A Supervised Learning Approach to Appliance Classification Based on Power Consumption Traces Analysis

J R Cuñado¹ and N B Linsangan²
¹ Department of Engineering Education, University of Mindanao – Tagum, Visayan Village, Tagum City, Davao del Norte 8100, Philippines
² School of EECE, Mapua University, 658 Muralla St., Intramuros, Manila 1002, Philippines

Abstract. Electric appliances are everyday major power consumers. Management and control of these appliances can only be possible with appliance classification infrastructures. An appliance classification smart meter, with a provision for remote control, is developed based on time-dependent power features drawn by an appliance, from power-up to its steady state. The kNN classifier is highly accurate at 95.55% in classifying the appliance class.

1. Introduction
The role that electric appliances play in enhancing the living styles of consumers can no longer be discounted at this time and age. With these devices central to the increasing everyday power consumption in both homes and offices, power consumption management is of growing interest to reduce electricity bills. Also, instances of neglect to turn off appliance power which instigated fire had been reported several times. The need, therefore, to provide a means to connect and control electric appliances remotely is seen to be of timely importance.

Appliance Load Monitoring (ALM) has been the subject of recent researches. Plug load identification is tricky, with each appliance distinct, depending on the model, type and brand [1]. Two existing approaches are used in identifying appliances based on power consumption traces: Non-Intrusive Load Monitoring (NILM) and Intrusive Load Monitoring (ILM). Hart introduced the NILM approach where a meter-extension collar containing Non-Intrusive Load Monitor is used to check for load signatures to infer appliances’ activities. He based his unique appliance signature on the real and reactive power consumed by each appliance in steady state [2]. With NILM, aggregate electricity consumption is measured at a central point using a single sensor; disaggregation is then done to determine the contribution of individual appliance. The use of decentralized meters to measure power consumption at appliance level is a relatively new research interest, with researchers looking for better ways to manage and control household devices. Ito was one of the first researchers who undertook a study in this field by matching the feature parameters of current of the appliance to the feature parameters stored in a database [3]. Reinhardt improved on Ito’s work by extracting current-based and harmonics features from medium to higher frequency data [4]. High-frequency measurements of voltage and current can be used to build a coarse-grained model to detect attached appliance type and a fine-grained model to differentiate between products of an appliance type [5]. However, the use of medium- to high-frequency sampling is resource demanding, the reason why other researchers sampled data at low frequency using low-end sensors with promising results in appliance identification. Ridi further added a User Interaction Layer in the context of appliance monitoring [6]. Machine learning approaches have been used for appliance identification, and most are based on supervised learning techniques [7]. Hidden Markov Models outperformed Dynamic Time Warping on
a 6-category task [8]. One-class Support Vector Machines with Gaussian Kernels achieved a 99.9% accuracy for MS-ILM [9]. kNN systems’ performance on a 2-class task is 87.5% [10]. The accuracy of kNN for category recognition for signatures sampled at 4.4 kHz was at 85.5% [11]. The choice of machine learning algorithms in appliance classification however varies with appliance categories and with different types of measurements. Recent researches proposed load monitoring systems based on different frameworks, e.g. using Interactive Voice Response [12], IoT [13], and a home energy saving network [14], and Murali proposed load monitoring via social media [15]. The scope, however, as to accuracy and specific appliance types and feature sets considered, especially those that used machine learning approaches, is not clear.

Most commercial energy meters do not come with appliance identification features which can be useful for remote load monitoring and control. With ILM a recent research area of interest, publicly available datasets for ILM are scarce, with signatures recorded at low frequency and with very limited appliance types, influencing significantly the reported performance of scientific works. To date, no ILM study has been conducted yet in the Philippines. This study explores the concept of developing a system for appliance classification using a supervised learning technique based on power consumption traces following the ILM approach.

The main objective of the study is to develop a system for appliance classification based on power signatures. Specifically, the study aims to: (1) develop an individual appliance metering device to be used in appliance classification; (2) identify appliances based on power signatures using a supervised learning technique; and (3) create a load library of appliance power signatures as corpus data for training and recognition.

With the availability of an appliance classification system, with provisions to connect and control electric appliances remotely, acquisition of data can be made convenient for monitoring an appliance’s consumption and usage patterns which can be used in managing appliance usage. Also, hazards caused by neglect to turn off appliance power which may lead to fire can be avoided.

Five appliance types are considered in this study, namely, computer set, electric fan, laptop (via charger), printer, and mobile phone (via charger). The choice of the appliance to be included in the study is based on the available appliances common to ten offices in UM Tagum College. This study covers the use of time-dependent power features drawn by an appliance, from power-up to its steady state, as bases for its classification. Also, this study assumes that the appliance to be subjected for classification is in good working condition.

2. Methodology

2.1. System architecture
The architecture of the appliance classification system is shown in figure 1. The smart meters are attached between the home outlets and the appliance. These meters perform data acquisition, feature extraction and appliance classification, in that order. A user interface (computer monitor) can be used to access a meter for maintenance purposes. The load library of appliance power signatures to be used for classification is local to the meter. Classification results are relayed to the server through the gateway, which contains drivers and protocols to enable the meter to communicate over the public Wide Area Network – the Internet. The smart meters have cloud-based back-end to provide information for views using a mobile device. The mobile device can control the turning on or off of the connected appliance, allowing not only remote user access but also control.

2.2. Smart meter hardware
As shown in figure 2, the appliance is connected in parallel with the main voltage source. The potential transformer (PT) and the current transformer (CT) step down the high value voltage and current levels, respectively, to safe values compatible with digital circuitries. The Raspberry Pi is used as the central processing unit of the smart meter because of its low-cost and reliability, appropriate for wireless sensor network monitoring system [16]. But since the Raspberry Pi does not come with a built-in analog-to-digital converter, a high precision Analog-to-Digital Converter (ADC) is used to convert the sensed analog voltage and current values to their digital equivalents. The
latching relay acts as an actuator to turn on or off a plugged appliance when a control signal from the Raspberry Pi is received. The Raspberry Pi has a built-in wifi module to enable the meter to communicate to a mobile device via the Internet.

![Figure 1. Appliance classification system architecture.](image1)

**2.3. Appliance classification framework**

Appliance classification seeks to determine the class of a new observation, based on a training data set containing observations whose classifications are known. The framework, whose major components are discussed below, is used in this study as shown in figure 3.

![Figure 3. Appliance classification framework.](image2)

**2.3.1. Voltage and current samples collection.**

Data acquisition is crucial in classification problems in that the results of the succeeding processes are greatly dependent on the soundness of the collected data. For data involving continuous signals, the timing and length of collection are critical. The smart meter developed for the purpose of this study is used to acquire voltage and current readings, with the appliance initially turned off. Seventy five plug loads from five appliance classes, fifteen of each class, including computer set, electric fan, laptop (via charger), printer, and mobile phone (via charger) were recorded for the training data set. Two thousand samples per appliance class were collected at a rate of 100 samples per second. The data were collected from the time the appliance was turned on until its steady state. Data for appliances via chargers were taken while the appliance was charging.
2.3.2. Feature extraction. Feature extraction is the process of deriving useful and distinct quantities from an initial set of measured data. In the context of electrical metering devices, the voltage and current readings obtained from the analog-to-digital converter were in the form of counts. These data still need to be pre-processed, based on the ADC’s resolution and the system voltage and current, to determine the analog voltage and current values measured, respectively. These values are then used to derive relevant electrical quantities to form part of an appliance’s power signature. The feature set used in this study includes real power (P), peak current (I_{PEAK}), RMS current (I_{RMS}) and the power factor (PF). Frequency and RMS voltage are not included since these quantities are considered constant for all devices and so are considered redundant features.

2.3.3. Classification model. The classification model is the library of appliance power signatures, appended with a corresponding appliance classification label for each signature set that will be used by the classifier to predict the appliance class.

2.3.4. Classification algorithm. The k-Nearest Neighbors classifier works by comparing the most similar piece of data to the training set and returns the class label [17]. The k value is the number of most similar pieces of data from the training set to be resolved by a majority vote. The classifier takes in four inputs: the input vector to classify, the matrix of training data set, the vector of labels and the value of k nearest neighbors to use in voting. To deal with values that lie in different ranges, normalization is done.

3. Results and discussions
The smart meter developed for this study was used for data gathering. After collection and the data pre-processed, the current values were plotted to visualize the actual current drawn by an appliance at start-up until its steady state for analysis.

The time it takes for the current to get to steady state differs from one appliance class to another. The computer set took approximately eight seconds after turning on to get to its steady state, and its current peaked, on average, at 0.464569155 A, as shown in figure 4.

![Figure 4](image4.png)

**Figure 4.** Sample power consumption traces of a computer set.

The mobile phone however immediately stabilized after turning on, with its peak current vary by only about 0.01 A from its average peak steady state current, as shown in figure 5.

![Figure 5](image5.png)

**Figure 5.** Sample power consumption traces of a mobile phone.
The electric fan peaked, on average, at 0.343963255 A and took approximately three seconds to get to steady state, as shown in figure 6.

![Figure 6. Sample power consumption traces of an electric fan.](image)

In figure 7, it can be observed that the laptop current steadied after about two seconds from turning on and peaked at 0.46006135 A.

![Figure 7. Sample power consumption traces of an electric fan.](image)

Very noteworthy is the large start-up current drawn by the printer shown in figure 8 which peaked at 3.704620815 A and significantly dropped approximately to almost 0 A eight seconds after turning on.

![Figure 8. Sample power consumption traces of a printer.](image)

It can be observed in the figures that the computer set, electric fan, mobile phone, and laptop have bigger steady state current to peak current ($I_{PEAK}$) ratios whereas the printer has a very small steady state to peak current ratio, which means that the printer draws a much bigger current upon start-up compared to its current draw during steady state.

The RMS current ($I_{RMS}$) of an appliance was computed using the average peak current values, per 100 samples, at steady state. This value is central to the computation of an appliance’s power consumption, with their relation linear. At steady state, a consistent current consumption pattern for each appliance class can be observed.

A sample power signature of a computer set at steady state is shown in figure 9, with an RMS current value of 0.37881478 A.
Figure 9. A sample power signature of a computer set at steady state.

A sample RMS current of a mobile phone is 0.074185947 A, evident in its power signature shown in figure 10.

Figure 10. A sample power signature of a mobile phone at steady state.

Figure 11. A sample power signature of an electric fan at steady state.

Figure 11 shows a sample power signature of an electric fan at steady state with an RMS current value of 0.276010293 A.

A laptop has a sample RMS current value of 0.339869267 A with its power signature at steady state shown in figure 12.

Figure 12. A sample power signature of a laptop at steady state.

A sample power signature of a printer is shown in figure 13, with an RMS current value of 0.127470846 A, a very remarkable drop from 3.704620815 A during start-up.
Figure 13. A sample power signature of a printer at steady state.

The value of the real power (P), which is the actual power consumed by an appliance, can only be obtained if the power factor value is known. For the power factor (PF) to be determined, the time difference of $t_2$ and $t_1$ in figure 14, that is when the voltage and current signals cross zero, was measured. To increase the measurement accuracy, the average of the zero-crossing time differences for the entire sampling duration was used.

Figure 14. Zero crossings of voltage and current signals.

After all the needed power features were collected and the training set prepared, the k-Nearest Neighbors classifier was tested on the data set with $k = 1$. The estimated error rate of the classifier on the training set was 0%.

Actual testing results are shown in table 1. The overall classification accuracy is 95.55%. The kNN classifier was able to correctly predict, in all instances, three appliances. The case when the printer was incorrectly classified was when the classification process was initiated and the printer was not turned on immediately, which significantly affected the results of the extracted features from the testing data for classification. The classification model for unknown appliance however slightly affected the overall accuracy with some unknown appliances’ power consumption traces closely resembled those of the known classes.

Table 1. Confusion matrix.

|               | Computer Set | Electric Fan | Laptop | Mobile Phone | Printer | Unknown |
|---------------|--------------|--------------|--------|--------------|---------|---------|
| Computer Set  | 14           |              |        |              |         | 1       |
| Electric Fan  |              | 15           |        |              |         |         |
| Laptop        |              |              | 15     |              |         |         |
| Mobile Phone  |              |              |        | 15           |         |         |
| Printer       |              |              |        |              | 1       | 14      |
| Unknown       |              |              |        |              | 1       | 13      |

4. Conclusion and future works

An appliance classification smart meter was developed and a library of five appliance classes, each with fifteen training sets, was created. A supervised learning approach using the kNN classifier was used in the appliance classification, with high accuracy. Based on the results, the timing of data acquisition is critical, especially for appliance whose steady state RMS current to peak current ratio is very small. For future works, the load library needs to be populated with more appliance classes, of different models and brands, to gradually capture all appliance types. Also, for appliance with finite states, the changes in power consumption with state transitions should be considered.
5. References

[1] Reddy, R S, Keesara, N, Pudi, V and Garg, V 2015 Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association (India: Hyderabad)
[2] Hart, G W 1992 Proceeding of the IEEE 80 12
[3] Ito, M, Uda, R, Ichimura, S, Tago, K, Hoshi, T and Matsushita, Y 2004 Proceedings of the 2004 IEEE International Symposium on Applications and the Internet pp 291
[4] Reinhardt, A, Burkhardt, D, Zaheer, M and Steinmetz, R 2012 Proceedings of the 7th IEEE International Workshop on Practical Issues in Building Sensor Network Applications (USA: Florida)
[5] Englert, F, Kobler, S, Reinhardt, A and Steinmetz, R 2013 Proceedings of the 4th International Conference on Future Energy Systems (USA: California)
[6] Ridi, A, Gisler, C and Hennebert, J 2015 proceedings of the 13th International Conference on Pervasive Computing and Communications (Percom 2015), Workshop on Pervasive Energy Services (PerEnergy) pp 323 – 328
[7] Ri di, A, Gisler, C and Hennebert, J 2014 Proceedings of the 22nd IEEE International Conference on Pattern Recognition pp 3702 – 3707
[8] Zaidi Adeel Abbas, F K and Palensky, P 2010 Proceedings of the 36th IEEE Annual Conference on Industrial Electronics Society (IECON) pp 2442-2447
[9] Kato, T, Cho, H, Lee, D, Toyomura, T and Yamazaki, T 2009 Ambient Assistive Health and Wellness Management in the Heart of the City 5597 pp 150–157
[10] Fitta, M, Biza, S, Lehtonen, M, Nieminen, T and Jacucci, G 2010 Proceedings of the 4th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UbiComm) pp 471–477
[11] Saitoh, T, Aota, Y, Osaki, T, Konishi, R and Sugahara, K 2008 Proceedings of the 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC ) pp 349–352
[12] Pandiri, S and Srikanth, G 2016 International Journal of Innovative Technology and Research (IJITR) 4 4 pp 3350 – 3353
[13] Kurde, A, Kulkarni, V S 2016 International Journal of Scientific and Research Publications 6 9
[14] Manjusha, P, Saran Kumar, T, Venkateshwarullu, V and Nageshwar Rao, I R S 2017 International Journal of Science, Engineering and Technology Research (IJSETR) 6 9
[15] Murali, M J 2017 International Journal of Advance Research and Innovative Ideas in Education 3 3
[16] Cabacan, C N, Cruz, F R G and Agulto, I C 2017 Proceedings of the 2017 IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (Philippines: Manila)
[17] Harrington, P 2012 Machine learning in action Manning Publications Co. New York

6. Acknowledgements
The researchers wish to extend their sincerest gratitude to all the persons who, in one way or another, contributed to the success of this endeavor. This research was supported by the Commission on Higher Education.