Developing User Centric HEMS through Automated Appliance Recognition Framework

Daphne H.Z. Tang, Yewguan Soo

Faculty of Electronics and Computer Engineering (FKEKK)
Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
daphne.tang@hotmail.com

***Abstract*—Home Energy Management Systems (HEMs) have been proven to help home users manage their power consumption and improve usage habits. With more advanced HEMs incorporating appliance recognition technology to enable tracking of appliances via its unique electrical signature, there still exists the drawback of requiring complex yet time consuming appliance registration stages. To curb this problem, this paper presents the framework required to automate the appliance registration process to create a much more user centric system. By demonstrating the working of the framework using one-class support vector machine with additional principal component analysis feature extraction using 10 household appliances, the classification rate of unregistered appliance into its rightful class was 100% with a recall rate of 67.04% for registered appliances. The results were obtained based on leave-one-out cross validation technique, excluding the results of the training dataset.**

***Index Terms*—Appliance Recognition, Energy Management System, One-class Support Vector Machine, Principal Component Analysis, User Centric Systems.**

I. Introduction

Ever since the green-house effect was proven at the end of 1980s, electricity conservation has gained public interest. Electricity production and usage caters to greenhouse gas emission, especially the carbon dioxide gas (CO_2). According to [1], electricity production accounts to 38% of CO_2 emitted compared to the four other sectors. Out of the produced electricity, 40% is used to power the residential and commercial buildings.

While a researcher identified several main sources of wasting electrical energy that could account up to 41% in total expenditures within a residential household [2], the capability for improvement in home power management is highly prominent with the estimation of savings of up to 43% with installation of home energy management systems (HEMs) [3]. Unlike basic energy monitors that display only immediate readings, HEMs encapsulates the ability to monitor and control power supply of singular home appliances whilst providing users with alert features.

However, several studies show that power saving using HEMs was only temporary from a few weeks to four months [4], [5]. Nevertheless, the saving potential of the HEMs was found to be significantly higher compared to using energy monitors, which may only achieve 5-15% savings only [6]. The difference in potential of power saving monitoring and management systems highlight the importance of user centric design in achieving efficient power management.

Recently, the addition of appliance recognition functions into HEMs has allowed the possibility of tracking and controlling home appliances according to specific appliance signatures. This means that appliances are able to be moved around the premises, and it can still be correctly controlled by the system. This drastically improves the applicability of HEMs in consumers' households as it minimizes pre-configuration when appliances are shifted. However, systems with appliance recognition still require initial appliance registration, which is done manually by the home user. Every appliance that the user wants the system to recognize has to be registered into the system. This can be done by operating the appliance a few times so that the characteristics of the appliance can be recorded [7], [8]. For an average consumer, the adaptation of this technology can be very confusing and difficult to setup.

To solve this problem, [9] applied a one-class support vector machine (OCSVM) to incorporate a systems' decision making process in identifying new unregistered appliances. Upon the identification of new appliances, the system would be able to automatically acquire appliance characteristics for the appliance training process and classify the new appliance. The author used the Principal Component Analysis (PCA) to extract useful components that reduce the dimensions of appliance characteristics. 24 appliances were used in their experiment. Using the leave-one-out cross-validation, 97.7% classification accuracy of unregistered appliances has been achieved. Errors were mainly caused by appliances with near similar waveforms.

II. Objectives

Even though researchers from [9] implemented OCSVM for automating appliance recognition, the framework required for its inclusion is still unclear. This paper describes the framework required for automating appliance training process in a system performing appliance recognition. Moreover, the paper presents experiment results in classifying the registered and unregistered appliances using PCA feature extraction and OCSVM on ten carefully selected appliances.

III. SYSTEM FRAMEWORK

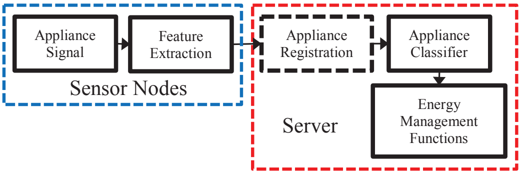


Figure 1: Automated Appliance Training Block Diagram

With user centricity in mind, the system design was aimed at providing an independent platform, which will be easily scalable for future system advancements. The hardware setup is similar to most HEMs that consist of sensor nodes and a server. However, the novelty of the model focuses on how human aided appliance training could be exempted from the setup process of HEMs ([8], [10]–[12]) using an additional 'appliance registration' block as shown in Figure 1.

One-class classification model is implemented inside the 'appliance registration' block to help the system identify new appliances introduced into the system. It also helps to register its features for future references. Feature extraction is applied as it helps to extract useful hidden components from the appliance's signal to eliminate common similar components available in the appliance's signal. The extraction of features does not only reduce the dimension of incoming data, but it also increases the processing speed in the system as the amount of data being processed is reduced.

IV. AUTONOMOUS APPLIANCE TRAINING

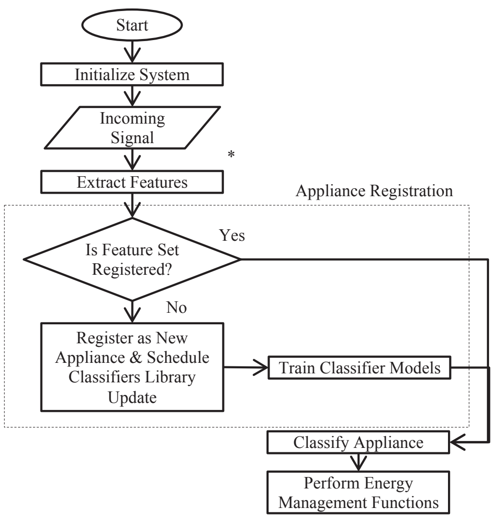


Figure 2: Flowchart Showing an Automatic Appliance Training Framework

Figure 2 describes how the 'appliance registration' block automates the appliance training process. The framework is applied in the server side of the system. After the feature extraction, the data is sent to a decision model (in this case, OCSVM) which will determine if the data has been registered before. If the data belonged to a registered appliance, the system shall proceed to classify and identify the specific appliance. On the other hand, if the data was not registered, the system will register it using the acquired features and schedule to update the classifier models in the system. A typical setup of such framework would only require an OCSVM model for registrative decisions. The multi-class classifier used after the registration serves only to classify the signals into specific appliances. It and may use algorithms, such as the k-Nearest Neighbor, SVM, artificial neural network or etc.

The process of feature extraction shown in Figure 2 may also be performed in the hardware level and placed before the 'incoming signal' stage to decrease transmission traffic in the sensor network. Performing the feature extraction after receiving appliance waveforms at the server level increases the amount of data to be transmitted from the sensor nodes, leading to a drastic increase in the network traffic.

V. RESEARCH METHODOLOGY

To demonstrate the operation of the framework, this section presents the experiment performed in this study. The study was performed in a 240 V_{AC}, 50Hz powered premise using 10 household appliances. The data processing techniques consist of feature extraction from the appliance signals and the classification using the OCSVM, and it was adopted to determine the registered and unregistered appliances.

A. Data Acquisition and Organization

Ten appliance signals were obtained using an electronic oscilloscope operating at 5 kHz clamped to a current transformer. Table I shows the obtained appliance information together with its indexes.

For validation purposes, the collection of data was broken down into 16 datasets, where every dataset consists of 100 unique waveform samples per appliance for all 10 appliances.

Table 1
Appliance Information

Index	Type	Rated Power (W)	Irms (A)
1	Iron 1	1550	6.425
2	Iron 2	1000	4.077
3	Refrigerator	200	0.713
4	Hair dryer (low)	500	1.909
5	Hair dryer (high)	1200	4.773
6	Kettle	1400	5.540
7	LCD	45	0.178
8	Mat burner	8	0.032
9	Oven	1400	5.575
10	Washing machine	330 4 (standby)	0.013 (standby)

B. Feature Extraction

In this study, we perform PCA on the acquired appliance waveforms to obtain the two most significant components. PCA is an eigenvector based multivariate analysis method that helps to reduce data dimensions to a few that best describes the data. We perform PCA from the MATLAB library on non-normalized appliance waveforms (refer to Figure 3) to take into account the large variation of current draw in the different appliance.

In PCA, eigenvalues show the significance of every calculated eigenvector. Two eigenvectors with the highest eigenvalues were obtained by performing the PCA with waveforms from dataset 1 as shown in algorithm 1. These two eigenvectors act as the feature extraction multiplier for all types of appliances to reduce the dimensions of original waveforms into two dimensions only. Hardware wise, these eigenvectors should be preset in the sensor nodes, where the feature extraction is performed. For the experiment, the eigenvectors were multiplied with the appliance waveforms throughout dataset 1 to 16 to reduce the dimensions from 100 data points to 2 data points per waveform. The value of these two data points are the direct scalar values of the two principal components calculated as shown in algorithm 2.

Input: Appliance waveforms in dataset 1
Output: Two eigenvector components

Collect one dataset of appliance waveforms stored in a *p*-by-*p* matrix, **X**, where every waveform is stored in a column.
Compute Principal Component Analysis to obtain a *p*-by-*p* matrix, **Y** where its columns are eigenvectors of $\mathbf{X}^T\mathbf{X}$.
Select two columns from **Y** that have the highest eigenvalues as $\mathbf{Y_2}$.

Algorithm 1: Pseudocode for extraction of two most significant PCA eigenvectors

Input: A single appliance waveform
Output: Two principal components

Acquire appliance waveform, **Z**, which its values are stored in a 1-by-*p* vector.
Acquire the two most significant principal components, $\mathbf{T} = [\mathbf{t_1}, \mathbf{t_2}]$, where $\mathbf{T} = \mathbf{Z} \times \mathbf{Y_2}$.

Algorithm 2: Pseudocode for obtaining principal components values from appliance waveforms

C. One-class Support Vector Machine

OCSVM works by training the model with data from a known class, such as the class of registered appliances. Our study utilizes OCSVM from the LIBSVM library for MATLAB developed by [13]. The radial basis function (RBF) kernel was selected to accommodate the multi-modal dataset with different appliance data. RBF kernel is widely used as it is easy to tune with one width parameter only, γ as shown in Equation 1.

$$K(x_i, y_j) = e^{-\gamma\left\|x_i - x_j\right\|^2} \qquad (1)$$

Tuning of the width parameter was performed empirically to obtain the best results. While dataset 1 was used to train the model, only 9 out of 10 appliances were used for each run. This method corresponds to leave-one-out cross validation in which every appliance is left out once. The final accuracy was calculated based on the results provided by dataset 2 to 16, leaving out the results of the training set.

D. Performance Index

Classification using OCSVM produces results that are clearly stated on a confusion matrix. By analyzing the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) rates, the recall (eq. 2) and precision (eq. 3) values of the classification can be known. As the study uses leave-one-out cross-validation, 90% of the appliances in a dataset belong in the registered (positive) class while the other 10% belongs in the unregistered (negative) class.

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

Based on the confusion matrix results, TP represents correct classification of the registered appliance class, while TN represents correct classification of the unregistered class. FP represents wrong classification of the registered class, while FN represents wrong classification of the unregistered class.

While the aim is to correctly identify unregistered appliances, emphasis is put on FP rate (Type I error) where it is ideally 0%, showing that all unregistered appliances were detected. Analysis of FP result is directly reflected in the percentage of precision rate which is ideally 100%. Apart from that, FN rate (Type II error) is considered less critical as it can be mended by assuming that the previous training samples were underfitted; hence, more training samples are required to improve the classification accuracy. This way, the detection of a registered signal as an unregistered signal serves to increase the training sample of that appliance in the database of the OCSVM model.

VI. RESULTS

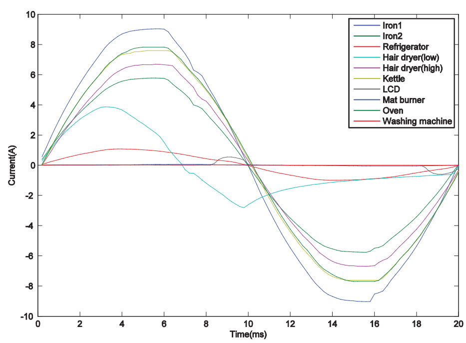


Figure 3: Appliance Waveform in a Single AC Cycle

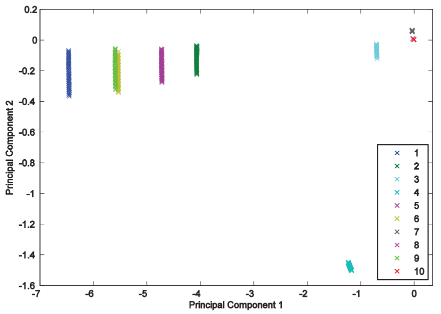


Figure 4: Feature Map of Waveforms in Dataset 1 after performing PCA

Figure 3 shows the waveforms of 10 appliances used in this study. As shown, six of the appliances have sinusoidal waveform which acts to test the classifier in telling apart appliances with near similar waveforms. The remaining appliances have power ratings ranging down to 4 Watts.

After the eigenvector transformation, the two most significant components of appliance waveforms were mapped on a 2-dimensional plot as shown in Figure 4. The index used in the figure corresponds to the appliance index in Table I. From the figure, it can be observed that the feature scatters of kettle and oven are situated very closely, though they are not overlapped. Further, low powered appliances, such as the LCD, mat burner, and washing machine (sleep mode) are situated close together near the axis origins.

Figure 5 shows the decision boundary created with the RBF kernel in OCSVM using gamma, γ of 407. The value was obtained by accessing the average classification accuracy as shown in Figure 6 in which the gamma of 407 showed 100% accuracy in classifying unregistered appliances (TN value). The training model achieved an overall training accuracy of 95.39% before performing validation. Upon training of the OCSVM model using features from dataset 1, the average classification accuracy (excluding training dataset) was 70.33% when the data were classified into registered and unregistered appliance class.

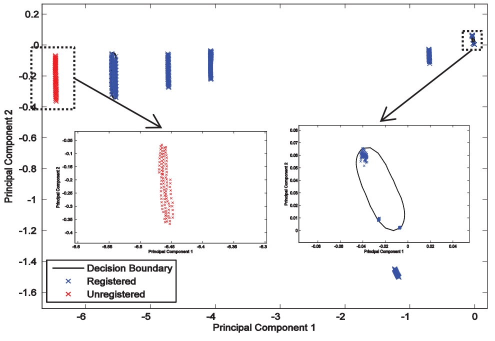


Figure 5: Boundary Fitting In OCSVM Model Using Training Data

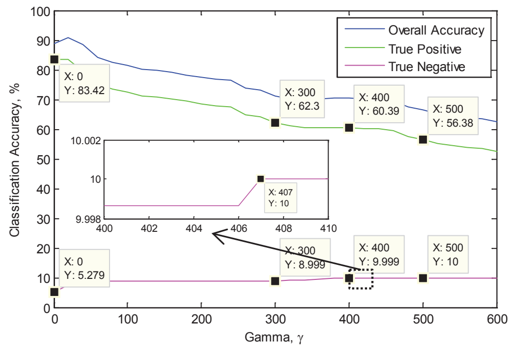


Figure 6: Graph of Gamma, γ versus Classification Accuracy

From the confusion matrix in Figure 7, the calculated precision rate was 100% while the recall rate was 67.04%. This means that there were no misclassified unregistered appliances, and all of the 'new' appliances were detected by the OCSVM model. On the other hand, there was 29.67% of misclassified registered appliance, which the system thought was unregistered even though the signal was previously registered. With 100% classification of unregistered data, OCSVM is proved to work well with the PCA in the application of the proposed framework.

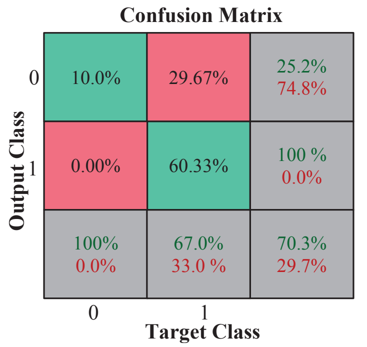


Figure 7: OCSVM Validation Confusion Matrix

VII. Discussion

A. *Preprocessing of appliance signals*

The experiment utilized original (non-normalized) appliance waveforms as opposed to [9] for the classification of unregistered appliances. While [9] correctly classified 97.7% of the unregistered appliance data, our experiment achieved 100% accuracy using the original waveforms. This result highlights the importance of appliance waveform amplitude in distinguishing the different appliances.

In our study, the values of principal components were found to be based on amplitude and shape. This can be observed with the scatters of the kettle and oven being so close together even though both had similar power ratings and profile shape (Figure 8). By emphasizing both the amplitude and shape of appliance profiles, the feature scatters of all 10 appliances were found to be clearly grouped without overlapping the class areas.

Using the same appliance library, when PCA was performed on normalized appliance waveforms, the resulting feature scatters showed various closely situated groups especially in the 'zoomed in' area in Figure 9. Those appliances were found to belong to those with sinusoidal waveforms shapes.

As shown in the confusion matrix in Figure 10, the overall classification accuracy of the normalized PCA was higher by 23.4% compared to the non-normalized PCA in Figure 7. The processing of the normalized waveforms also helped to achieve a precision rate of 95.5% and a recall rate of 97.6%. While the recall rate improved, the precision rate decreased as 41.8% of the unregistered appliances were misclassified. With the large amount of undetected unregistered signals, it can be concluded that the normalized PCA is less suitable for the application of detecting unregistered appliances.

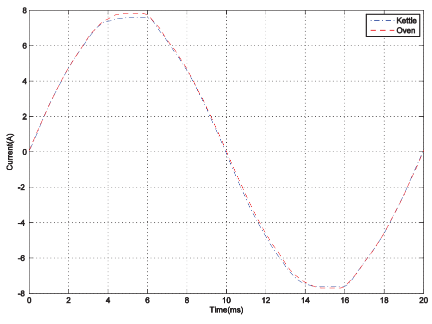


Figure 8: Graph of Kettle and Oven Waveforms

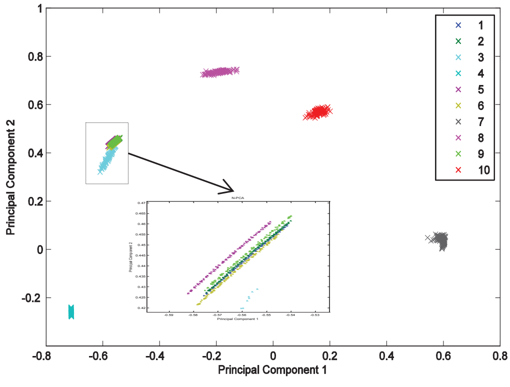


Figure 9: Feature Map of Normalized Waveforms after PCA

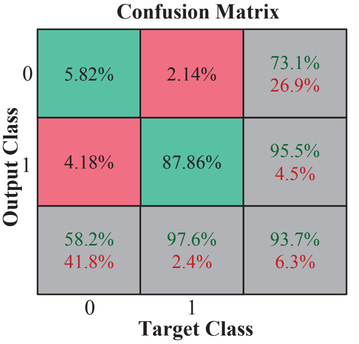
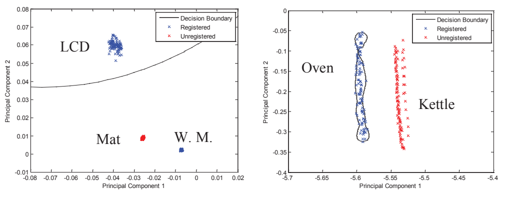


Figure 10: OCSVM Validation Confusion Matrix for Normalized Waveforms

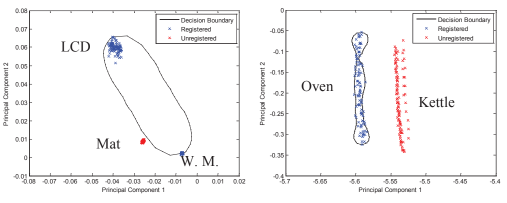
From the results, we find that normalization deletes the amplitude information in those waveforms; thus causing difficulty in differentiating appliances with near similar shapes. However, normalization helps emphasize differences of low powered appliances as it increases the importance of waveform shapes. Overall, normalization works well when classifying appliances is based on waveform shapes, while cancelling out the relevance of its amplitude.

B. *Value of γ*



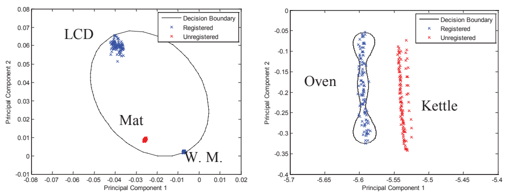
(a) (b)

Figure 11: Gamma, γ = 600



(a) (b)

Figure 12: Gamma, γ = 407



(a) (b)

Figure 13: Gamma, γ = 200

This section discusses the effect of γ on classification accuracy. Differences of decision boundary width between 3 gamma values are shown in Figures 11 to 13. While gamma of 600 caused exclusion of the washing machine scatters situated below in Figure 11(a), gamma of 200 underfits the scatters and creates a wide decision boundary which unintentionally includes the scatters of the unregistered class (Mat burner) and causes misclassification. On the other hand, gamma variation does not show much effect towards the decision boundary of the oven, which in all three cases did not misclassify the nearby scatters of the kettle. From Figure 6, gamma of 407 shows the optimum result in preventing type I classification error (FP) while giving 29.67% type I error (FN).

VIII. CONCLUSION

This paper presents the study of creating a user centric home energy management system by adding in an automated appliance registration block into systems that uses appliance recognition technology. The adaptation of the framework was presented with additional experimentation of the decisive model which uses OCSVM to detect unregistered appliances. Results from the study show a successful implementation of PCA feature extraction technique with the OCSVM, achieving 100% precision in detection of unregistered appliances via leave-one-out cross validation of ten household appliances. However, 29.67% of the registered appliances were misclassified, resulting in a recall rate of 67.04%. Other feature extraction techniques could be tested in the future to improve this result.

ACKNOWLEDGMENT

This work was supported in part by the Centre for Research and Innovation Management (CRIM) of Universiti Teknikal Malaysia Melaka (UTeM) under grant PJP/2012/FKEKK/ (13B)/ S01017.

REFERENCES

[1] C. C. D. US EPA, "U.S. Greenhouse Gas Inventory Report." [Online]. Available: http://www.epa.gov/climatechange/ghgemissions/usinventoryreport.html . [Accessed: 14-Oct-2013].

[2] E. Williams and H. Matthews, "Scoping the potential of monitoring and control technologies to reduce energy use in homes," in *IEEE International Symposium on Electronics & the Environment*, 2007, pp. 239–244.

[3] R. Teng and T. Yamazaki, "Construction and Experimental Evaluations of User-Centered Power Consumption Management Systems in Home Environments," *Int. J. Energy, Inf. Commun.*, vol. 2, no. 1, pp. 1–15, 2011.

[4] S. S. van Dam, C. a. Bakker, and J. D. M. van Hal, "Home energy monitors: impact over the medium-term," *Build. Res. Inf.*, vol. 38, no. 5, pp. 458–469, Oct. 2010.

[5] M. a. Alahmad, P. G. Wheeler, A. Schwer, J. Eiden, and A. Brumbaugh, "A Comparative Study of Three Feedback Devices for Residential Real-Time Energy Monitoring," *IEEE Trans. Ind. Electron.*, vol. 59, no. 4, pp. 2002–2013, Apr. 2012.

[6] D. S. Parker, "Pilot Evaluation of Energy Savings from Residential Energy Demand Feedback Devices," Florida, 2008.

[7] M. Baranski and J. Voss, "Genetic Algorithm for Pattern Detection in NIALM Systems," in *IEEE International Conference on Systems, Man and Cybernetics*, 2004, no. 2, pp. 3462–3468.

[8] A. G. Ruzzelli, C. Nicolas, A. Schoofs, and G. M. P. O'Hare, "Real-Time Recognition and Profiling of Appliances through a Single Electricity Sensor," *2010 7th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Networks*, pp. 1–9, Jun. 2010.

[9] T. Kato, H. Cho, D. Lee, T. Toyomura, and T. Yamazaki, "Appliance recognition from electric current signals for information-energy integrated network in home environments," in *7th International Conference on Smart Homes and Health Telematics, ICOST*, 2009, pp. 150–157.

[10] M. Dong, P. C. M. Meira, W. Xu, and W. Freitas, "An Event Window Based Load Monitoring Technique for Smart Meters," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 787–796, 2012.

[11] C. Belley, S. Gaboury, B. Bouchard, and A. Bouzouane, "An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances," *Pervasive Mob. Comput.*, Mar. 2013.

[12] T. Saitoh, Y. Aota, T. Osaki, R. Konishi, and K. Sugahara, "Current sensor based non-intrusive appliance recognition for intelligent outlet," in *ITC-CSCC*, 2008, pp. 349–352.

[13] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–39, 2011.