Test set validation for home electrical signal disaggregation

Luc Dufour, Dominique Genoud, Gianluca Rizzo, Antonio J. Jara [1],
Pierre Roduit [2]
Jean-Jacques Bezian and Bruno Ladevie [3]

1 Institute of Information Systems, University of Applied Sciences Western Switzerland (HES-SO), Sierre, Switzerland
2 Institute of Industrials Systems, University of Applied Sciences Western Switzerland (HES-SO), Sion, Switzerland
3 Mines Telecom, Albi, France
Email: jara@ieee.org, (Luc Dufour, Dominique Genoud, Gianluca Rizzo, Pierre Roduit)@hevs.ch,
(Jean-Jacques Bezian and Bruno Ladevie)@mines-albi.fr

***Abstract*—In order to enable demand response schemes for residential and industrial users, and to get power consumers involved in the general effort to save energy and decrease the carbon footprint of human activities, it is crucial to be able to monitor each component of the total power consumption of a household or of an industrial site over time. A cost effective and readily available solution to obtain these data is to exploit Non-Intrusive Load Monitoring (NILM) techniques over low frequency smart meter readings. Such methods have still important limitations in the detection of some types of appliances. For that reason, this work presents a new non intrusive load disaggregation method for residential households, based on low frequency measurements of active and reactive power. Our method allows detecting some families of common household appliances which were undetectable through low frequency NILM methods proposed so far. The proposed method has been validated on an experimental setting and using direct measurements of appliances consumption, proving that it allows achieving a high level of accuracy in load disaggregation.**

I. Introduction

The last decade has been characterized by a sharp increase in fuel prices and in energy demand. At the same time, energy consumption awareness has been raised over the impact which the greenhouse effect is having on the climate and on world economy. These factors have driven research of solutions for sustainability in energy production, distribution, storage, and consumption. In energy distribution, new "smart" solutions have been proposed, centered on the idea that exploiting properly data on power generation, distribution and consumption, a substantial increase in efficiency is achievable in power production and distribution.
At the user level, several load management and demand response techniques have been proposed, aiming at flattening the peaks of power consumption over time, and at adapting demand to variations in supply due to renewable. These strategies imply, on behalf of consumers, awareness of the amount of energy consumed by each device and of its relative impact on the total energy bill of the household. Such knowledge, besides enabling users to respond appropriately to power price variations over time, allows also identifying inefficiencies and decreasing the overall consumption, helping them to contribute to the decrease of their carbon footprint.
The electric consumption increased 120% in 50 years [1]. The objectives in Europe are the 3* 20 which includes a part of renewable energy in 20%. So, we have more and more a photovoltaic panel but their production peaks engage problems for the distributors and producers of electricity. One possible answer is to increase the auto-consumption but the problem is to pilot the different devices in function of weather, scenarios, identification of powered-on devices, cost of electricity, storage...It's to have a good prediction for the next 15-minutes and to pilot devices in order to increase auto-consumption in our microgrid.
Among the techniques proposed to monitor appliance consumption, non-intrusive ones (called also NILM [7]) are of particular interest in households, since they do not require specialized, costly hardware and installation and maintenance of a sensor network. Moreover, they adapt over time in changes in households (such as changes in appliance number and type) without requiring new installations or reconfiguration of existing hardware and software. The extensive deployment of smart meters which is planned in many countries for the near future will enable a large scale deployment of NILM techniques [2]. Such deployment will make available measurements of the total active and reactive power consumed, typically sampled at low frequencies, allowing non-intrusive load monitoring without the use of additional hardware.

NILM methods have been first proposed in [3], and they are typically structured in three phases: feature extraction, events detection, and events classification. They make use of a database of electric signatures of appliances, and they are based on the measure of the total active power consumed, sampled at frequency of one Hertz. Later methods [4] [5] try to decrease the duration of the training period. Indeed, the main drawbacks of these techniques reside either in the need of a learning phase requiring intrusive measurements, and/or on the fact that they cannot detect appliances whose power consumption patterns vary drastically over time (e.g. washing machines, whose power consumption varies substantially during a washing cycle). Indeed, a fine granularity and a good accuracy in load disaggregation are crucial in order to enable useful feedback to users, to set up appropriate measures for changing consumption patterns, and to enable detection of anomalies and appliance malfunctioning. Many of the techniques proposed in order to overcome these drawbacks imply a substantially higher sampling frequency, and therefore expensive hardware [10] [9] [11].

Our work shows how is it possible, with a relatively simple setup and inexpensive, readily available hardware, to verify the started use of electrical appliances. In this paper, three appliances were selected to verify the model: heat pumps, refrigerators and freezers, and dishwasher. We have a verification around 94%.
We chose these appliances because there power can be sifted for a predetermined time. The peak consumption can be move 15-minute for the cold and one hour for the heat pumps without major impact on the users comfort. We remove in a first time the detection of washing machine and dryer : their usage are planned in the Swiss building and used practically every time.There are one washing machine and one dryer for ten apartments in a Swiss building.
The paper is organized as follows. In Section II we present our method, and detail the algorithm for load disaggregation. In Section III we describe the setting we used for the validation of our method, and in Section IV we present the results of the different algorithms. Finally we conclude and discuss future directions of research in Section V.

II. Verification method

In this section, we describe the data required for our method and the prediction model to verify the started devices in use the global power curve.
Cross-validation is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. It is worth highlighting that in a prediction problem, a model is usually given a dataset of known data on which training is run (training dataset), and a dataset of unknown data against which the model is tested (testing dataset).
The goal of cross validation is to define a dataset to ”test” the model in the training phase in order to limit problems like overfitting, give an insight on how the model will generalize to an independent data set.
One round of cross-validation involves partitioning a sample of data into complementary subsets, performing the analysis on one subset, the training set, and validating the analysis on the other subset, the testing set. To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the rounds.

In fact, there are at least three techniques of crossed validation: test set validation, k-fold cross-validation and leave-one-out cross validation.In this paper, we begin in used the test set validation. We partitioning the sample of size n into training dataset (¿ 60%) and a testing set. The model is built on the training set and validated on the testing set. The error is estimated by calculating a test, a measure or a score of performance of the model on the testing set, for example the error quadratic average.In our case, we define a contingency matrix representing the possible outcomes of the classification. We interpret this matrix like a good or bad prediction for each algorithms and for each devices.
Our method requires data for every second for the voltage, the amperage, active power,reactive power for each phase. This allows us to understand the dependency of the target variable to the input vectors. The target variable are the difference between two active power points and in a second part,the difference between two reactive power points.

Feature extraction: We have in inputs three values every second for voltage, three for amperage, three for active and reactive power, three for the power factor and the accumulated active energy and a time reference for the global load curve. We have too in inputs the active power and the time reference for the different devices connected inside the house. We use edge detector method , which consists on detecting variations of active and reactive power with regards to a predefined threshold [3]. The different work on this subject and our analysis working enable to know the different characteristics to describe a device [9], [10] [11].
We begin by filtering the interferences that we must identify on the global diagrams as well as on individual connected appliances diagrams.To reduce the effects of noise and transients on the measured data, we apply a median filter of size 60 to the time series of measured values.We have empirically determined the optimal size of the filter via a set of tests, by attempting to balance the elimination of spurious load level variations and the performance of the event detection algorithm, in terms of average number of undetected events.
We use a median filter on each entry signal to remove high frequencies. This is used on the global load curve for each phase and on the different devices’ signals. The threshold used depends mainly on the type of device that we want to detect.If the objective is to detect all devices, the active threshold must be the finest possible, but the variations detected can lead to errors corresponding to residual noise.
We have now a list of variations of active and reactive powers variations by phases which represent the different devices connected in the test house. A time level recognition is realized to group the potential vectors (deltaP / deltaQ) for the three observed devices.

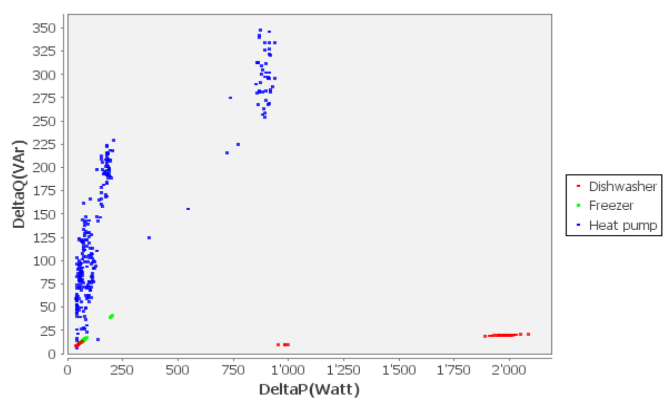

Fig. 1. Active and Reactive power vectors detected on phase 1 for three devices, (Dishwasher,Freezer and Heat pump),11 February 2013 - 17 February 2013

Predictors: To have a good indicator, it’s crucial not to use part of the training data to test the method. We use a random sampling of all active/reactive power vectors for the partitioning: the input table is split into two partitions,train and test data. To estimate the training data, we use three known classification algorithms : k- Nearest Neighbor (KNN) [12], Support Vector Machine (SVM) [13], and a decision tree [14]. The nearest neighbor algorithm derives the class labels or the numeric target values of new input objects from the most similar training examples, where similarity is measured by distance in the feature space. The prediction is computed by a majority vote of the nearest neighbors or by averaging their numeric target values.

The number k of neighbors to be taken into account is a parameter of the algorithm.The best choice of which depends on the data and the prediction task. The kNN node allows setting the number of neighbors to be considered and if the distance should be used to weigh in on those neighbors. Furthermore, The kNN is very sensitive to the chosen distance function so we should make sure to normalize the data and use the exact same normalization procedure for both the training and test data.This can be achieved by using the normalizer function and Normalizer (Apply) function , which copies the settings from the first node. We then feed those two data tables into the K Nearest Neighbor node which adds a column with the predicted class to the test data.

Support vector machine algorithms have not been around as long as neural networks but in the meantime, many implementations of support vector machines (SVM) have shown up in commercial and open software packages. Our analysis consists of two steps, one learning the SVM model offering the choice of a few well known kernels and the second one allowing to apply the model to a second data set. The native SVM implementation currently offers three different kernels functions: a polynomial, hyper tangent and RBF kernel. In contrast to e.g. decision trees, kernel functions (or at least the settings of their respective parameters) are rather sensitive to the normalization of the input data. For instance, the sigma parameter of the RBF kernels controls the width of the basis functions in relation to the Euclidean Distance of the input vectors to the support vectors. Hence it is critical to adjust this parameter accordingly.

When constructing decision trees in data mining software, a number of options are available. We need to first and foremost select the target attribute (it has to be a nominal attribute). After that,we can choose between two different ways to compute the information gain (Gini index and Gain ratio) and if a pruning of the tree is to be performed (The data mining software offers to either skip pruning or performs a minimum description length (MDL) [14] based pruning). Noteworthy is the last option number threads, which allows to control how many threads the software can use to execute the learning method in parallel on, e.g. a multi core machine. Once the node is run, we can display the resulting decision tree.

In order to evaluate the decision by scored, we define a contingency matrix representing the possible outcomes of the classification, namely the true positives (TPpositive examples classified as positive), the True Negatives (TNnegative examples classified as negative), the False Positives (FPnegative examples classified as positive) and the False Negative (FN-positive samples classified as negative). The recall is defined as TP/(TP+FN) and the precision is TP/TP+FP.

III. EXPERIMENTAL SETTING

In this section, we describe how our information system collects data, how the parameters are used in our analysis software and how the data set for prediction model.

Information System: We use the schneider electric system to collect active and reactive power by phase on the global electric meter [15]. This smart meter is the PowerLogic Series 800 PM810 of the Schneider Electrical Company [16]. As outputs, we have the amperage, voltage, active and reactive power and energy consumed with in a one second interval per phase. We use the same device to collect data from the principal Heat pump. We also have amperage, voltage, active and reactive power per second for each phase for the heat pump 1.
The different devices of houses are connected using the Geroco smart plugs. This smart meter collects the amperage, voltage, the active power and timestamps. The protocol of communication used is a Zigbee [18]. This enables piloting after a processing step in order to define activation and disactivation of one or many devices. An integrated pre-programmable code enables the recognition of variations of active and reactive powers on the global load curve for each device.

- Phase 1: Dishwasher, electric ovens, induction plates, heat pump1, heat pump 2, fridge, microwave ovens, light, Modem, kettle, coffee machine,overhead projector;
- Phase 2: Two freezer, light, heat pump 1, induction plates, electric ovens;
- Phase 3: A washing machine, a dryer, light, heat pump 1, induction plates, electric ovens;

The data travels through a modbus communication.These datas are stored on mini-pc in csv files and to send on server in HES SO at Sion. We use the open source data analysis software KNIME [18] to connect the database and process the data. At this point, we have one year of data (November 2012-November 2013) for the global charge charts and the devices connected. 10houses are today connected.

Data set: We chose a traditional Swiss house, in a mountainous area to study the possibility of energy management. In fact, this region benefits of a warm and dry climate,(approximately 2 500 hours of sunlight a year) and the house has a production of electricity by photovoltaic panels. Furthermore,the two heat pumps and the dishwasher are piloted.These devices are activated automatically at 12 am in priority one. The aim is to decrease electricity invoice because the cost of electricity is low the night between 9pm and 6 am. For our test, we chose a random week in the winter season to study the activation of two heat pumps(11 February 2013 and the 17 February 2013). We test our prediction model in a first time with the active power and in a second time with active and reactive power.

Feature extraction: We tested two threshold for the active and reactive power to 50W/15VAR and 40W/5VAR to study the impact of transitional effects on electrical signals. If we

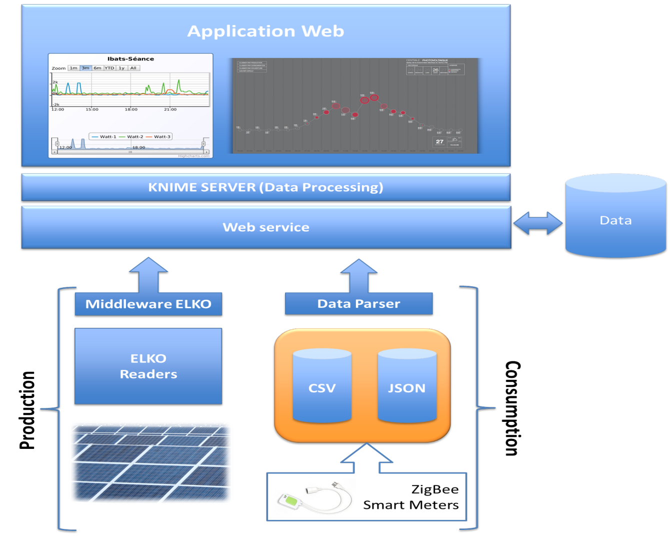

Fig. 2. I-BAT information system architecture based No-SQL databases

use a median filter with a size of 60 and an edge detector of 40W/5VAR, we detect all variations for the three devices (heat pump, refrigerator/freezer, dishwasher). The active threshold, 40Watt or 50 Watt, is sufficient to detect a dishwasher, a freezer or a heat pump on/off jumps. But the reactive power threshold is too high to detect a fridge or freezer. That is why we used a 40W and 5VAR threshold in our methodology. The devices not detected with this active power threshold are the modem and the computers.

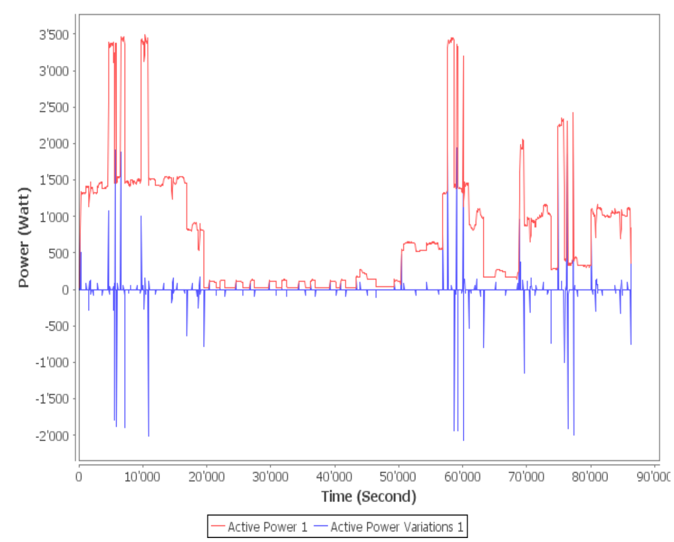

Fig. 3. Active power variations on phase 1, 15 February 2013

Predictors: We use a training set of 70 percent and a test set of 30 percent of the collected data. We chose this partitioning to balance the number of cycles studied by device. For a week of datas, we have 7 cycles for the dishwasher, 19 cycles for the heat pump and 146 cycles for the freezer.
The training set vectors are normalized by Min-Max Normalization Method. The model of normalization is applied for the test set. The three values normalized are the time reference,

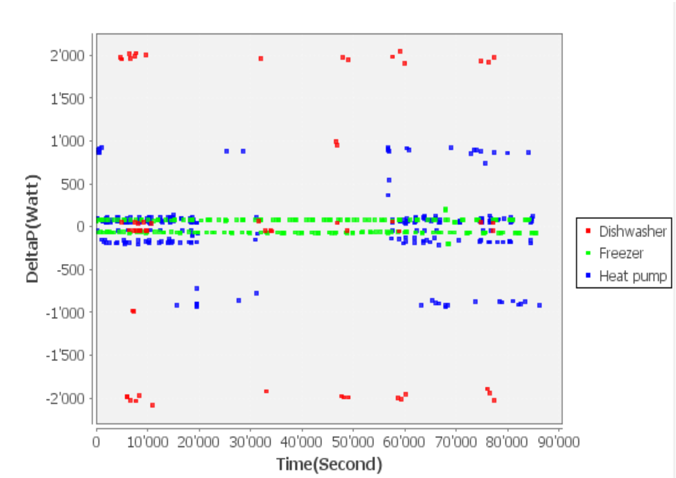

Fig. 4. Active power variations on phase 1 for the three devices study, 15 February 2013

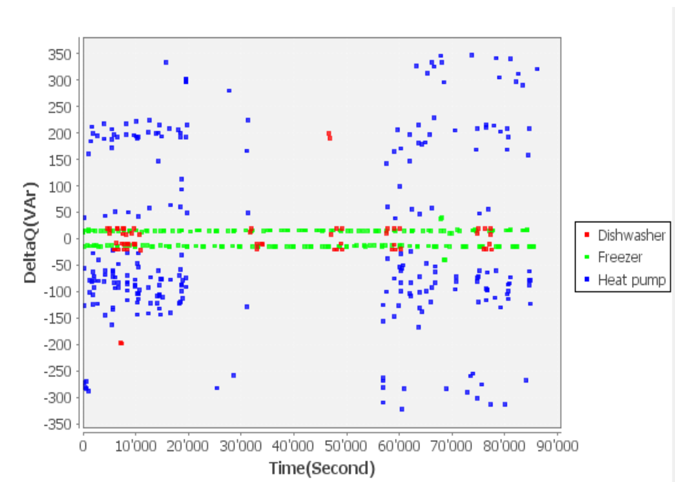

Fig. 5. Reactive power variations detected on phase 1 for the three devices study, 15 February 2013

active power and reactive power. For the kNN, we fix k = 3 and we chose the weight neighbors by distance. For the DCT, the pattern finding increase the DCT performance because three data are analyzed by split of the tree so we don't use a post pruning method like "Minimal Description Length" (MDL). The quality measures for split calculation used is the gini index. We unchecked the split value by average compute. For the SVM, we use a RBF kernel .

IV. Classification Results

In this section, we discuss our results by algorithms. The active and reactive power variations enable the identification a class of devices or a state of a device. In our case, a freezer has one state but the dishwasher and the heat pump have many states. These two devices are one or many pumps. The functioning of these pumps are the same: the event on are followed by an event off. We can consider that these events are in the same class. These pumps are started between the significant variations of active and reactive power. The intervals of the one and off jumps are significant in every device.For example, the pump of the heat pump started after the jumps of 900 watts (heat pump 1). These pumps can have different signals according to the installation-selected player and it will be interesting to separate them in future works. The different algorithms can separate these significant variations, but the difficulty is to separate the different pumps. However,

Devices	Active Power Peak	Reactive Power Peak
Heat pump 1	926.5	348
Heat pump 2	536	220
Freezer	203.7	30.7
Fridge	254.2	40.2
Dishwasher	2050.1	52.4

TABLE I. AVERAGE OF ABSOLUTE VALUE ACTIVE AND REACTIVE POWER PEAK IN WATT BY DEVICES FOR A WEEK DATA SET

3-NN	SVM	DCT
70.4 $\pm$2.3	77.8 $\pm$2	94.9 $\pm$1.5

TABLE II. AVERAGE RESULTS BY ALGORITHMS WITH REACTIVE AND ACTIVE POWER DATA SET FOR A WEEK

the power (active and reactive) of these pumps depends of the installations of the house, the type of function (for example,there are several ways of functioning the dishwasher). In our case the pumps of the dishwasher and the heat pump are more powerful than that of the freezer. Furthermore, the variations of the reactive power of the pumps of the heat pump are more powerful than those of the refrigerator and dishwasher.
Our results are suitable except for the KNN.The associated average accuracy was 70.4% . For the svm, The average accuracy is good for hot air pumps with 80% : the active/reactive power is very different in particular the reactive power to compare two others. Half of the variations of the refrigerator are confused with a heat pump. Only three variations of the refrigerator are classified with the dishwasher. For the classification of the dishwasher,the margins of error are the same for a refrigerator and a heat pump. The algorithm which obtains the accuracy with highest 94.9% is the decision tree. Under these circumstances, we can conclude that the performed methods are suitable for verification. Moreover, the composed signature can be an accurate description for each of the appliances in the dataset. Nevertheless, the small number of electrical appliances might be a limitation that should be further considered.

Devices	3-NN	SVM	DCT
Heat pump	67.9 $\pm$2.6	80.3 $\pm$1.2	96.3 $\pm$1.5
Dishwasher	80.2 $\pm$1.4	62.6 $\pm$1.8	82.3 $\pm$2.5
Freezer	72.2 $\pm$3.1	61.9 $\pm$2.2	96.2 $\pm$2.6

TABLE III. RESULTS BY ALGORITHMS AND BY DEVICES WITH REACTIVE AND ACTIVE POWER DATA SET FOR A WEEK

V. CONCLUSIONS

We propose a verification method which uses active and reactive power to differentiate three devices : heat pumps dishwasher and freezer. We have more than 92% percent of recognition on the studied devices with the Decision tree algorithm. The results are encouraging and we must focus our job on the association between the different pumps detected and a device.
We can use a clustering step or use HMM algorithm for example.Furthermore, we can focus our job on the heating and the dishwasher which remain the potential of the most importing flexible energy.
In fact, the devices are started the night but the production of electricity is summit the daytime, the hours depends of the seasons. In our case,the sunrise at 8am and sunset at 6 pm, the peak of production may be between 12am and 2pm. In general, we have no presence the day in the house when the solar production is present.

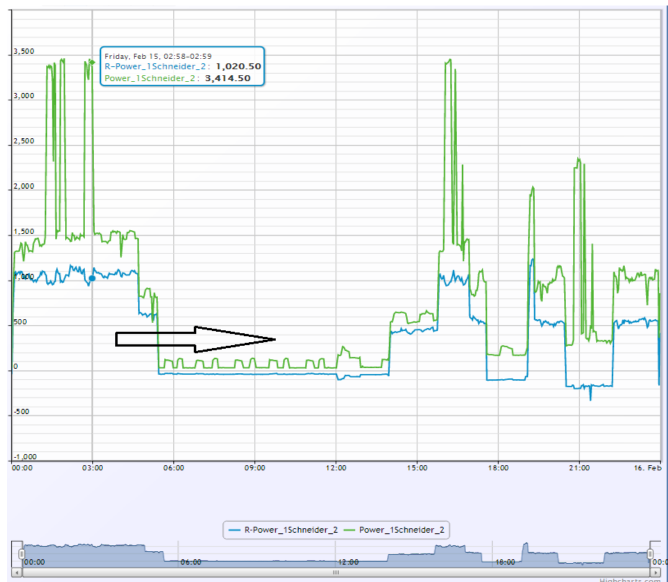

Fig. 6. Active and Reactive power by phases for the 15 February 2013 in use the schneider2 smart meter

Hours	Flexible energy	Detected Devices
00-01	2.8	Heat pump 1,2 + Cold
01-02	3.4	Heat pump 1,2 + dishwasher + Cold
02-03	3.7	Heat pump 1,2 + dishwasher + Cold
03-04	1.6	Heat pump 1,2 + Cold
04-05	1.2	Heat pump 1 + Cold

TABLE IV. TOTAL FLEXIBLE ENERGY IN KWH BETWEEN 12.PM AND 5.AM FOR THE 15 FEBRUARY 2013

For example, if we chose a random day in the data set, we notice the two heat pumps and the dishwasher, which are present during 8 hours in the day.The maximum total flexible energy is 3.4 Kwh when the two heat pumps and the dishwasher are started. In decentralized electrical production, such as in a microgrid, this potential of flexible energy is important. In your case, it's the heat pump 1 presents on the three phases which consume most energy with 1.6KWh. Next, it's the heat pump 2 and the dishwasher. The cold (freezer and fridge) are small consumptions around 40KWh.

VI. ACKNOWLEDGEMENTS

This research was undertaken as part of the I-BAT project and has been funded by The Ark Energy (Project No. 712-07). We would liketo thank all the contributions from the partners involved our colleagues from HES-SO (Pierre Roduit, Milene Fauquex, Pascal Pitteloud, Pierre Roduit, Rene Vuilleumier, and Simon Martin), Icare Institute (Pierre Kenzelman,Francois Chabbey), Alro Communication SA (Frederic Revaz),Geroco SA (Eric Nussbaum), Sierre Energie (Laurent Antille, Jean-Marc Salamin, and Christophe Darioly),Centre Suisse d'Electronique et Microtechnique - CSEM (Andreas Hutter, and Philippe Renevey), and Centre de Recherches Energtiques et Municipales - CREM (Thierry Bernhard, and Gaetan Cherix). Finally, we would like to thank the STREP European Projects "Universal Integration of the Internet of Things through an IPv6-based Service Oriented Architecture enabling heterogeneous components interoperability (IoT6)" from the FP7 with the grant agreement no: 288445.

REFERENCES

[1] IEEE (2012) OFEN, Statistique suisse de l'lectricit en 2012. *The Edison Foundation Institute for Electric Efficiency*. URL = http://www.strom.ch/fr/dossiers/graphiques-electricite.html.

[2] IEEE (2006) Utility-scale smart meter deployments, plans and proposals. *The Edison Foundation Institute for Electric Efficiency*.

[3] Hart, G.W. (1992) Nonintrusive appliance load monitoring. *Proceedings of the IEEE*. vol. 80, no. 12, pp. 1870-1891.

[4] Farinaccio, L. (1999) The disaggregation of whole-house electric load into the major end-uses using a rule-based pattern recognition algorithm. *Concordia University*, PhD Thesis.

[5] Marceau, M.L., Zmeureanu, R. (2000) Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. *Energy Conversion and Management*. vol. 41, no. 13, pp. 1389-1403.

[6] Luan, S. W., Teng, J. H., Chan, S. Y., Hwang, L. C. (2009). Development of a smart power meter for AMI based on ZigBee communication. *In Power Electronics and Drive Systems, PEDS 2009. International Conference on*. pp. 661-665, IEEE.

[7] Zoha, A., Gluhak, A., Imran, M.A., Rajasegarar, S. (2012) Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. *Sensors*. vol. 12, no. 12, pp. 16838-16866.

[8] Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M., Patel, S. (2011). Disaggregated end-use energy sensing for the smart grid. *Pervasive Computing, IEEE*. Vol. 10, no. 1, pp. 28-39.

[9] Patel, S., Robertson, T., Kientz, J., Reynolds, M., Abowd, G. (2007) At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line. *UbiComp 2007: Ubiquitous Computing. Lecture Notes in Computer Science*. Vol. 4717, Springer Berlin Heidelberg, pp. 271-288.

[10] Srinivasan, D. and Ng, W. S. and Liew, A.C. (2006) : Neural-network-based signature recognition for harmonic source identification, Vol. 21,no. 1,pp. 398-405.

[11] El Guedri, Mabrouka (2009) Caractérisation aveugle de la courbe de charge électrique: Détection, classification et estimation des usages dans les secteurs résidentiel et tertiaire

[12] T. Cover,P. Hart : Nearest Neighbor Pattern Classification(1967), journal=The Edison Foundation Institute for Electric Efficiency,*The Edison Foundation Institute for Electric Efficiency* pp. 21-27.

[13] Fast Training of Support Vector Machines using Sequential Minimal Optimization, URL = http://research.microsoft.com/en-us/um/people/jplatt/smo-book.pdf.

[14] John Shafer, Rakesh Agrawal, Manish Mehta,SPRINT, A scalable parallel classifier for data URL = http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.104.152rep=rep1type=pdf.

[15] title = Power meter series 800 data sheet,URL = http://www2.schneider-electric.com/resources/sites/ SCHNEIDER ELECTRIC/content/live/FAQS/140000/FA140345/enUS/PM80020DataSheet.pdf. ,

[16] Power meter series 800 user guide,URL = http://download.schneider-electric.com/files?p File Id= 27600253p File Name=63230-500-225A2 PM800 User Guide EN.pdf.

[17] Knime,URL = http://knime.com/.

[18] Ecowizz,URL = https://www.ecowizz.net/.