Building a Vector-based load taxonomy using electrical load signatures

H Y Lam, K H Ting, W K Lee and G S K Fung
Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong

Abstract

A load signature of an appliance describes the appliance behaviour in an electrical circuit. By comparing load signatures of various loads, we can separate the loads in terms of their differences, and cluster them to the same class in terms of their similarities. The process shall generate a load taxonomy. Traditional metrics are ineffective to distinguish electronic loads. This paper introduces a new methodology of classifying appliances by extracting vectors form load signatures.

Keywords: load signatures, vector, load taxonomy

1 INTRODUCTION

Load signatures contain useful information with respect to operational characteristics of the loads in an electrical circuit. Applications including power quality and management can be derived through utilization of the knowledge-base of a load taxonomy.

Table 1 List of measured appliances

Type	Appliance
Resistive	Kettle, incandescent lamp, hair dryer, soldering iron, rice cooker, iron, fan heater, column heater, humidifier, water boiler
Motor-driven	Fan, vacuum cleaner, blender, washer
Pump-operated	refrigerator, dehumidifier, water pump, air conditioner
Electronics	Energy-saving light bulb, fluorescent lamps with electronic ballast, CD player, mobile phone battery charger, battery charger, desktop PC, LCD monitor, CRT monitor, ink jet printer, oscilloscope, notebook computer, television, video recorder, desktop fluorescent lamp, scanner, laser printer, LCD television, DVD player, radio, hi-fi, router, PS2, projector, plasma TV
Fluorescent lighting	Fluorescent lamp with conventional ballast, aquarium lighting
Others	microwave oven, induction cooker

In this project, load signatures of 74 pieces of home and office appliances, operating in their various designed modes, were measured and investigated. Traditional metrics such as r.m.s current, power, power factor, harmonics and transient duration, were extracted form their signatures. Using classification proposed by F. Sultanem [2], appliances were classified as Table 1. Such classification showed that these metrics could not distinguish electronic appliances effectively. Therefore, another approach based on vector was used to classify these electronic appliances. This paper presents this methodology and its extension in building a vector-based load taxonomy.

2 METHODOLOGY

2.1 Load Signature

The current and voltage of the appliances operating in their designed modes are measured. Each data is sampled at a high frequency, thus its current and voltage waveforms can be analyzed up to hundreds of harmonic components.

2.2 Vector Extraction

The raw current and voltage waveforms of the appliances were sampled at a high frequency, and thus carried a large set of time dependent variables. Dominant variables were separated from redundant variables by Principal Component Analysis (PCA). PCA reduced the size of variables by extracting a set of vectors from the raw data. Eventually a small set of orthogonal waveforms which represented the most significant properties of the original waveforms were used.

In this study, Singular Value Decomposition (SVD) was used to perform PCA. Let X be a m×n matrix containing the original data of all appliances; each column of the X is the original one-cycle steady-state current waveform of one appliance. The equation of SVD is $X=USV^T$, where U is a m×n matrix, S and V^T are n×n matrices. The rows of V^T are right singular vectors and form a set of orthonormal basis. S is a diagonal matrix containing the singular values. S and

V^T can be calculated by diagonalizing X^TX, where $X^TX=VS^2V^T$. The matrix SV^T contains the principal components score, which is the weighting of each principal component in the original signal. The columns in U are left singular vectors and form an orthonormal basis for the load current waveforms. U can be calculated by $U=XVS^{-1}$.

The data used for vector extraction was the one-cycle current waveform for each of the appliances. By PCA, a set of principal components were generated. The principal components form an orthogonal basis of the data. They were uncorrelated to each other and contained no redundant information. Each principal component was treated as a single axis in space. The original current signal was projected on the axis to obtain the weighting of the principal component in the original signal.

2.3 Hierarchical Clustering

Hierarchical clustering was performed to group similar appliances together with respect to a particular vector. The clustering result was represented by a tree called dendrogram as shown in Figure 1. The dendrogram shows how similar objects are grouped together. The horizontal axis shows the objects that required to be clustered. Similar objects are linked by upside-down U-shaped lines; the height of a U-shaped line represents the distances between the pair of objects that the U links to. Smaller distance indicates that the objects are more similar, so more similar objects are linked by shorter U.

The steps for hierarchical clustering on a set of data are as follow:

(a) *Measure the similarity or dissimilarity between every pair of objects in the data set.* This is done by measuring the distances between each pair of object in a particular vector space. In this project, the distances were calculated by using the Euclidean metric.

(b) *Group objects into a binary hierarchical tree.* Originally, each single object, such as the objects numbered from 1 to 10 in Figure 1, is treated as a single cluster. Pairs of similar objects, with smallest distances between them are linked together to form larger clusters. These clusters are continued to pair with each other until all objects are linked into one cluster. For example, object 3 and 7 are linked together to form a larger cluster. Object 2 is the most similar object to this large cluster, so this large cluster is linked with object 2 to form another large cluster. Eventually, all the 10 objects are linked together to form a single cluster.

(c) *Divide the hierarchical tree into clusters*. The hierarchical tree is divided by a cut off line. The cut off line is determined based on two considerations: (1) The number of clusters that we want to create; (2) The separation distance between the clusters, which is indicated by the height of the U-shaped link. In Figure 1, the U-shaped line linking object 9 and the cluster that contains object 1 to 8 is the longest, so object 9 and the cluster are the most dissimilar. A cut off line can be drawn between 0.8 and 1.6 on the vertical axis to cut the U-shaped line and create 2 clusters, with one cluster contains object 9 and the other contains object 1 to 8. But, only 2 clusters are too less and we want to have more clusters. So, a cut off line is drawn between 0.4 and 0.6 on the vertical axis as shown in Figure 1. In this way, four clusters are formed and the distances between the four clusters are largest.

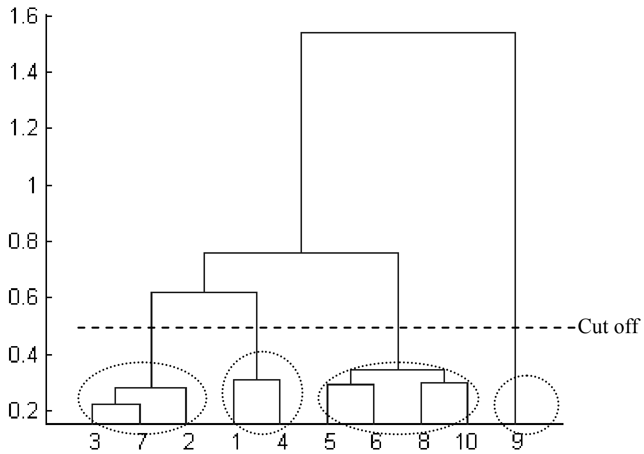


Figure 1 An example of a dendrogram – The number 1 to 10 represents the ten objects required to be clustered; vertical axis shows the distance between pairs of objects.

2.4 Building Load Taxonomy

With all the clustering results from every vector, a load taxonomy was formed in a hierarchical structure (Figure 2).

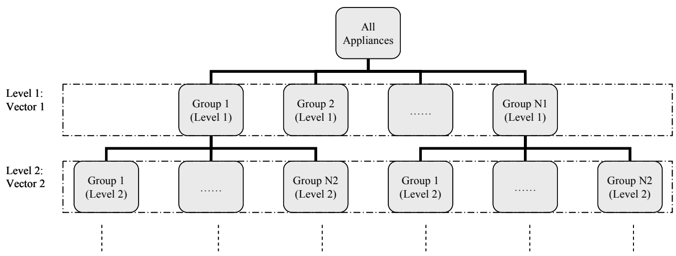


Figure 2 Structure of load taxonomy tree

The load taxonomy contains a number of levels, where each level corresponds to the clustering result of each vector extracted. The top of the taxonomy is a single group containing all appliances. For example, a number of vectors called Vector 1, Vector 2 and so on, are extracted from the original data. In level 1, the appliances are first clustered into different groups based on the clustering result of Vector 1. These groups are further divided into more groups based on the clustering result of Vector 2 as shown in level 2. The groups are further divided into smaller groups until the clustering results of all vectors are included in the taxonomy. Finally, similar appliances are grouped into the same group at the bottom of the taxonomy.

3 RESULTS

3.1 Load Signature Acquisition

A number of electronic appliances with a wide range of harmonic components were chosen to perform vector analysis. They included desktop PC, LCD monitors, CRT monitor, laser printer, ink jet printer, energy-saving light bulb, television, DVD player etc. Steady state data was used in vector analysis, despite transient data was also measured.

3.2 Vector Extraction

Vectors were extracted from the one-cycle steady-state current waveforms of the appliances. The extracted 16 most significant vectors are shown in Figure 3. Vector 1 is the most significant vector, which contributes the most in the original current waveforms. Vector 2 is less significant than Vector 1, and Vector 3 is less significant than Vector 2 and so on.

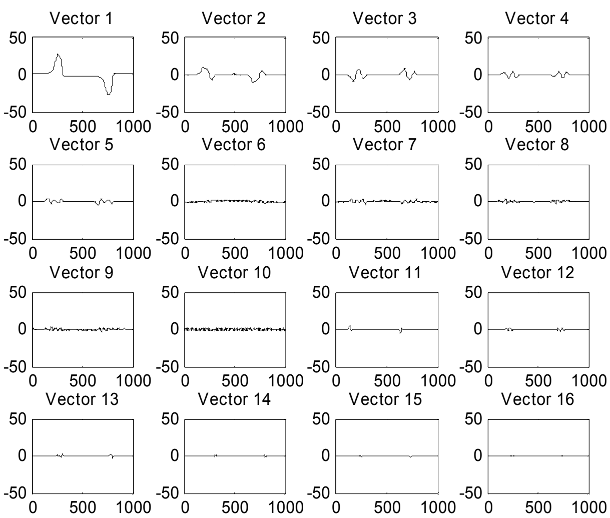


Figure 3 The 16 most significant vectors

Table 2 showed the percentage of total energy when different numbers of largest vectors were used. It was observed that the first six largest vectors contributed more than 95% of total energy. In other words, using Vector 1 to Vector 6 would be enough to approximately reconstruct the original waveforms sufficiently. For this reason, only these 6 vectors were used for clustering.

Table 2 Percentage of total energy with different number of largest vectors used

No. of largest vectors used	% of total energy
1	72.198
2	83.367
3	90.228
4	92.518
5	94.278
6	**95.684**
7	96.951
8	97.590

3.3 Hierarchical Clustering

The appliances were clustered according to the weightings of each vector in the original waveforms, and hierarchical trees for each vector were formed (Figure 4). For each hierarchical tree of each vector, a cut off line was used to divide the tree into clusters. Histograms were also used to show the number of appliances in different values of weighting of a particular vector in the original waveforms. The horizontal axis of the histogram was the weighting of vector, and the vertical axis was the number of appliances. A name was given to every cluster for each vector that would be used in constructing the taxonomy.

In the hierarchical tree of Vector 1 as shown in the first row in Figure 4, five clusters were formed by drawing a cut off line on the hierarchical tree. The five clusters can be observed in the histogram, and they were named as VS (Very small), S (Small), M (Medium), L (Large) and VL (Very large) as shown at the bottom of the histogram. Most of appliances were grouped into cluster VL, and their values of component score for Vector 1 were greater than 0.085. From Vector 1 to Vector 6, two to six clusters were formed for each vector.

3.4 Building Load Taxonomy

By gathering all the clustering results of the six vectors, a load taxonomy tree was created as shown in Figure 5. Abbreviations were used to represent the appliance names in the taxonomy tree (Table 3). In Figure 5, the top of the taxonomy tree is a large cluster including all the appliances. This large cluster is further divided into smaller clusters in the six levels below. Some clusters contain only one appliance, such as the mobile phone battery charger (MPC2O) in level 1, are not required to further divide and this single appliance forms one single class. The clusters at the bottom of the tree are the final clustering result using all six vectors. The appliances in the same cluster belong to the same class and they all have similar waveform shapes.

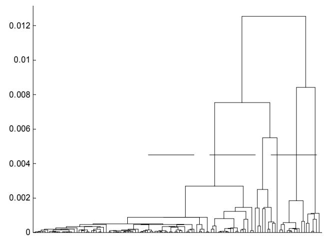
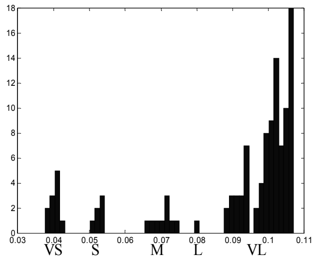
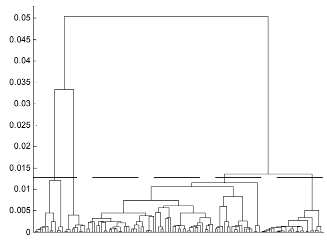
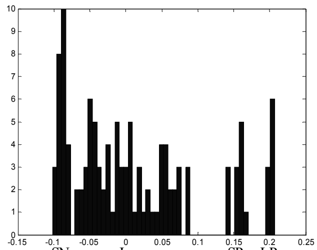
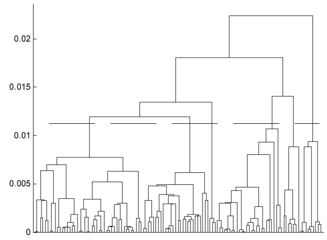
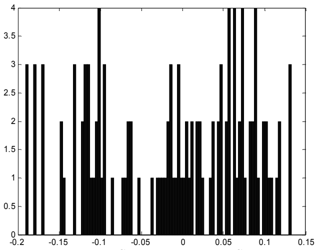
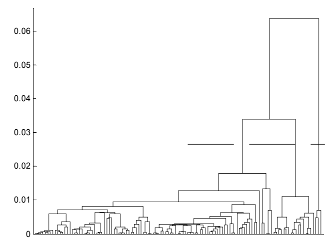
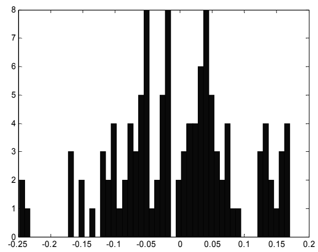
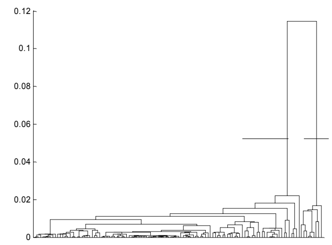
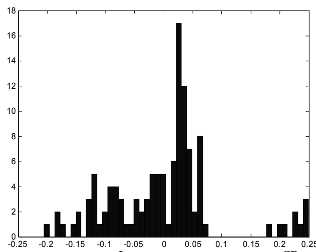
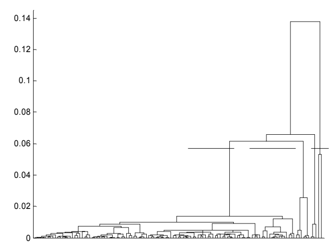
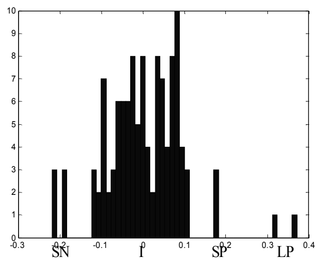
Vector	Hierarchical Tree	Histogram	Ranges
1			**VS**: Very small $0.03 \le vector1 < 0.045$ **S**: Small $0.045 \le vector1 < 0.06$ **M**: Medium $0.06 \le vector1 < 0.075$ **L**: Large $0.075 \le vector1 < 0.085$ **VL**: Very large $vector1 \ge 0.085$
2			**SN**: Small negative $-0.15 \le vector2 < -0.07$ **I**: Insignificant $-0.07 \le vector2 < 0.12$ **SP**: Small positive $0.12 \le vector2 < 0.18$ **LP**: Large positive $vector2 \ge 0.18$
3			**LN**: Large negative $-0.2 \le vector3 < -0.16$ **MN**: Medium negative $-0.16 \le vector3 < -0.08$ **SN**: Small negative $-0.08 \le vector3 < -0.05$ **I**: Insignificant $-0.05 \le vector3 < 0.03$ **SP**: Small positive $0.03 \le vector3 < 0.125$ **LP**: Large positive $vector3 \ge 0.125$
4			**SN**: Small negative $-0.25 \le vector4 < -0.2$ **I**: Insignificant $-0.2 \le vector4 < 0.1$ **SP**: Small positive $vector4 \ge 0.1$
5			**I**: Insignificant $-0.25 \le vector5 < 0.15$ **SP**: Small positive $vector5 \ge 0.15$
6			**SN**: Small negative $-0.3 \le vector6 < -0.15$ **I**: Insignificant $-0.15 \le vector6 < 0.15$ **SP**: Small positive $0.15 \le vector6 < 0.25$ **LP**: Large positive $vector6 \ge 0.25$

Figure 4 Hierarchical trees and histograms for Vector 1 to Vector 6

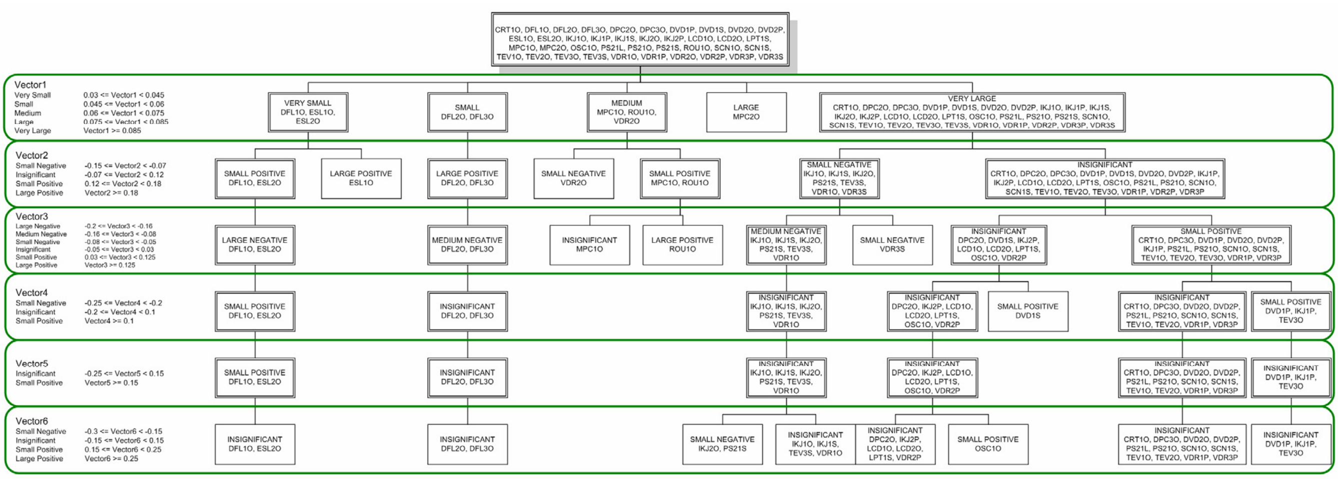


Figure 5 Load Taxonomy tree

Table 3 Abbreviation of the appliances

Abbreviation	Appliance
CRT1O	CRT monitor, on
DFL1O	desktop fluorescent lamp 1, on
DFL2O	desktop fluorescent lamp 2, on
DFL3O	desktop fluorescent lamp 3, on
DPC2O	desktop PC 2, on
DPC3O	desktop PC 3, on
DVD1P	DVD player 1, playing DVD
DVD1S	DVD player 1, stand by
DVD2O	DVD player 2, on
DVD2P	DVD player 2, playing DVD
ESL1O	energy-saving light bulb 1, on
ESL2O	energy-saving light bulb 2, on
IKJ1O	ink-jet printer 1, on
IKJ1P	ink-jet printer 1, printing
IKJ1S	ink-jet printer 1, stand by
IKJ2O	ink jet printer 2, on
IKJ2P	ink jet printer 2, printing
LCD1O	LCD monitor 1, on
LCD2O	LCD monitor 2, on
LPT1S	laser printer, on
MPC1O	mobile phone battery charger 1, on
MPC2O	mobile phone battery charger 2, on
OSC1O	oscilloscope, on
PS21L	PS2, loading disc
PS21O	PS2, on
PS21S	PS2, stand by
ROU1O	router, on
SCN1O	scanner, on
SCN1S	scanner, scanning
TEV1O	television 1, on
TEV2O	television 2, on
TEV3O	television 3, on
TEV3S	television 3, stand by
VDR1O	video recorder 1, on
VDR1P	video recorder 1, playing
VDR2O	video recorder 2, on
VDR2P	video recorder 2, playing
VDR3P	video recorder 3, playing
VDR3S	video recorder 3, stand by

3.5 Observation

Vectors are unique combination of harmonics

It is observed that the harmonics of the six vectors are mainly odd harmonics. Observing from Vector 1 to Vector 6, the dominant frequency components shift from low frequency to high frequency as shown in Figure 6. So, the vectors measure the significance of its frequency components from low to high frequency in the original waveforms.

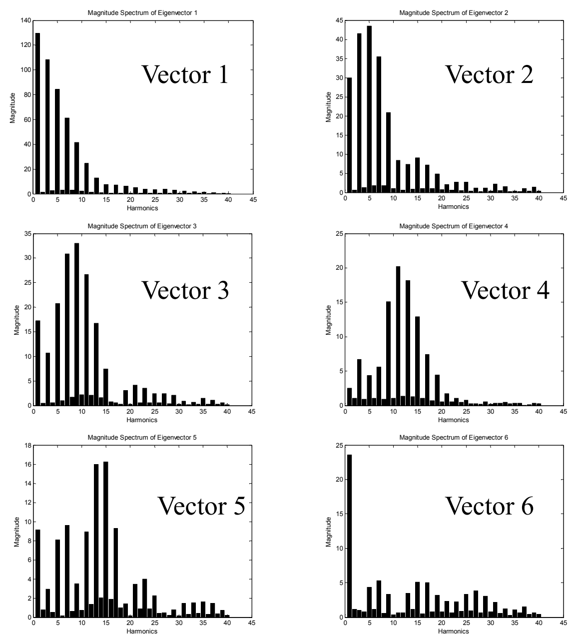


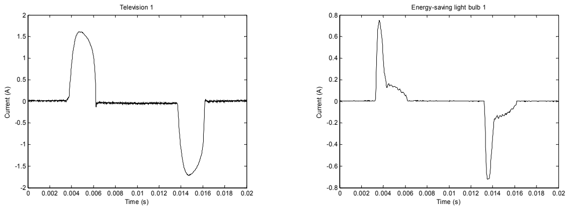
Figure 6 Harmonics of Vector 1 to Vector 6

Vector-based taxonomy is able to distinguish appliances with different waveform shapes

Vector 1 and Vector 2 are powerful in separating two main kinds of appliances according to the main difference of their waveforms. There are mainly two types of waveforms:

Simple peak – Simple sharp peak, for example, the current waveform of a television (Figure 7a)

Sharp peak with decaying curve – Sharp peak with a small decaying round curve after it, an example is the current waveform of an energy-saving light bulb (Figure 7b).



(a) Television (b) Energy-saving light bulb

Figure 7 Two different types of waveforms

Appliances having simple peak waveform are all have large weighting of Vector 1, and they are grouped into the "Very large" cluster as shown in the first level of Figure 5. They include televisions, ink jet printer, DVD player etc.

Appliances having waveform of sharp peak with decaying curve are all have small weighting of Vector 1 and large weighting of Vector 2. They include energy-saving light bulb and desktop fluorescent lamp.

4 CONCLUSIONS

In this study, the load signatures of appliances were transformed into vectors, and these vectors contained the important properties of the appliances. Appliances having similar weighting of each vector can be assigned to the same group using hierarchical clustering. Levels of groups based on the clustering results shall create a taxonomy tree. The result showed that the vector-based taxonomy was effective in distinguishing different types of electronic appliances based on their differences in current waveforms. This load taxonomy provides a quick reference to understand similarities and differences among different appliances. It can also help both power providers and power users to understand the effect of appliances on the power system and make improvements on power quality, load identification, load management and smart metering.

REFERENCES

[1] W K Lee, G S K Fung, H Y Lam, F H Y Chan, M Lucente, "Exploration on Load Signatures", Proc. of International Conference on Electrical Engineering (ICEE2004), vol. 2, July 4-8, 2004, pp.690-694.

[2] F. Sultanem, "Using appliance signatures for monitoring residential loads at meter panel level," IEEE Transactions on Power Delivery, Vol. 6, pp.1380-1385, 1991

[3] I. T. Jolliffe, Principal Component Analysis, Springer, New York, 1986

[4] L. Kaufman and P. J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, Wiley, New York, 1990.