second step is to generate frequent patterns from this FP-tree as presented in Algorithm 4.

Algorithm 3 FP-tree construction

```
Require: TDB D,I,s_{min},FP-Tree T with root labeled null
 1: F_{frequent} \leftarrow \emptyset.
 2: C_1 \leftarrow \text{set of 1-itemsets}.
 3: for all item in I do
         Support(item) \leftarrow ComputeSuport(item).
 4:
 5:
         if Support(item) \ge s_{min} then
             Add item to F_{frequent}.
 6:
         end if
 7:
 8: end for
 9: Sort F in support-descending order
10: Sort all transaction in TDB according order as in F_{frequent}.
11: for all Transaction [p|P] in TDB (p is the first element and P is the remaining list do
         Call InsertTree(p|P,T).
12:
13: end for
```

Algorithm 4 InsertTree(p|P,T)

```
      Require:
      TDB,I,S<sub>min</sub>,FP-Tree T with root labeled null.

      1:
      for all item in p|P do

      2:
      if T has a child N such that N.item-name = p.item-name then.

      3:
      N.count ← N.count + 1.

      4:
      else

      5:
      Create a new node N.

      6:
      N.count ← 1.

      7:
      end if

      8:
      end for
```

3.5.2 finding frequent patterns via incremental extraction

The Algorithms 3 and 4 require a huge memory space to store the frequent patterns in the database. To overcome this problem, the authors in [19] propose an incremental procedure to

determine the frequent patterns in a bottom-up way. The main idea is to divide the complete database into 24-sub-database of size 24 itemSet (here the time granulation is one hour). After this step, the FP-grouth algorithm will be applied to the first 24 sub-databases to determine the frequent items. This procedure will be incremented for the next 24-sub database.

The construction of FP-Tree is done in two steps: The first step is to scan all transactions in the 24-sub-database and determine the support of each item. Then, the procedure will sort all items and transactions based on their support value in ascending order. The item with the highest support value will then be placed at the beginning and so on for the others. The second step is to go through all the transactions in each 24-sub-database and determine for each transaction if all the items are accessible via the roots of the current tree. If so, the support of each item visited is incremented by 1. In the opposite case (i.e. an item inaccessible via the current tree) we build a new root. The second step will be repeated for all the transactions of the 24-sub-database.

3.6 Appliances' Utility.

The objective of this phase is to study Appliances' Utility in each building. The utility of each device will be calculated for each fraction of time (here the fraction chosen is one hour). During each fraction of time, the device will be considered active with the value 1 and 0 otherwise. To calculate the utility value of each device, we use the following formula [20]:

$$U_a = \frac{\sum_{h=1}^{24} IU_a^h}{\sum_{i=1}^{N} \sum_{h=1}^{24} IU_i^h},\tag{3}$$

where IU_a^h is the Internal Utility and is calculating as follow:

$$IU_a^h = \sum_{i=1}^n S_{i,a}^h, (4)$$

where:

$$S_{i,a}^{h} = \begin{cases} 1, & \text{If Appliance is active in hour } h. \\ 0, & \text{Otherwise.} \end{cases}$$

4 Results and discussion

In this section, the algorithms presented in the previous sections will be implemented on the MORED dataset to discover the different frequent patterns and Inter-Appliance Association. The obtained result allows us to understand the behavior of a Moroccan occupant in terms of his energy consumption and thus the total consumption in the grid. The results will be presented for 2 types of buildings and 2 houses in total.

4.1 Frequent patterns and rules associations

Table 2 presents the frequent patterns for House 5 and House 1 in the Apartment in the Affluent Neighborhood (AFN) and Apartment in a Disadvantageous Neighborhood (ADN) respectively. The results obtained confirm that the use of a TV in building 5 is the most frequent with a percentage of 76% then water heater with a percentage of 20.4%. For the patterns of a size greater than 1 we note that the occupant uses a water heater and hairdryer at the same time with a percentage of 1.3%. For House 1 we can see that the use of the Laptop

ItemSet Support Sr. Size TV monitor 1 1 0.760 2 Water heater 1 0.204 House 5 AAFN 3 Oven 1 0.051 4 Hair straightner 1 0.040 5 Hair dryer 1 0.037 6 TV monitor, Water heater 2 0.037 Washing machine 1 0.016 8 Water heater, Hair dryer 2 0.013 9 2 TV monitor, Oven 0.012 2 10 TV monitor, Hair straightner 0.011 2 11 TV monitor, Hair dryer 0.010 2 12 TV monitor, Washing machine 0.013 2 13 Water heater, Hair straightner 0.008 Hair straightner, Hair dryer 0.008 Laptop ASUS 1 0.587 1 2 Laptop MSI 1 0.565 3 Laptop ASUS, Laptop MSI 2 0.178 4 Tv monitor 1 0.070 5 2 Laptop MSI, TV monitor 0.035 2 6 Laptop ASUS,TV monitor 0.029

Table 2: Frequent patterns in House 5 and House.

ASUS is the most frequent with a percentage of 58.7% and the Laptop MSI with a percentage of 56.5%. The use of other devices remains negligible with a percentage lower than 0.7%.

3

1

0.014

0.005

0.004

Laptop ASUS, Laptop MSI, TV monitor

Washing machine

Panini grill

To understand the behavior of the occupant in each building, it is convenient to draw the rules association between the different appliances for each building. Table 3 presents the different association rules and their support and confidence. As we see, the occupant turns on the TV before using any appliance (rules 2,7,6,10,12) which confirms the result obtained in Table 2. For building 1, the association rules between the AUSI Laptop and the MSI Laptop is higher with confidence of 30% which means that the occupant frequently uses both laptops at the same time.

Appliances' utility are derived from the frequency of usage of the respective appliance. As shown in the table 4, the TV has the most utility value in house 5 with 68% then water heater with 18.27%. Laptop MSI and ASUS have an almost equal utility value because they are used at the same time most of the time.

4.2 Energy consumption and Peak Hours

8

The data set generated above allows us to determine the energy consumption of each device for each unit of time of one hour. The objective of this section is to summarize this dataset into another one where each column presents the hours of a day and the rows are the different days that are presented in the event file. For example Building 1 in AFNH contains 33 days. The next step is to calculate for each hour in the interval [1,24] the frequency where this hour

Table 3: Association rules and their support and confidence in House 5 and House 3.

	Sr.	Association Rule	Support	Confidence	Lift
	1	Water heater⇒TV monitor	0.037407	0.183421	0.241502
Z	2	TV monitor⇒Water heater	0.037407	0.049252	0.241502
AF	3	Hair dryer⇒Water heater	0.013281	0.355631	1.743780
House 5 AAFN	4	Water heater⇒Hair dryer	0.013281	0.065120	1.743780
e S	5	Washing machine⇒TV monitor	0.012525	0.786813	1.035957
sno	6	TV monitor⇒Washing machine	0.012525	0.016491	1.035957
H	7	TV monitor⇒Oven	0.012497	0.016454	0.323057
	8	Oven⇒TV monitor	0.012497	0.245363	0.323057
	9	Hair straightner⇒TV monitor	0.010720	0.269197	0.354438
	10	TV monitor⇒Hair straightner	0.010720	0.014114	0.354438
	11	Hair dryer⇒TV monitor	0.010188	0.272812	0.359198
	12	TV monitor⇒Hair dryer	0.010188	0.013414	0.359198
	1	Laptop ASUS⇒Laptop MSI	0.178363	0.304113	0.538526
1 ADN	2	Laptop MSI⇒Laptop ASUS	0.178363	0.315847	0.538526
	3	TV monitor⇒Laptop MSI	0.035370	0.503527	0.891651
	4	Laptop MSI⇒TV monitor	0.035370	0.062634	0.891651
House 1	5	TV monitor⇒Laptop ASUS	0.029230	0.416119	0.709492
noj	6	Laptop ASUS⇒TV monitor	0.029230	0.049838	0.709492
工	7	TV monitor⇒Laptop MSI	0.013691	0.468391	0.829430
	8	TV monitor⇒Laptop ASUS	0.013691	0.387081	0.659983
	9	Laptop ASUS⇒TV monitor	0.013691	0.076760	1.092750
	10	TV monitor⇒Laptop ASUS	0.013691	0.194906	1.092750
	11	Laptop ASUS⇒TV monitor	0.013691	0.023344	0.659983
	12	Laptop MSI⇒TV monitor	0.013691	0.024244	0.829430

Table 4: Appliances' Utility.

	Sr.	Appliances	Appliances' Utility(%)
\overline{z}	1	TV monitor	68
AF	2	Water heater	18.27
Ą	3	Hair dryer	3.59
e 5	4	Hair straightener	3.83
House 5 AAFN	5	Oven	4.18
H	6	Washing machine	2.13
	1	Laptop MSI	45.59
-	2	Laptop ASUS	48.09
	3	Washing machine	0.17
l A	4	TV monitor	5.76
House 1 ADN	5	Panini grill	0.31
ons	6	Hair dryer	0.06
H	7	Blender	0.01

has been selected as the peak hour for each day. The peak hours are the hours with the highest frequency. The Tables 5, 6 show the first six peak hours for both houses 5 and 3. The means

energy consumption for each peak is presented in Tables 7 and 8 respectively. For house 5, we notice that the hours 15 a.m, 10 p.m is the peak hours with a frequency of 4 times. Table 7 informs us that the appliance with the highest energy consumption for each peak is the water heater. For house 1, we can see that 18 a.m and 7 p.m are the peak times and the MSI and ASUS Laptops are the devices responsible for this peak, which confirms the results obtained previously.

Table 5: House 5: Peak Hour fre-Table 6: House 1: Energy Consumpquency.

Sr	Hour	Frequency
1	15	4
2	10	4
3	12	3
4	3	3
5	22	3
6	19	3

Sr	Hour	Frequency
1	18	4
2	19	2
3	12	1
4	14	1
5	21	1
6	8	1

Table 7: House 5:Energy consumption per Peak Hour.

Sr	Hour	HD	HS	OV	TV	WM	WH
1	15	2.86	1.62	24.62	4.12	4.94	61.85
2	10	3.21	1.73	18.94	14.14	6.65	55.33
3	12	3.01	1.78	0.98	8.23	6.59	79.41
4	3	5.00	2.14	52.11	13.24	7.93	19.57
5	22	1.67	0.62	87.09	7.12	3.21	0.29
6	19	6.24	2.05	32.28	24.48	8.23	26.73

Table 8: House 1:Energy consumption per Peak Hour.

Sr	Hour	BL	HD	ASU	MSI	PG	TV	WM
1	18	1.35	3.71	13.49	45.92	4.84	7.48	23.20
2	19	1.33	3.31	0.21	63.75	4.42	3.34	23.63
3	12	0.99	5.03	21.10	50.56	2.65	2.01	17.65
4	14	1.39	3.15	11.01	57.26	4.05	2.80	20.34
5	21	1.60	3.73	48.67	10.68	4.89	4.31	26.12
6	8	1.10	3.27	12.23	0.12	62.01	1.89	19.38

To complete our energy analysis, the Figures 3,4 show the duration of use of each machine for the two houses 1 and 5. The duration of use of a device does not have a direct influence on the total energy consumption during the peak hours because a device is used frequently but its power is small and subsequently its energy consumption is low. For example, the TV is the most used device in house 5 if we compare it with the Water heater but it is responsible for the peak hours due to its high power.

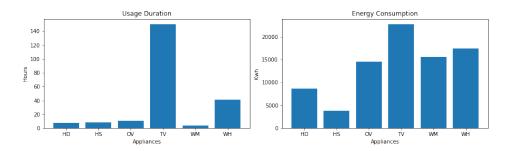


Figure 3: House 5: Energy consumption and Usage duration.

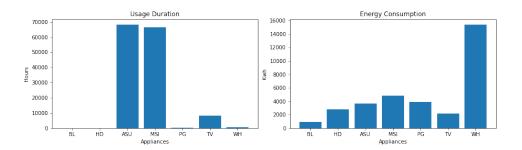


Figure 4: House 1: Energy consumption and Usage duration.

In Figures 6 and 5 we show the peak hours for each house and the percentage of consumption of each machine during each peak hour. For house 5 we notice that the Water Heater is always the device that consumes a lot of energy and consequently increases the customer's bill. For house 1, the MSI laptop washing machine are the most energy-consuming devices. These analysis results help to guide the development of an Energy Management System to control the devices to minimize the energy consumption and subsequently minimize the load on the grid.

5 Conclusions and future work

In this paper, we present Data Mining Techniques to detect different items in a Moroccan energy consumption database. These techniques allow to study and understand the behavior of a Moroccan consumer at the level of energy consumption and to determine the preferences of each occupant for two types of houses. The result and analyses were carried out to allow other researchers to develop an Energy Management System to optimize and pilot the energy consumption and subsequently minimize the bills and load in the grid.

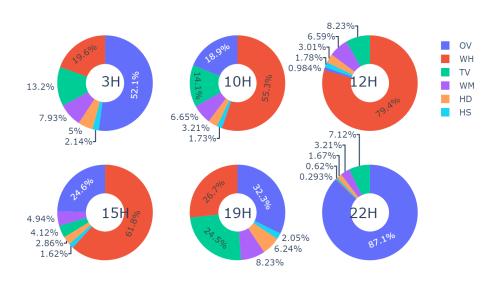


Figure 5: House 5: Energy consumption Per peak Hour.

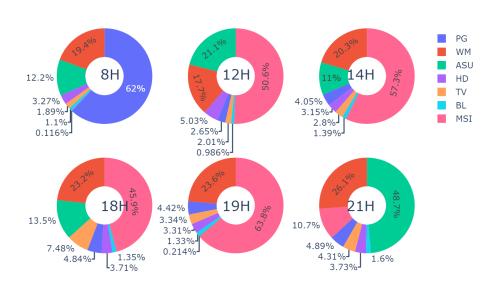


Figure 6: House 1: Energy consumption Per peak Hour.

Bibliography

- [1] M. LLC, Energy Policies beyond IEA Countries: Morocco 2019 (2019), https://www.iea.org/
- [2] J. Page, D. Robinson, N. Morel, J.L. Scartezzini, *A generalised stochastic model for the simulation of occupant presence* (2008), Vol. 40, pp. 83–98, ISSN 0378-7788
- [3] G. Singla, D. Cook, M. Schmitter-Edgecombe, *Recognizing Independent and Joint Activities Among Multiple Residents in Smart Environments* (2010), Vol. 1, pp. 57–63
- [4] L. ORTA, Réalisation d'un logiciel pour l'optimisation énergétique du bâtiment (2016), p. 48
- [5] F. Pallonetto, M. De Rosa, F. Milano, D.P. Finn, *Demand response algorithms for smart-grid ready residential buildings using machine learning models* (2019), Vol. 239, pp. 1265–1282, ISSN 0306-2619
- [6] Y. Peng, A. Rysanek, Z. Nagy, A. Schlüter, Using machine learning techniques for occupancy-prediction-based cooling control in office buildings (2018), Vol. 211, pp. 1343–1358, ISSN 0306-2619
- [7] C. Fan, F. Xiao, Y. Zhao, A short-term building cooling load prediction method using deep learning algorithms (2017), Vol. 195, pp. 222–233, ISSN 0306-2619
- [8] J.L. Gomez Ortega, L. Han, N. Whittacker, N. Bowring, A machine-learning based approach to model user occupancy and activity patterns for energy saving in buildings, in 2015 Science and Information Conference (SAI) (2015), pp. 474–482
- [9] T. Yu, Modeling Occupancy Behavior for Energy Efficiency and Occupants Comfort Management in Intelligent Buildings, in 2010 Ninth International Conference on Machine Learning and Applications (2010), pp. 726–731
- [10] B. Dong, B. Andrews, K.P. Lam, M. Höynck, R. Zhang, Y.S. Chiou, D. Benitez, An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network (2010), Vol. 42, pp. 1038–1046, ISSN 0378-7788
- [11] T. Ekwevugbe, N. Brown, V. Pakka, D. Fan, Real-time building occupancy sensing using neural-network based sensor network, in 2013 7th IEEE International Conference on Digital Ecosystems and Technologies (DEST) (2013), pp. 114–119
- [12] N. Li, G. Calis, B. Becerik-Gerber, Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations (2012), Vol. 24, pp. 89–99, ISSN 0926-5805
- [13] H. Hajj, W. El-Hajj, M. Dabbagh, T.R. Arabi, An Algorithm-Centric Energy-Aware Design Methodology (2014), Vol. 22, pp. 2431–2435
- [14] S. Mamidi, Y.H. Chang, R. Maheswaran, *Improving Building Energy Efficiency with a Network of Sensing, Learning and Prediction Agents* (2012)
- [15] M.A. Ahajjam, D. Bonilla Licea, C. Essayeh, M. Ghogho, A. Kobbane, MORED: A Moroccan Buildings' Electricity Consumption Dataset (2020), Vol. 13, ISSN 1996-1073
- [16] J. Han, J. Pei, M. Kamber, *Data Mining: Concepts and Techniques* (Elsevier Science, 2011), The Morgan Kaufmann Series in Data Management Systems, ISBN 9780123814807, https://books.google.co.ma/books?id=pQws07tdpjoC
- [17] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, A.I. Verkamo et al., Fast discovery of association rules. (AAAI/MIT Press Menlo Park, CA, 1996), Vol. 12, pp. 307–328
- [18] J. Han, J. Pei, Y. Yin, Mining Frequent Patterns without Candidate Generation (Association for Computing Machinery, New York, NY, USA, 2000), Vol. 29, p. 1–12, ISSN 0163-5808

- [19] S. Singh, A. Yassine, Mining Energy Consumption Behavior Patterns for Households in Smart Grid (2019), Vol. 7, pp. 404–419
- [20] S. Osama, M. Alfonse, A.B. M. Salem, *Mining Temporal Patterns to Discover Inter-Appliance Associations Using Smart Meter Data* (2019), Vol. 3, ISSN 2504-2289, https://www.mdpi.com/2504-2289/3/2/20