Available online at www.sciencedirect.com
ScienceDirect
Energy Procedia 105 (2017) 2499 – 2505
Energy Procedia

The 8th International Conference on Applied Energy – ICAE2016

A Methodology to Identify Multiple Equipment Coordinated Control with Power Metering System

Fan Feng[a], Zhengwei Li[a*]
[a]: *College of Mechanical Engineering, Tongji University, Shanghai, China*

Abstract

Modern buildings are equipped with advanced Building Energy Monitoring System (BEMS) for energy supervision and management. To fully leverage the power of this system, a dynamic real time data analysis to understand the building operational strategies is critical. Considering the significant volume of building energy data, a computer aided technique is necessary for any this kind of analysis. However, until now, there still lacks an effective method to automatically identify building operational strategies. In this paper, a new methodology targeting at identify multiple equipment coordinated control is proposed, which adopts Symbolic Aggregation AproXimation (SAX) as data pre-processor and performs a weighted association rule mining algorithm (WARM) to identify the coordinated control strategy. Case studies show that the proposed framework can effectively identify schedule based operational strategies and detect abnormal energy use behavior.

Peer-review under responsibility of the scientific committee of the 8th International Conference on Applied Energy.
keywords: Building energy monitoring system, coordinated control, association rule mining, automatic identification

1. Introduction

Building sector consumes more than 30% of the total energy worldwide (IEA, 2014). An efficient way to alleviate global warming and improve environmental sustainability is to enhance building energy efficiency.

Power metering system is an effective way of monitoring and supervising the energy consumption in buildings. Today, two types of power metering systems have been actively developed: Nonintrusive Load Monitoring (NILM) system, and detailed power submetering system. While the former is mainly used to monitor residential houses due to its non-intrusive nature; the latter simultaneously monitors a large number of electric equipment, and is mostly used in public buildings.

A typical power submetering system measures the electricity consumption of four major building service systems: lighting and plug loads; Heating, Ventilation, and Air Conditioning (HVAC) systems; power system (fan, pump, etc.); and special equipment (such as kitchen, information center, etc.). Due to the many advantages of power submetering system, it has been adopted as a recommended energy

Peer-review under responsibility of the scientific committee of the 8th International Conference on Applied Energy.
doi:10.1016/j.egypro.2017.03.721

monitoring system in many countries. Take China as an example, it is estimated that by 2015, the area of public buildings with power metering system has reached 60 million square meters (MOC 2011). With the popularity of power metering system, to quickly analyze the measured power data and send helpful feedbacks to facility managers becomes important. Due to the signification volume of the collected data, manual analysis is no longer a feasible approach, and computer aided techniques have to be resorted.

Data mining (DM) is a powerful approach to deal with 'big data', and has been successfully applied to understand human preferences. In the area of building energy research, DM methods have been heavily used for energy demand prediction (Li and Huang 2013; Magoules et al. 2013). In contrast, studies using DM to analyze building operational characteristics are relatively few.

Yu firstly adopt Association Rule Mining (ARM) to examine the relationship between building operational data, around 500 if-then rules were obtained, which were then filtered manually to find the problematic behavior (Yu et al. 2012). Obviously, manually filtering a set of candidate rules is painful given a large number of rule sets. To reduce the number of rule candidates, Fan proposed to conduct an energy use profile based clustering as a preliminary step, so that ARM is performed only for a representative data set (Fan et al. 2015). Another way to simplify the rule filtering process is to set a specific analysis target. Cabreba implemented a pattern recognition algorithm to identify abnormal light use behavior in classrooms, and found a high correlation between high energy consumption and light on during unoccupied period (Cabrera and Zareipour, 2013). It can be seen from above that, although DM methods are powerful in identifying hidden relationships, if the inputs for the DM process are not well controlled, the rules generated by ARM method are extremely difficult to manage.

On the other hand, regarding the operation strategies in buildings, there are typically two types: simple rule based strategy and more advanced model based strategy. Although model based strategy may lead to higher performance, its use is still limited due to reliability and robustness issues (Sun et al. 2013). For the rule based strategies, DM based methods (such as ARM) are perfect match to identify them. For this reason, it is argued in this study that DM method are better suited to identify normal operational strategies, which will be proved in the rest of this paper.

The content of this paper is organized as following: first, the principle of multiple equipment coordinated control is presented; second, this is followed by a detailed introduction of the proposed identification methodology; third, cases studies are conducted to validate the effectiveness of the proposed method; finally, conclusion remarks are given.

2. Multiple equipment coordinated control

Strategies to control multiple equipment of different types are referred to as coordinated control in this study. This type of strategy is typically set to guarantee the safety of equipment. A typical example is the coordination of chillers and pumps, since a water flow mismatch could lead to chillers' malfunction.

This control strategy can be executed either based on rules, or based on more advanced models with the aim to optimize certain variables. For rule based strategies that can be expressed in a way similar to 'if parameter A exceeds a, then equipment B does b', the DM based methods are perfect match for identification purpose, which will be shown below.

3. Proposed identification framework

The proposed framework is illustrated in Fig. 1, which consists of three main steps: preprocessing (to detect and diagnose the data quality of the power metering data), strategy identification and postprocessing.

3.1. Data preprocessing

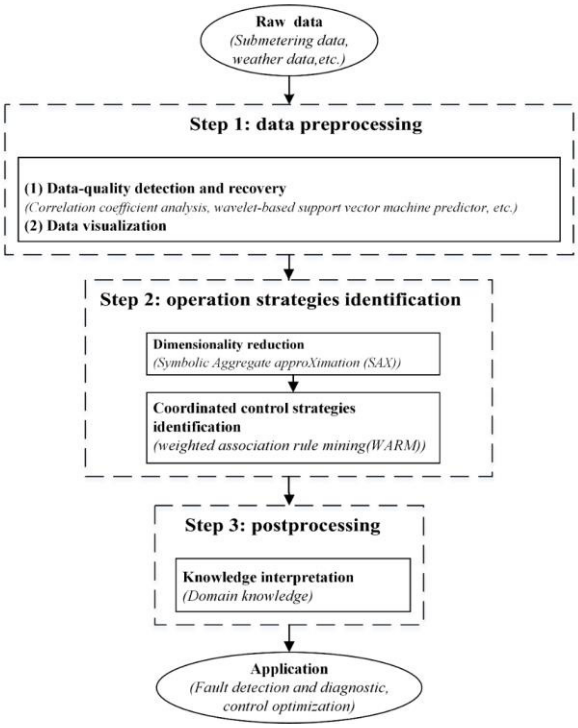

Figure 1. Framework for rule based strategy identification

In the preprocessing step, four types of faults are detected and diagnosed: (1) NAN (or zero fault), caused by hardware faults (including faults meter fault, data collection system failure, or transmission system failure); (2) negative fault, where negative data is recorded, caused by an inverse current flow due to improper wiring; (3) proportional deviation fault, typically due to the losing phase in either current or voltage; (4) disturbance fault, mostly due to environmental disturbance around working meters.

Among these four types of faults, type 1 and type 2 can be easily detected and corrected, thus are not discussed in this study. To detect and diagnose the proportional deviation fault, a combination of mismatch analysis and correlation analysis is proposed. While the mismatch between the total power usage and the sum of all metered power usage is used to detect the proportional deviation fault, the mismatch value is further correlated with all meters to identify the problematic meter. To detect and diagnose the type 2 fault, a wavelet based support vector machine (SVM) power predictor is applied. For more details regarding the data Fault Detection and Diagnostics (FDD) techniques, please refer to this paper by the authors (Fu et al., accepted).

3.2. Strategy Identification

As introduced above, this study aims at multiple equipment coordinated control. This framework utilized the Weighted Association Rule Mining(WARM) method, a variant of conventional association rule mining method, to fulfill this function.

Association rule mining (ARM) is a DM technique applied mainly to identify hidden relationships in the form of rules (Yu et al. 2012). Two parameters are critically important to identify the potential rules: support and confidence. While support is the joint probability of the antecedent and consequent, and confidence is the conditional probability of consequent, given the antecedent. Typically, rules associated with high support and confidence values are considered as important rules, and should be giver first consideration when checking the rule validity. However, it is found that when the volume of input data exceeds a certain amount, the speed of the commonly used Apriori algorithm decreases significantly. To speed up the process, two preprocessing techniques are applied: (1) hourly input data is aggregated to multiple hour data series with Symbolic Aggregate approXimation (SAX) technique (Clayton et al. 2015), which will be introduced below; (2) daily operational data is clustered into several groups based on their profiles. With these two techniques, the ARM algorithm is applied to the data set consists of only typical daily profiles, whose size is much smaller than the original data set.

The SAX technique is composed of several steps. First, the original hourly data set is normalized with Z-score normalization (as formulated in Eqn.1). Second, the hourly normalized data is broken down into n individual non-overlapping subsequences D, whose length is typically 24 (a day) in the context of building energy research. Each subsequence is further divided into W equally sized segments, and an alphabetic character is assigned to each of these segments according to their data averages. For example, character “a” means the lowest level, character “b” means a higher level than “a”, and so on. Furthermore, all of the W*D data segments are evenly divided into n bins, where n equals to the number of the characters, and the boundary of these bins are chosen to be the thresholds for transforming the continuous power usage data to discrete symbols. To illustrate the SAX word creation process, data of a variable during a day is transformed into a SAX words in Fig. 2, based on its z-score normalizations. If there are multiple variables, their SAX words can form a 'sentence', which describes the profiles of multiple equipment in the building (as in Fig. 3). It should be noted that in Fig. 3, the bar height denotes the frequency of the SAX word. In this way, the operation data of multiple equipment is transformed into several 'sentence'.

However, after performing the SAX pre-processing, the original dataset will be converted into weighted dataset, in which each typical day is assigned with a frequency representing its occurrence in this year. Correspondingly, a modified ARM is required to handle this weighted dataset. Here, we adopt the Weight Association Rule Mining method(WARM) for this purpose (Tao et al. 2003).

$$Z(t) = \frac{x(t) - \mu}{\sigma} \tag{1}$$

3.3. Knowledge Interpretation

After rules are identified, domain knowledge is still needed to interpret the results obtained in the knowledge discovery process. It should be noted that, the interpretation work required here is significantly less than the conventional approaches where all operational data is analyzed, due to the target specific analysis as introduced in Section 3.2.

4. Case Study

From the building studied in the second case study, data sets consisting of four chillers, six chilled water pumps, and six cooling water pumps are extracted and analyzed in this section. From the summary of the operational data of these equipment, it can be found that three chilled water pumps (CPSB11, CPSB 13, and CPB11) and two cooling water pumps (CNPSB12 and CNPSB14) don’t use any energy at all. Therefore, these equipment are excluded during the analysis.

To apply the WARM algorithm, the first step is to discrete the original data. With the number of discretization region set to 2 and the number of daily segments set to 6, the original data is transformed into a set of SAX sentences .

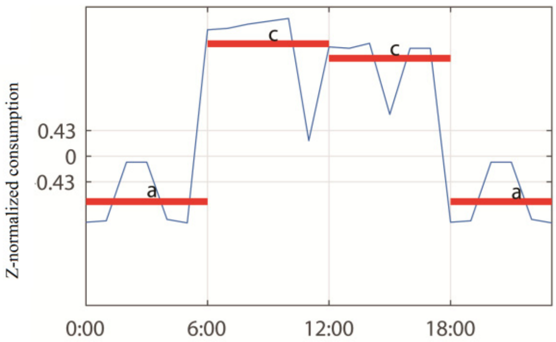

Figure 2. A SAX word creation example

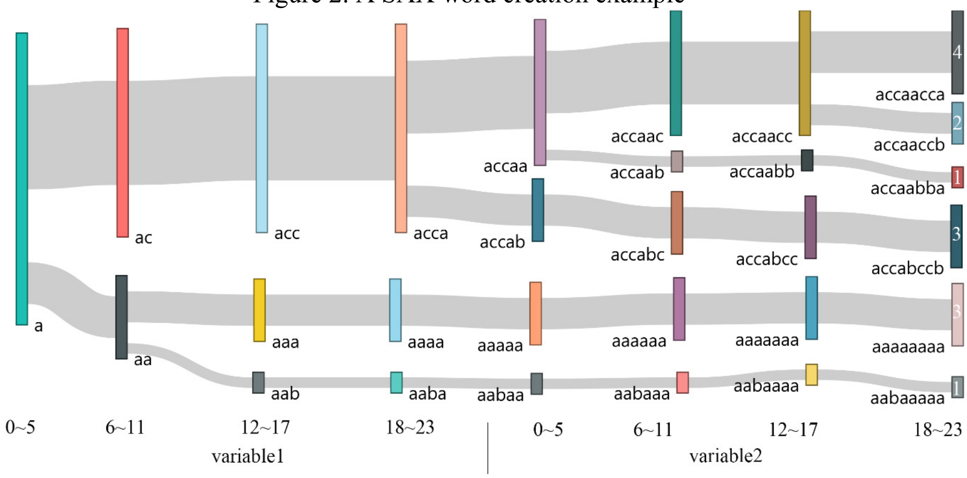

Figure 3. SAX words of two variables with two day data in a Sankey diagram

When the SAX transformation is done, WARM approach is applied to the derived data set. During the analysis, the support and confidence are set to 15% and 90% respectively, to keep infrequent but strong rules. As a result, 9 pairs of potentially useful rules are found and they are given in Table 1. Clearly, the antecedent and consequent of the two rules in each pair are exactly reversed, thus shows that the antecedent and consequent nearly occurred at the same time in this year. The first four rules show the association between two equipment: 2#Chiller and ChilledPumpB1-2. The first pair of rules indicates when the electricity usage of 2#Chiller is low, that of ChilledPumpB1-2 is always low as well, and the support of these rules is 0.544. The second pair of rules show that “2#Chiller = high” and “ChilledPumpB1-2 = high” occur at the same time. The total support of these two pairs of rules is 0.988, suggesting 2#Chiller and ChilledPumpB1-2 are turned on/off simultaneously. Similar phenomenon can be observed for rules 5~16. In Sum, ChilledPumpB1-2, ChilledPumpB1-3 and ChilledPumpB1-4 are dedicated to 2#Chiller, 3#Chiller and 4#Chiller, respectively

Table 1. Rules generated by WARM method

No.	antecedent	consequent	Sup	Conf
1	2#Chiller =low	ChilledPumpB1-2 = low	0.544	0.978
2	ChilledPumpB1-2 = low	2#Chiller =low		1
3	2#Chiller=high	ChilledPumpB1-2 = high	0.444	1
4	ChilledPumpB1-2 = high	2#Chiller=high		0.973
5	3#Chiller = low	ChilledPumpB1-3 = low	0.499	0.974

6	ChilledPumpB1-3 = low	3#Chiller = low		1
7	3#Chiller = high	ChilledPumpB1-3 = high	0.488	0.976
8	ChilledPumpB1-3 = high	3#Chiller=high		0.914
9	3#Chiller = low	CondPumpB1-4 =low	0.499	1
10	CondPumpB1-4 =low	3#Chiller = low		0.973
11	3#Chiller = high	CondPumpB1-4 =high	0.453	0.904
12	CondPumpB1-4 =high	3#Chiller = high		0.973
13	4#Chiller = low	ChilledPumpB1-4 = low	0.685	1
14	ChilledPumpB1-4 = low	4#Chiller = low		1
15	4#Chiller = high	ChilledPumpB1-4 = high	0.315	1
16	ChilledPumpB1-4 = high	4#Chiller = high		1

5. Conclusion

In this paper, a new analysis methodology to identify multiple equipment coordinated control. This methodology consists of mainly two parts: Symbolic Aggregation ApproXimation as preprocessor, and Weighted Association Rule Mining (WARM) as data miner. Testing of this methodology show that, it can successfully identify schedule based strategies for multiple equipment coordinated control. Compared with conventional ARM method, the proposed approach significantly reduces the work needed to filter out useless rules. Further work is still needed to simplify the interpretation of the generated rules.

Acknowledgement

The authors would like to thank the funding support from Shanghai Pujiang Talent Program (Project No. 15PJ1408100) and Chinese National Science Fund for Young Scholars (Project No. 51508394).

Reference

[1] International Energy Agency (IEA), Key world energy statistics, 2014.
[2] Z. Li, G. Huang, Re-evaluation of building cooling load prediction models for use in humid subtropical area, Energy and Buildings 62 (2013) 442-449.
[3] F. Magoules, H. Zhao, D. Elizondo, Development of an RDP neural network for building energy consumption fault detection and diagnosis, Energy and Buildings 62 (2013) 133-138.
[4] Z. Yu, F. Haghighat, B. C. M. Fung, et. al., A novel methodology for knowledge discovery through mining associations between building operational data, Energy and Buildings 47 (2012) 430-440.
[5] C. Fan, F. Xiao, C. Yan, A framework for knowledge discovery in massive building automation data and its applications in building diagnostics, Automation in Construction 50 (2015) 81-90.
[6] D. F. M. Cabrera, H. Zareipour, Data association mining for identifying lighting energy waste patterns in educational institutes, Energy and Buildings 62 (2013) 210-216.
[7] Y. Sun, G. Huang, Z. Li, et al., Multiplexed optimization for complex air conditioning systems, Building and Environment 65 (2013) 99-108.
[8] Y. Fu, Z. Li, F. Feng, et. al., Data-quality detection and recovery for building energy management and control systems (EMCSs): Case study on submetering, Science and Technology for the Built Environment, accepted (online).
[9] M. Clayton, Z. Nagy, A. Schlueter, Automated daily pattern filtering of measured building performance data, Automation in Construction 49 (2015) 1-17.
[10] Tao, Feng, Fionn Murtagh, and Mohsen Farid. "Weighted association rule mining using weighted support and significance framework." Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003.

Biography

Dr. Zhengewi Li graduates from Georgia Institute of Technology in 2012, with a focus on HVAC system fault detection technology. He has working experience in Lawrence Berkeley National Lab (LBNL) and City University of Hong Kong. He now leads the building energy system management direction in Tongji University.