An Intelligent System for Mining Usage Patterns from Appliance Data in Smart Home Environment

Yi-Cheng Chen, Yu-Lun Ko and Wen-Chih Peng
Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan 300
E-mail: ejen.cs95g@nctu.edu.tw jimmy0301.cs99g@nctu.edu.tw wcpeng@cs.nctu.edu.tw

***Abstract*—In the last decade, considerable concern has arisen over the electricity saving due to the issue of reducing greenhouse gases. Previous studies on usage pattern utilization mainly are focused on power disaggregation and appliance recognition. Little attention has been paid to utilizing pattern mining for the target of energy saving. In this paper, we develop an intelligent system which analyzes appliance usage to extract users' behavior patterns in a smart home environment. With the proposed system, users can acquire the electricity consumption of each appliance for energy saving easily. In advance, if the electricity cost is high, users can observe the abnormal usage of appliances from the proposed system. Furthermore, we also apply our system on real-world dataset to show the practicability of mining usage pattern in smart home environment.**

Keywords- abnormal detection;energy saving;usage pattern; smart home

I. INTRODUCTION

Recently, concern over global climate change has motivated efforts to reduce the emissions of CO2 and other GHGs (greenhouse gases). Many researchers focus on the reduction of electricity usage, especially, the residential sector which is a significant contributor of GHGs. With the consideration of electricity saving, we also can reduce the generation of GHGs. However, it is not easy for residents to conserve the electricity, since, in general, residents only can know the information of electricity usage from the electricity bill which reveals the total electric cost. In fact, they do not know the electric consumption of each appliance in the house. If the electricity bill is expensive this month, we only can know it is expensive instead of "why" it is expensive.

To the best of our knowledge, few studies utilized the behavior pattern to detect abnormal user behavior for target of energy saving. Previous researches of usage patterns mainly focus on energy disaggregation [2, 8, 9, 12] and appliance recognition [3, 4, 5, 6, 11]. Kim et al. [8] investigated the effectiveness of several unsupervised disaggregation methods on low frequency power measurements collected in real homes. Author proposed a usage pattern which consists on-duration distribution and dependency between appliances. Farinaccio et al. [3] used the pattern, such as, number of data point and ON-OFF switch, to disaggregate whole-house electricity consumption into its major end-uses. Suzuki et al. [12] use new NIALM technique based on integer programming to disaggregate residential power use. Lin et al. [9] used a dynamic Bayesian network and filter to disaggregate the data online. Goncalves et al. [4] explored an unsupervised approach to determine the number of appliances in the household, including their power consumption and state, at any given moment. Chen et al. [2] disaggregated utility consumption from smart meters into specific usage that associated with human activities. Authors proposed a novel statistical framework for disaggregation on coarse granular smart meter readings by modeling fixture characteristic, household behavior, and activity correlations.

Prudenzi [11] utilized an artificial neural network based procedure for identifying the electrical signatures of residential appliances. Ito et al. [5] extract features from the current (e.g., amplitude, form, timing) to develop appliance signatures. For appliance recognition, Kato et al. [6] used Principal Component Analysis to extract features from electric signals and classified them by Support Vector Machine. Some of these works and the characteristics of workable solutions were discussed by Matthews et al. [10].

In this paper, we develop an intelligent system, **HAUBA** (**H**ousehold **A**ppliance-**U**sage **B**ehavior **A**nalysis), which analyzes appliance usage to extract users' behavior patterns in a smart home environment. With the proposed two usage patterns, users can acquire the information of power consumption and representative behaviors of each appliance. In advance, if the electricity cost is high, from HAUBA, users can observe the abnormal usage of appliances for conserving electricity easily, as shown in Fig. 1.

Fig. 1: User interface on a smart phone of proposed system

II. USAGE PATTERN ANALYSIS

In this section, we will introduce two types of usage patterns, time-slot probability usage pattern (TPUP) and

978-0-7695-4919-4/12 $26.00 © 2012 IEEE
DOI 10.1109/TAAI.2012.54

daily behavior-based usage pattern (DBUP). We discuss two types of patterns and corresponding mining algorithms in detail as follows.

Definition 1 (usage-point and usage-point log)
Let $\mathfrak{I} = \{ON, OFF\}$ be the set of states. Without loss of generality, we define a set of uniformly spaced time points based on the natural number N. We say the pair $(s_i, t_i) \in \mathfrak{I} \times N$ is an usage-point, where $s_i \in \mathfrak{I}$, $t_i \in N$. The usage-point log of an appliance is $U_P = \langle D_1, D_2, \ldots, D_n \rangle$, where D_i is the daily usage-point sequence of the appliance, $1 \leq i \leq n$. D_i is a sequence of usage-point, i.e., $D_i = \langle (s_{i1}, t_{i1}), (s_{i2}, t_{i2}), \ldots, (s_{im}, t_{im}) \rangle$.

Given a time slot size Z, we partition a day into h slots. The time slot probability usage pattern is defined as a probability sequence of the probabilities of each time slot, $\langle P_1(ON), P_2(ON), \ldots, P_h(ON) \rangle$. The pseudo code of mining TPUP is as shown in Fig. 2. For extracting time slot probability usage pattern, we partitions each day into h time slots and the size of each time slot is z (line 1, algorithm 1). Then, we evaluate ON probability of each time slot for n days data. The duration time T_{ij} (the j_{th} slot in the i_{th} day) is summed together (line 5, algorithm 1), divided by time slot size S, and multiplied by N (line 6, algorithm 1). For example, the time slot size S is 3 hours, so the number of time slots h is 8 (i.e., 24/3 =8). The time slots are 0:00~03:00, 03:00~06:00, 6:00~09:00 ...etc. Evaluating the ON probability of the first time slot, we sum the usage duration T_{i1}, where i is from 1 to n. Then, the summation is divided by $3n$.

Algorithm 1: TPUP Miner (U_p)

Input: An usage-point log: $U_p = \langle D_1, D_2, \ldots, D_n \rangle$ where D_i is the daily usage-point sequence $\langle (s_{i1}, t_{i1}), (s_{i2}, t_{i2}), \ldots, (s_{im}, t_{im}) \rangle$, time-slot size: z
Output: A probability sequence: $\langle P_1(ON), P_2(ON), \ldots, P_h(ON) \rangle$

```
01: h ← 24/z ;
02: for j = 1 → h do
03:     for i = 1 → n do
04:         T_ij ← get the duration time of jth time-slot in D_i ;
05:         Slot_Duration[j ] ← Slot_Duration[j ] + T_ij ;
06:     P_j(ON) ← Slot_Duration[j ] / (z*n);
07: output ⟨P_1(ON), P_2(ON), …, P_h(ON)⟩;
```

Fig. 2: The pseudo code of TPUP mining algorithm

TPUP only reveals the usage time distribution of the appliance instead of the general usage behaviors of users. In this paper, we propose another usage pattern, daily behavior-based usage pattern (DBUP) to describe the representative usage behavior. The concept of DBUP is shown as in Fig. 3. For extracting DBUP from an usage point log, first, we treat a daily usage-point sequence as a daily behavior of an appliance. Then, similar daily usage behaviors are clustered in the same group. A hierarchical cluster method is proposed to group similar behaviors together. Finally, we evaluate the centroid behaviors of each cluster and output as the daily behavior-based usage patterns.

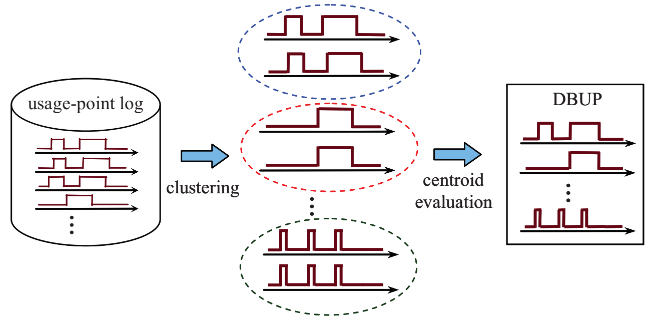

Fig. 3: The concept of daily behavior-based usage pattern

An efficient method, DBUP Miner, is developed for mining daily behavior-based usage patterns. The pseudo code is elaborated in Fig. 4. DBUP Miner first calls sub-procedure Hierarchical_Cluster to enumerate the clusters of all daily usage-point sequences (line 1, algorithm 2). Hierarchical_Cluster is designed to cluster the similar daily usage behavior. At first, each daily usage-point sequence is considered as a cluster (line 4, algorithm 2). To cluster usage-point sequences (i.e., 0 and 1 time series sequences), we need to measure similar among all input time series sequences. In the clustering process, each piece of time series data can be viewed as a point located in an abstract space, and the distances between these points are usually figured by similarity function. A time series similarity qualifies the distance between the sequences of time series data as points in the clustering space. One important point for similarity function is that the time-shifting constrain needs to be considered, i.e., the range of local time shift should be limited. In this paper, we adopt EDR as the similarity function that can deal with local time shifting under a time shifting constrain, and can which deal with noise, but which does not allow too much amplitude shifting.

Algorithm 2: DBUP Miner (U_p)

Input: An usage-point log $U_p = \langle D_1, D_2, \ldots, D_n \rangle$, where D_i is the daily usage-point sequence $\langle (s_{i1}, t_{i1}), (s_{i2}, t_{i2}), \ldots, (s_{im}, t_{im}) \rangle$
Output: A set of usage-point sequences $\{S1, S2, \ldots\}$

```
01: CU_p ← Hierarchical_Clustering (U_p);
02: {S1, S2,…} ← evaluate the centroid of each cluster in CU_p ;
03: output {S1, S2,…};

Procedure Hierarchical_Clustering (U_p)
04: Let each usage-point sequence in U_p be a cluster;
05: for each cluster C_j ∈ U_p do
06:     for each cluster C_k ∈ U_p − C_j do
07:         if (distance (C_j, C_k) ≤ σ) and (distance (C_j, C_k) is minimum in U_p) then
08:             merge C_j and C_k to a new cluster C_r;
09:             update distance between C_r and other clusters;
10:             U_p ← U_p − {C_j, C_k} ; U_p ← U_p ∪ C_r ;
11: output U_p ;
```

Fig. 4: The pseudo code of DBUP mining algorithm

For each two clusters, if their distance is smaller than or equal to the threshold σ and is minimum, DBUP Miner merges two clusters to a new cluster and updates the similarity between new cluster and other clusters (Lines 5-9, algorithm 2). The setting of threshold σ is usually heuristic; we will discuss in detail in experimental results and show the effect upon the mining results. Finally, after clustering, we evaluate the centroid of the each cluster and output these representative daily usage behaviors (lines 2-3, algorithm 2). We compute mean of each cluster as the representative centroid. With DBUP Miner, we can obtain the daily behavior-based usage patterns efficiently.

III. System Architecture

The architecture of proposed system, HAUBA (Household Appliance-Usage Behavior Analysis), is shown in Fig. 5. We attach smart meters to all appliances in the smart home environment and setup a cloud server to collect usage data. Smart meters will send the log data of appliance to server every constant time (about 1 to 5 seconds). When a user wants to know the information of an appliance, he/she can use the smart phone connect to cloud server and check two proposed usage patterns. Furthermore, if the electricity bill is high this month, users can observe the abnormal usage of appliances comparing with the discovered pattern from the proposed system. User can modify the range of parameter setting to control the tolerance of abnormality (i.e., the number of generated abnormal patterns). The user interface on a smart phone of proposed system is as shown in Fig. 1.

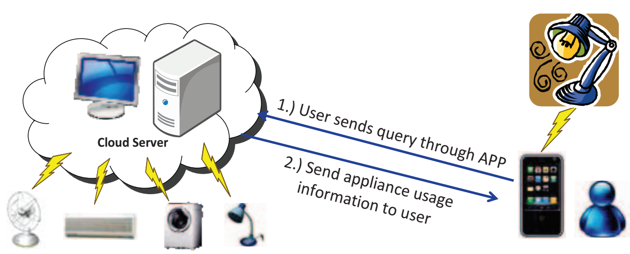

Fig. 5: System Architecture

IV. Experimental Results

To indicate the applicability of our system, we have performed an experiment on real-world dataset. In our smart home environment, we attach the wireless power meter on the six appliances, i.e., microwave, dish-washer, washer-dryer, light, oven and air-conditioner. The meter sends log data for every 3 or 4 seconds. The cloud database used in the experiment consists a collection of 24,692,250 data points for three years. Fig. 6 shows the discovered time-slot probability usage pattern of six appliances. We can observe that different appliance has different usage characteristic, different high usage frequency at different time slot. From these patterns, user can find the abnormal usage of appliances easily from the alarm on the smart phone by our system.

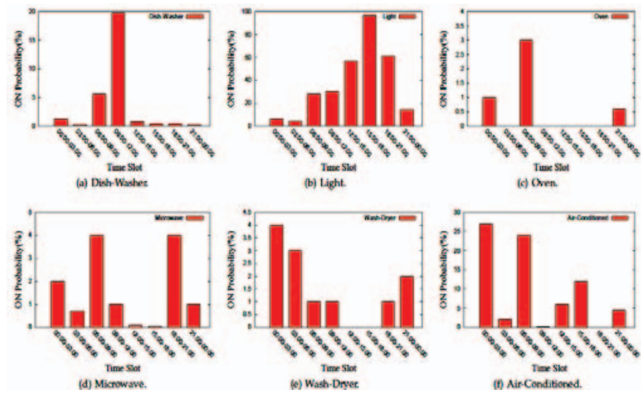

Fig. 6: Usage patterns of six appliances in smart home environment

For the DBUP, we implement EDR to be the similarity function. The advantage and property of EDR have been already introduced in Section II. Fig. 7 shows the discovered DBUP. The curves of different colors represent different daily behaviors. For example, Fig. 7(a) shows the DBUP of the microwave and its eight different representative daily usage behaviors.

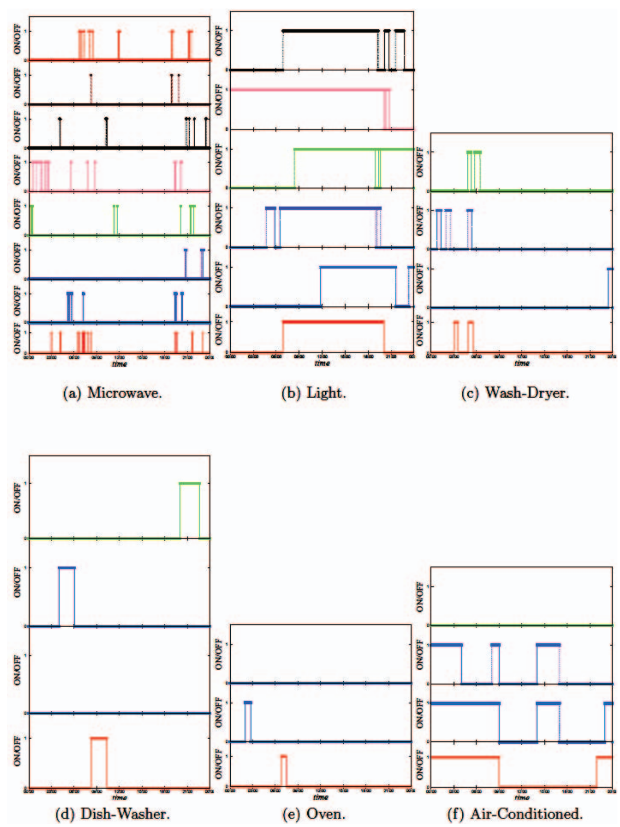

Fig. 7: Usage patterns of six appliances in smart home environment

V. Conclusion

Recently, some researches of usage pattern focused on power disaggregation and appliance recognition; however, little attention has been paid to the target of energy conservation. In this paper, we develop an intelligent system, HAUBA, which utilizes the analysis of appliance usage to detect the usage behavior. With the proposed usage pattern, users can acquire the information of power consumption and the representative usage behaviors of each appliance. In advance, if the electricity cost is high, users can observe the abnormal usage of appliances from the proposed system for electricity conservation easily. Furthermore, we also apply our system on real world dataset to show the practicability of mining usage pattern in smart home environment.

Acknowledgement

Wen-Chih Peng was supported in part by the National Science Council, Project No. 100-2218-E-009-016-MY3 and 100-2218-E-009-013-MY3, by Taiwan MoE ATU Program, by ITRI JRC, Project No. B352BW3300, by D-Link and by Microsoft.

References

[1] Aritoni, O., Negru, V.: A Methodology for Household Appliances Behavior Recognition in AmI Systems Integration. In: 7th International Conference on Automatic and Autonomous Systems, pp. 175–178 (ICAS'11). (2011)

[2] Chen, F., Dai, J., Wang, B., Sahu, S., Naphade, M., Lu, C.T.: Activity Analysis Based on Low Sample Rate Smart Meters. In: 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 240–248 (KDD'11). (2011)

[3] Farinaccio, L., Zmeureanu, R.: Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. Energy and Buildings, vol. 30, no. 3, pp. 245–259. (1999)

[4] Goncalves, H., Ocneanu, A., Bergés, M.: Unsupervised disaggregation of appliances using aggregated consumption data. KDD workshop on Data Mining Applications in Sustainability (SustKDD'11). (2011)

[5] Ito, M., Uda, R., Ichimura, S., Tago, K., Hoshi, T., and Matsushita., Y.: A method of appliance detection based on features of power waveform. In: 4th IEEE Symposium on Applications and the Internet, pp. 291–294 (SAINT'04). (2004)

[6] Kato, T., Cho, H., Lee, D., Toyomura, T., Yamazaki, T.: 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)

[7] Kolter, J.Z., Johnson, M. J.: REDD: A public data set for energy disaggregation research. KDD workshop on Data Mining Applications in Sustainability (SustKDD'11). (2011)

[8] Kim, H., Marwah, M., Arlitt, M., Lyon, G., Han, J.: Unsupervised disaggregation of low frequency power measurements. In: 11th SIAM International Conference on Data Mining, pp. 747–758 (SDM'11). (2011)

[9] Lin, G., Lee, S., Hsu, J., Jih, W.: Applying power meters for appliance recognition on the electric panel. In: 5th IEEE Conference on Industrial Electronics and Applications, pp. 2254–2259 (ISIEA'10). (2010)

[10] Matthews, H., Soibelman, L., Berges, M., Goldman, E.: Automatically disaggregating the total electrical load in residential buildings: a profile of the required solution. Intelligent Computing in Engineering, pp. 381–389. (2008)

[11] Prudenzi., A.: A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel. IEEE Power Engineering Society Winter Meeting, vol. 2, pp.491–496. (2002)

[12] Suzuki, K., Inagaki, S., Suzuki, T., Nakamura, H. and Ito., K.: Nonintrusive appliance load monitoring based on integer programming. In International Conference on Instrumentation, Control and Information Technology, pp, 2742–2747. (2008)