Zoha, A., Gluhak, A., Nati, M., Imran, M. A., and Rajasegarar, S. (2012) Acoustic and Device Feature Fusion for Load Recognition. In: 2012 6th IEEE International Conference Intelligent Systems, Sofia, Bulgaria, 06-08 Sep 2012, ISBN 9781467322782.

Acoustic and Device Feature Fusion for Load Recognition

Ahmed Zoha, Alexander Gluhak, Michele Nati, Muhammad Ali Imran
Centre for Communication Systems Research
University of Surrey, Guildford, GU2 7XH, UK
a.zoha@surrey.ac.uk

Sutharshan Rajasegarar
Department of Electrical and Electronic Engineering
University of Melbourne, Melbourne, Australia
r.sutharshan@ee.unimelb.edu.au

***Abstract*—Appliance-specific Load Monitoring (LM) provides a possible solution to the problem of energy conservation which is becoming increasingly challenging, due to growing energy demands within offices and residential spaces. It is essential to perform automatic appliance recognition and monitoring for optimal resource utilization. In this paper, we study the use of non-intrusive LM methods that rely on steady-state appliance signatures for classifying most commonly used office appliances, while demonstrating their limitation in terms of accurately discerning the low-power devices due to overlapping load signatures. We propose a multi-layer decision architecture that makes use of audio features derived from device sounds and fuse it with load signatures acquired from energy meter. For the recognition of device sounds, we perform feature set selection by evaluating the combination of time-domain and FFT-based audio features on the state of the art machine learning algorithms. The highest recognition performance however is shown by support vector machines, for the device and audio recognition experiments. Further, we demonstrate that our proposed feature set which is a concatenation of device audio feature and load signature significantly improves the device recognition accuracy in comparison to the use of steady-state load signatures only.**

Keywords;Non-intrusive Load Monitoring (NILM); energy reduction; energy monitoring; audio features; Support Vector Machines (SVM)

I. INTRODUCTION

The energy consumption in residential spaces and offices is increasing every year [1], which is a growing concern because the energy resources are limited as well as it has negative implications on the environment (e.g. CO2 emissions). As a result, we see recent initiatives taken by governments across Europe and USA for the large scale deployment of smart meters for improved energy monitoring. . Smart meters however can only measure energy consumption on a house level granularity, providing little information on the breakdown of the energy spent. A major challenge is to acquire appliance-level information, providing details which appliances have been used, how much they have consumed as well as when and why they are operated. Such fine-grained energy monitoring is essential for providing meaningful information in real-time to the consumers about their energy consumption behavior, which requires identification of energy hungry devices. But it is still a challenge for appliance-specific LM solutions to be accurate and cost effective altogether, especially in the recognition of low-power consumer appliances. It is because low-cost solutions make use of steady-state energy consumption characteristics of the devices in order to recognize them, whereas most of the low-power consumer appliances have similar steady-state behavior which makes it difficult to perform reliable automatic load identification. Motivated by this, in this paper we propose a multi-layer decision architecture that combines acoustic features and device steady-state energy features to reliably identify low-power consumer appliances commonly found in an office environment. Our proposed method correlates device energy consumption characteristic and acoustic activity within an environment to facilitate steady-state Non-Intrusive Load Monitoring (NILM) method. In addition, the acoustic information obtained from the environment can further be used to identify user's inefficient energy behavior; detect situations when unattended appliances are consuming energy purposelessly. In particular we make the following contributions with our presented research:

- We report performance comparison of steady-state feature sets used in NILM for discerning low-power office appliances. In our experiment, we further demonstrate the limitation of using these feature sets for identifying appliances with overlapping load signature.
- We investigate the discriminative ability of time-domain and frequency based audio features for the task of identifying machine and user-generated acoustical events, and further performs optimal classification model selection.
- We present a multi-layer decision architecture that combines acoustic and device energy features to reliably identify office appliances, providing increased detection performance for low power devices with overlapping load signatures compared to existing load monitoring methods that only rely on steady state signatures derived from energy consumption measurements.

The remainder of the paper is organized as follows. In the next section we review related research work, highlighting the limitation of the current approaches. In

Section III we provide an introduction to our proposed multi-layer decision framework, whereas we present our experimental evaluations for load identification and acoustic event recognition in Section IV. Finally, we provide conclusion and our future work in Section V.

II. Related Work

The existing approaches to appliance load monitoring can be classified into two categories: Intrusive and Non-intrusive load monitoring (NILM) respectively. The intrusive load monitoring makes use of multiple sensors per appliance to measure its energy consumption. Though we can achieve a high accurate measure of device energy usage through this approach but high installation complexity, cost, as well as calibration and data aggregation are outstanding issues that will not favor the use of this technique. In contrast to that the NILM approach as proposed by Hart's [2]makes use of feature extraction process to extract load signatures and apply machine learning algorithms to disaggregate device specific data from the aggregated data, acquired using single point measurement. The NILM methods are further classified into steady-state and transient methods based on the sampling frequency of the load signatures. The transient methods require high sampling rate for acquiring transient signatures that is shape, size and duration of the transient waveforms that occur during state transitions. The transient behavior of major appliances is found to be distinct and researchers [3] [4] [5] show that transient features are less overlapping in comparison to steady state signatures; providing much higher recognition accuracy for multi-state and low-power consumer appliances. However transient methods suffer from drawbacks such as sensitivity to the wiring architecture of the target environment and expensive hardware due to high sampling rate requirement, which both limit the applicability of this approach [3].

On the other hand the steady state methods identify devices based on variations in their steady state signatures that are present throughout the steady state operation of the load such as power (i.e., real power, reactive power), current and voltage signatures. One of the advantages of using steady-state method is that steady-state signatures can easily be extracted using cheap energy meters. Several commercial solutions (e.g., Plogg, Kill-A-Watt, and Watts Up) [6] are available in the market that can easily measure steady-state energy consumption characteristic of an appliance and they have the capability to wirelessly transmit the data, thus they can form a wireless sensor network. A similar approach is adopted by Jiang et al. [7]who provide design and implementation of a wireless sensor based AC Metering (ACME) network architecture for various power centric applications making use of steady-state load signature. Nevertheless, it has been reported in [2] [8] that, appliances with ON/OFF operating characteristics (e.g. toaster or a lamp) as well as large loads water heaters, refrigerators have distinct steady-state signatures but most of the low-power consumer appliances, as well as devices with multiple states are hard to recognize due to overlapping signatures in feature space. Norford [8] tried to reduce the ambiguous overlapping of steady-state signatures by using Fast Fourier Transform (FFT) to acquire harmonic content of the input current. However, this approach works on a limited number of appliances whereas accurate recognition of low-power consumer appliances is still a challenge. The side information could help in achieving high recognition accuracy such as, [9]developed an automated annotation system using external information from ambient sensors along with energy features, however the drawback is the use of multiple sensors. Similar to our proposed approach, in [10] a prototype has been developed that make use of audio and steady-state signatures for appliance recognition, however it considers very limited high-power sound emitting residential appliances and did not include any low-power consumer appliances in their evaluation.

In this paper, we have proposed a solution that improves low-cost steady-state methods by utilizing side information acquired from detecting acoustic events within an environment. The fusion of energy and acoustic features facilitates the NILM method to reliably detect low-power consumer appliances with overlapping load signatures, as discussed in the next section.

III. Multi-Layer Decision Framework for Smart Sensing

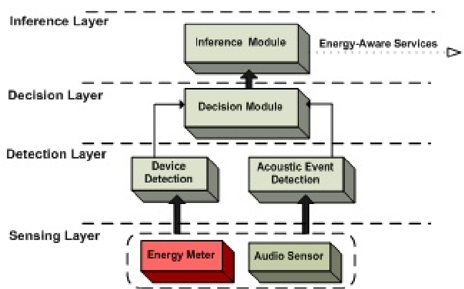


Fig. 1: Multi-Layer Decision Framework for Smart Sensing

We have already discussed in Section II that steady-state load signatures due to their low-sampling rate requirement can easily be extracted using cheap energy meters. However, due to the similarity in the steady-state energy consumption pattern of the low-power consumer appliances NILM methods fails to perform well. Therefore, we proposed a multi-layer decision framework as shown in Figure 1, which combines information from energy meter and the audio sensor in order to perform automatic and reliable detection of low-power consumer appliances. We briefly summarize the functionality of each layer

1) Sensing Layer: The sensing layer consists of an energy meter installed at a metering point or it can be a commercially available solution (i.e. Plogg unit [6]), and an audio sensor. The appliance-specific energy related information is acquired from the energy meter whereas the audio sensor provides us with the acoustic information from the environment. The energy consumption pattern of an active appliance in terms of current, voltage, real and reactive power draw, together with real-time acoustic events from the environment is sensed via the sensing layer. This information is provided to the detection layer for appliance and acoustic event recognition.

2) *Detection Layer:* The detection layer consists of device detection and acoustic event detection modules as shown in Figure 1. The functionality of each of these modules is described below.

a) Device Detection Module (DDM): The Device Detection Module (DDM) performs automatic appliance recognition using NILM load identification framework as shown in Figure 2.

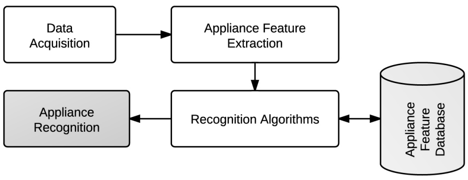


Fig. 2: NILM flow diagram for Load Identification

The NILM approach is commonly based on supervised learning, that requires extraction of load signatures and correspondingly labelling it with a device class to develop a appliance feature database. The labeled data or the training set is used to train the recognition algorithms so that any test input can be classified by matching it with training examples. As discussed earlier in Section II, research work in the past indicates that most of the low-power consumer appliances have similar or overlapping steady-state load signatures; therefore recognition algorithms fails to perform well when trained with these features. The role of DDM is to perform initial detection of appliances based on the steady-state load signatures acquired from the energy meter via the sensing layer. It further identifies the target device classes that are misclassified during the training phase as demonstrated in Section IV.A and pass on this information to the decision layer in order to perform accurate recognition of these confused classes.

b) Acoustic Event Detection Module (AEDM): The main functionality of Acoustic Event Detection Module (AEDM) is to perform acoustic event detection for the acoustic surveillance of the target environment. In our experimental evaluations as discussed in Section IV.B, we have considered machine and user-generated sounds as target acoustical events that most commonly occur within an office environment.

1)Decision Layer: The decision layer reevaluates the recognition results for the confused classes identified by the DDM. In Section IV.A of this paper, we first demonstrate the limitation of using steady-state load signatures in recognizing a set of target office appliances. Furthermore, we addressed this challenge in the decision layer by combining information from AEDM and DDM to improve the appliance recognition.

*2) Inference Layer:*The decision layer forward the appliance state information as well as the acoustic cues from the environment to the inference layer. The task of the inference layer is to correlate the acoustic activity of the environment with that of device activity in order to develop energy-aware applications such as user-specific appliance scheduling etc. However, our experimental evaluation in this paper did not demonstrate the functionality of the inference layer which we plan to address in our future work.

IV. Experimental Evaluations

For experimental evaluation of our proposed architecture as discussed in Section III, we have considered an office scenario. As for sensing layer, the real-time energy consumption statistics of the nine most common low-power office appliances are collected using Plogg [6]unit. Similarly, 12 target acoustical events which are most likely to occur in an office scenario including user-generated and machine specific sounds are included in the sound database. In order to identify most discriminative feature set for the task of appliance and acoustic event recognition, we have performed feature set and classification model selection experiments in the detection layer. We further analyze the classification performance of each algorithm that enable us to identify which of appliance and sound classes are hard to recognize due to ambiguous overlapping in the feature space. Finally based on the best feature set obtained from DDM and AEDM, we perform feature fusion in the decision layer to improve the low-power device recognition results. Our results and evaluations have been reported below.

A. DDM: Device Recognition based on Steady-State Load signatures

1)Feature Extraction: A load or a device signature is a unique energy consumption pattern of the device that characterizes its operation and distinguishes it from other loads also referred to as device features. The aim of NILM is to perform automatic recognition of the devices and their operational states based on their load signatures. However it requires an appliance feature database to be developed as shown in Figure 2, therefore we have collected six steady-state load signatures: real power (P), reactive power (Q), frequency (F), voltage (V_{RMS}), current (I_{RMS})) and phase angle (φ), by measuring the energy consumption target appliances which are listed in Table I.

TABLE I: Target Office Appliances

Class No	Devices	No of Products	Operational States
1	Fluorescent Lamp	3	1
2	Incandescent Lamp	2	1
3	Laptops	3	3
4	MAC	1	3
5	LCD Screen	3	2
6	Fan	2	2
7	Mobile Charger	2	1
8	Desktop Computer	2	3
9	Printer	1	2

We have tried to collect diverse samples of data within each class in order to achieve better generalization. We have included feature samples from four different laptops, two desktop computers, three mobile chargers, three LCD screens, one printer, one fan and three different lamps. It has helped us to identify intra-class variation of feature values. For example, we have found out that the real power consumption of MAC-book is 13W whereas for other laptops it is around 44W. Similarly within the lamp class, fluorescent lamps and incandescent lamp show different consumption characteristics. Hence, they are divided into separate device classes. We end up with nine different appliance classes as shown in Table I; for which six different load signatures have been collected. We combine these load signatures into three different feature sets as listed in Table II.

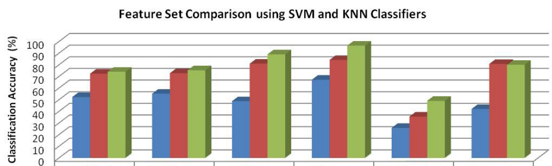
TABLE II: Device Feature Set

Feature Set Number	Features
Feature Set 1	P,Q
Feature Set 2	I, V, Frequency , Phaseangle
Feature Set 3	I,V, P,Q,Phaseangle

The selection of features is an important step because it is often a case that the performance of classifier is influenced by the presence of redundant features. In the feature set 1 we have included two most commonly used device features, real power (P) and reactive (Q), whereas in feature set 2 we completely neglected the P and Q features and instead used combination of V_{RMS}, I_{RMS}, Frequency and Phase angle. We want to analyze the discriminative ability of these feature sets for the classification of target appliances. Finally we combined feature set 1 and feature set 2 neglecting the Frequency feature because we have found out from the experimental results that almost all of the devices listed in Table I have similar high order harmonics in the range of 49.9-50.1 Hz, thus it has no impact on the classification performance.

2) Performance Evaluation of Device Feature Sets: The ability of each of these device feature sets (as listed in Table II) to discriminate between different device classes is evaluated by testing them with the state of the art classification algorithms. *k*-Nearest Neighbors (*k*-NN) because of its simplicity and ease of implementation is a popular choice for classification tasks [11], whereas Support Vector Machines (SVM) due to strong mathematical have proven to be effective in text and audio recognition experiments [12] . Therefore, we decided to test the performance of each of these classification algorithms on our target dataset. We have used MATLAB simulation environment for offline training and testing of the algorithms. *k*-NN uses Euclidean distance as a metric to partition the data. Thus it is easy to implement because training phase of the algorithm requires storing feature vectors and class labels. The new test sample is classified based on the majority voting of *k* nearest samples. The detail description of the algorithm can be found at [11]. On the other hand, SVM is a maximum margin classifier that tries to create a decision boundary between the two classes. One of major advantages of SVM is that it can model non-linear decision boundaries by applying kernel functions (i.e. polynomial, radial basis function etc.) that converts the non-linear input vector space to linear by transforming it to a higher-dimensional feature space. This allows the algorithm to find a decision boundary in the transformed feature space. The detail description of the algorithm can be found at [12]. SVM is originally a binary class classification algorithm. In order to extend it's applicability for multi-class classification problem we have used state of the art one-against one method. We have tested SVM with two kernel functions: polynomial with an exponential value of 2, and Radial Basis Function (RBF) whereas the gamma value (G) of the RBF kernel is found out to be 0.01 using 10-fold cross validation.

The overall recognition accuracy of the classifiers with different configurations is shown in Figure 3.



	KNN k=1	KNN k =10	SVM ker = Poly, Exp =2, C = 1	SVM ker = Poly, Exp =2, C = 100	SVM ker = RBF, G =0.01, C = 1	SVM ker = RBF, G =0.01, C = 100
Feature Set1	52.31	55.14	48.71	67.15	25.79	42.03
Feature Set2	72.31	72.78	80.94	84.14	35.503	80.72
Feature Set3	74.03	75.3	88.95	96.45	49.01	80

Fig. 3: DDM Feature Set Recognition Results

As for feature set 1, SVM with RBF kernel showed worst performance whereas *k*-NN with *k* value set to 10 show highest classification performance of 55%. Nevertheless, it has been found out that almost 45% percent of instances are misclassified by *k*-NN classifier due to high overlapping in the feature space and only six clusters are formed instead of nine. We did not except high performance from feature set 1 because in the early research work by Hart [2], devices had been categorized in the *P*-*Q* plane. It was found out that *P* and *Q* feature only works well for identifying high power devices such as air conditioners, microwave oven, cookers because they are well separated in the signature space.However, low power devices such as lamps, PC, and TV etc. heavily overlap in *P*-*Q* plane. Therefore , in feature set 2 we completely neglected these two features and used combination of I_{RMS}, V_{RMS}, Frequency and Phase angle. The reason for choosing root mean squared (RMS) values for current and voltages is that it has already been reported in [13] that I_{RMS} and V_{RMS} show much higher performance in comparison to I_{AVG} or I_{PEAK}.

In comparison to feature set 1, SVM and *k*-NN classifiers show significant improvement in the recognition accuracy when trained with feature set 1. As for *k*-NN, despite the improvement in recognition accuracy the classes such as fluorescent lamp, laptop, and incandescent lamp are highly overlapped because of similar current and voltage characteristics and hardly

40% of the instances from these classes are correctly classified. We found no significant improvement in performance of k-NN classifier when the value of k has been increased from 1 to 10. On the other hand SVM classifier with polynomial kernel shows superior performance over k-NN. In contrast to it, SVM with RBF kernel show poor performance; unless the cost function is increased to 100. It is evident from Figure 3 that increasing the cost function (C) improves the classification performance for the SVM classifier; however a SVM classifier with a lower cost function is highly desirable.

The increase in the value of C minimizes the training error forcing the decision boundaries to strictly follow the training data which decreases the generalization capability of the classifier. The consequence would be that the classifier would perform poorly if the test data is even slightly different from the training data which is always the case in a real-world scenario. As discussed earlier, we have observed that there is already a lot of variation in the data within each device class, thus a classifier with a good generalization capability is preferred. Therefore we decided to select an SVM model with high recognition accuracy and lower cost function; that is SVM classifier with polynomial kernel function having a cost value of 1.

The selected SVM model in combination with feature set 2, despite achieving overall recognition accuracy above 80%, less than 60% of MAC instances are correctly classified as they are confused with incandescent lamp test samples. Similarly 50% test instances from LCD screen class are misclassified as laptops. For the feature set 3 we carried out experiments by neglecting the frequency feature and found no difference in results because the target appliances selected for this experiment has similar harmonics. Neglecting the frequency feature, we combined P,Q, I_{RMS},V_{RMS}, and phase angle to form feature set 3.

The highest recognition accuracy of 89% is obtained using feature set 3 with SVM classifier (kernel = poly, cost =1) as shown in Figure 3. There is almost an 8% increase in accuracy in comparison to feature set 2, which highlight the fact that redundant feature, which in our case is frequency feature has a direct impact on the classification performance. On the other hand we found no significant improvement in results for k-NN classifier. In order to analyze the recognition results in detail, we have computed the Confusion Matrix (CM) using the best classification model that is SVM classifier (kernel = poly, cost =1) as shown in Table III. The bold values in the diagonal represent the recognition accuracy of respective class, whereas the row values in parallel show how a particular class is confused with other classes. Each of the device class is represented by a class number as shown in Table I. The confusion matrix provides an easy visualization of device classes that are confused during the classification. Out of 9 device classes, 5 of them achieve recognition accuracy above 85% using feature set 3.

However MAC, lamp incandescent, Laptop, and LCD Screen are the 4 device classes that are misclassified by our best classification model. The test instances of MAC device class is confused with incandescent lamp device category whereas LCD screen test instances are confused with laptop and desktop computers as shown in Table III. In order to minimize the confusion amongst identified classes, we apply our new proposed feature set with the SVM classification model as demonstrated in Section IV.C.

TABLE III: Confusion Matrix OF SVM Classifier (KER = POLY, COST = 2, C=1) Using FEATURE SET 3

CM	1	2	3	4	5	6	7	8	9
1	**85.3**	2.4	7.3	0	0	5	0	0	0
2	0	**65.7**	0	0	0	0	10	24.3	0
3	14.8	0	**80.2**	0	0	5	0	0	0
4	0	0.8	0	**99.2**	0	0	0	0	0
5	0	0	0	0	**100**	0	0	0	0
6	0	0	4.4	0	0	**95.6**	0	0	0
7	3.32	0	3.31	0	0	0	**93.37**	0	0
8	0	20.8	0	0	0	0	0	**79.2**	0
9	0	0	0	0	0	0	0	0	**100**

B. AEDM: Acoustic Event Detection in an office Environment

As discussed earlier, the role of AEDM is to perform acoustic monitoring of the environment. We have included user-specific acoustic events together with machines sound in our target sound classes as we have considered an office scenario. It allows us to test the performance of our classification models based on audio feature sets, for recognizing machine sounds in the presence of user-generated sounds.

1)Feature Extraction: We have developed an offline classification system, where MATLAB is chosen as a simulation environment for audio feature extraction and classification. We have decided to include 8 user generated and 4 machine generated acoustical events as target sound classes for our experiment as shown in Table IV.

TABLE IV: Target Acoustical Events

No	**User-specific AE**	**Source**	**Samples**
1	Chair Moving	I	12
2	Clapping	RWCP+I	100+7
3	Cough	I	47
4	Door Slam	I	80
5	Laughter	I	26
6	Music	I	38
7	Sneeze	I	40
8	Speech	ShaTR	52
	Machine Sounds	**Source**	**Samples**
9	Keyboard typing	I	45
10	Printer Active Mode	I+R	20
11	Mouse Click	I+R	13
12	Scanner Active Mode	I+R	20

There is a lack of sound data especially for machine sounds, therefore we have not obtained samples from the internet (I), ShATR Multiple Simultaneous Speaker Corpus [14],as well as we have recorded machine sounds using a microphone (indicated by R) as shown in Table IV. Only clapping sound samples belong to Real World Computing Partnership (RWCP) sound scene database [14]. The right most column of Table IV indicates the number of samples included in the sound database. Each sample has duration of almost 1 minute. In the pre-processing step we downsampled sound data to 8 kHz, normalized it in range of [-1 1] and frame-based segmentation is performed (frame length = 128, overlapping 50% Hamming window). We have removed the silence portion and further extracted the audio features within each frame. The time-domain features include Zero Crossing Rate (ZCR), Short Time Energy (STE), Fundamental Frequency (F), and Sub-band log energies (SLE). The details of each of them can be found at [14]. As for frequency based audio features we have computed 12 mel-frequency cepstral coefficients (MFCC) for each frame. The application of a second order filter $H(z) = z - z^{-1}$at the output of mel-scaled filter bands provide us with Frequency Filter Band Energies (FFBE) which is found out to be more discriminative than MFCC for the task of speech recognition [15]. We have combined time-domain and FFT based features into three feature sets, whereas content and size of each feature vector is shown in Table V.

TABLE V: Audio Feature Set

No	Feature Set	Contents	Size
1	F1	ZCR+STE+F+SLE	8
2	F2	E+MFCC	13
3	F3	FFBE	13

2)Performance Evaluation of Audio Feature Sets: In order to evaluate the performance each of the audio feature sets listed in Table V, we trained SVM and Gaussian Mixture Models (GMM) for classification of target acoustical events. For SVM, after experimentation RBF was found out to be the optimal kernel function with a G value of 0.001. GMM on the other hand is widely used in applications such as speech and music recognition. GMM tries to estimate the underlying probability density functions of the observations assuming that they can be modeled with mixture of Gaussians. We have tried fixed and variable number of Gaussians per class and the best results were achieved using variable mixture components for each sound class as the amount of data in each class is different. The detail description of GMM algorithm can be found at [16]. Figure 6 clearly shows that, for all three feature sets SVM classification model outperforms GMM.The highest overall recognition accuracy achieved is 89% using frequency based feature set F3 in combination with SVM classifier.Despite achieving high overall recognition accuracy, it is not a true indicator of classifier performance due to data unbalance issue as discussed earlier.

Therefore, in order to analyze the results in detail,

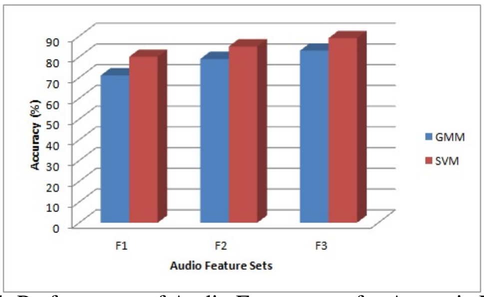


Fig. 4: Performance of Audio Feature sets for Acoustic Event Detection

we have computed CM and found out that the cough, laughter and sneeze are the most confused classes for the category of user-specific acoustical events. The feature set F3 however performs well in comparison to F1 and F2 in recognizing these confused classes. A comparison of GMM and SVM based classification models for the task of detecting user generated sounds has also been reported in [14], however our experimental evaluation also considers the presence of machine sounds within the environment. Albeit, SVM still outpeforms GMM in our scenario as well.In Figure 5, we report the performance of each feature set using SVM classification model for the task of identifying these machine specific acoustical events.

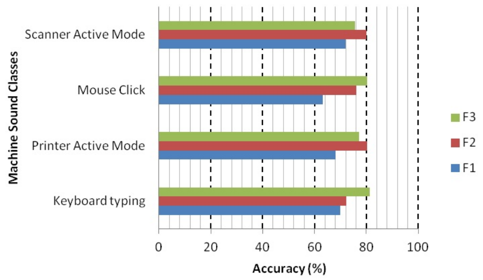


Fig. 5: Machine Sound Recognition results using SVM Model

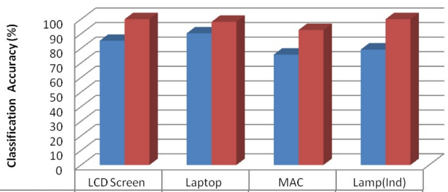
Although, F3 achieve an overall high recognition accuracy for all the target sound classes, however we can clearly see from Figure 5 that in comparison to F3, printer and scanner sounds are more accurately recognized using feature set F2. In contrast to it F3 performs much better than F2 and F1 for recognizing keyboard typing and mouse click sound. From this, we can conclude that not every feature sets works best for each sound class. The recognition of these machine sound will faciliate the device detection method as discussed below.

C. Decision Layer: Feature Fusion of Device and Acoustic Features

In the decision layer, we further improve the classification accuracy of those low-power consumer appliances which are hard to reliably classify due to overlapping of steady-state load signatures as demonstrated in Section IV.A. The experimental results obtained in the DDM showed that MAC, Laptop, LCD Screen and Incandescent lamp are the most confused classes. Two of the devices that include MAC and Laptop have a common acoustic signature that is the sound of keyboard typing. In this experimental evaluation, we assume that whenever the user is using a laptop or a MAC device he will generate a keyboard typing sound. Our AEDM can detect and classify this machine sound as already demonstrated in Section IV.B. The decision module can reevaluate the result obtained from the DDM by correlating the machine sound with the device features. Therefore, we propose to generate a new feature set that combines the audio feature with the device feature acquired from DDM and AEDM respectively.. From our previous experiment, SVM was found out to be the best classifier in the DDM, therefore we decided to train the same classifier with our new feature set F_{AD} that is defined as

F_{AD} = Feature set 3 (DDM) + F3(AEDM) (1)

The new feature set F_{AD} combines the best feature set selected from the device recognition and the acoustic event recognition experiments respectively.



	LCD Screen	Laptop	MAC	Lamp(Ind)
Feature Set 3	85.2	90.3	75.7	79.2
Feature Set AD	100	98.2	92.6	100

Fig. 6: Comparison of F_{AD} and Feature Set 3 for classification of Confused classes using SVM(ker = poly, exp =2, cost =1)

The fusion of acoustic feature with device feature easily separates the MAC and Laptop from incandescent lamp and LCD screen respectively in the feature space. We compared the performance of F_{AD} with feature set 3 which is found out to be the best steady-state feature set in the DDM.It is quite evident from the results shown in Figure 6 that our proposed feature set outperforms feature set 3 ; increasing the recognition accuracy of the confused classes almost by 16%.

V. Conclusion And Future Work

This paper has discussed a multi-layer decision framework for smart energy sensing that improves the device recognition accuracy of low-power consumer appliances by combining steady-state load features with audio features derived from the device usage. We investigated the use of time-domain and FFT based audio feature sets for recognizing acoustic activity within an office environment. We found out that FFBE are more discriminative than MFCC based audio features for most of the target sound classes used in the experiment, however few of the machine sounds are best recognized by MFCC based audio feature. SVM was found out to be the best classification model for both audio and device recognition tasks. In future, we will remove our assumptions made in the decision layer, and implement reasoning strategies in the inference layer for device usage context recognition.

Acknowledgment

We acknowledge the support from the REDUCE project grant (no:EP/I000232/1) under the Digital Economy Programme run by Research Councils UK - a cross council initiative led by EPSRC and contributed to by AHRC, ESRC, and MRC. This research is done during Dr S Rajasegarar's research visit to the University of Surrey, UK from the University of Melbourne, Australia.

References

[1] J. Uteley and L. Shorrock. Domestic energy fact file 2008. [Online]. Available: http://www.bre.co.uk/page.jsp?id=879

[2] G. W. Hart, "Nonintrusive appliance load monitoring," *IEEE Proceedings*, vol. 80, no. 12, pp. 1870–1891, 1992.

[3] Y. Du, L. Du, B. Lu, R. Harley, and T. Habetler, "A review of identification and monitoring methods for electric loads in commercial and residential buildings," in *Energy Conversion Congress and Exposition (ECCE),IEEE*, sept. 2010, pp. 4527 –4533.

[4] A. I. Cole and A. Albicki, "Data extraction for effective non-intrusive identification of residential power loads," in *IEEE Conference Proceedings on Instrumentation and Measurement Technology Conference (IMTC98)*, vol. 2, 1998, pp. 812–815.

[5] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox, "Nonintrusive Load Monitoring and Diagnostics in Power Systems," *IEEE Transactions On Instrumentation And Measurement*, vol. 57, no. 7, pp. 1445–1454, 2008.

[6] M. Hazas, A. Friday, and J. Scott, "Look back before leaping forward: Four decades of domestic energy inquiry," *IEEE Pervasive Computing*, vol. 10, no. 1, pp. 13 –19, 2011.

[7] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. Culler, "Design and implementation of a high-fidelity AC metering network," in *International Conference Proceedings on Information Processing in Sensor Networks (IPSN 09)*. IEEE Computer Society, 2009, pp. 253–264.

[8] L. K. Norford and S. B. Leeb, "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms," *Energy and Buildings*, vol. 24, no. 1, pp. 51–64, 1996.

[9] A. Schoofs, A. Guerrieri, D. Delaney, G. O'Hare, and A. Ruzzelli, "ANNOT: Automated electricity data annotation using wireless sensor networks," in *7th Annual IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks (SECON)*, june 2010, pp. 1 –9.

[10] Z. C. Taysi and M. A. Guvensan, "Tinyears : Spying on house appliances with audio sensor nodes," *The Proc. of 2nd ACM Workshop on Embedded Sensing Systems for Energy Efficiency in Building*, pp. 31–36, 2010.

[11] M. Cowling, "Non-speech environmental sound classification system for autonomous surveillance," Ph.D. dissertation, Faculty of Engineering and Information Technology, Griffith University, 2004.

[12] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, p. 43, 1998.

[13] T. Kato, H. S. Cho, and D. Lee, "Appliance Recognition from Electric Current Signals for Information-Energy Integrated Network in Home Environments," *Ambient Assistive Health and Wellness Management in the Heart of the City*, vol. 5597, pp. 150–157, 2009.

[14] A. Temko and C. Nadeu, "Classification of acoustic events using svm-based clustering schemes," *Pattern Recogn.*, vol. 39, pp. 682–694, April 2006.

[15] C. Nadeu, J. Hernando, and M. Gorricho, "On the decorrelation of filter-bank energies in speech recognition," in *Proc. Eurospeech*, 1995, pp. 1381–1384.

[16] J. A. Bilmes, "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," *International Computer Science Institute*, vol. 4, no. 510, p. 126, 1998.