A new route for energy efficiency diagnosis and potential analysis of energy consumption from air-conditioning system

Rong-Jiang Ma Nan-Yang Yu
PhD candidate Professor
School of Mechanical Engineering, Southwest Jiaotong University
Chengdu, China

ABSTRACT

Energy consumption of an air-conditioning system reflects the dynamic characters of system in an actual performance. It is extremely important to reduce system energy consumption, while maintains a comfort indoor air condition, by automatic monitoring, and controlling. However, the data collected in a large data repositories become "data disasters", for many operators, it is not possible to detect equipment, design, or operation issues because of data overload. Thus, this paper presents a new route to diagnose the faults of an air-conditioning system and identify the potential energy savings opportunities based on data mining. A case study is implemented to demonstrate the application of the new route and validate its feasibility and effectiveness. The results show that the approach can effectively identify system defaults and reduce the time spent on troubleshooting. It is a powerful and effective tool for diagnosis and potential analysis of energy consumption.

INTRODUCTION

Energy Consumption of Air-Conditioning System

Humanity faces serious energy and environment problems at present. The rapidly growing world energy use has already raised concerns over supply difficulties, exhaustion of energy resources and heavy environmental impacts (Pérez-Lombard et al. 2008). For instance, by increasing greenhouse gas emissions, which are contributed to concentrations in the atmosphere, they having already reached concerning levels in terms of their potential to cause climate change (Arroyo 2006). The international energy agency (IEA) has gathered frightening data on energy consumption trends. During 1984–2004, primary energy has grown by 49% and CO_2 emissions by 43%, with an average annual increase of 2% and 1.8% respectively (Gupta and Chandiwala 2009). Air pollution, acid precipitation and stratospheric ozone depletion are other serious environmental concerns. The severity of climate change impacts is predicted to increase if significant action is not taken to reduce greenhouse gas emissions (Croci et al. 2011). An important action to address energy and environmental challenges lies in the intelligent and efficient use of energy, including reducing energy waste and using low-carbon fuels (Ma et al. 2013).

The global contribution from buildings towards energy consumption, both residential and commercial, has steadily increased reaching figures between 20% and 40% in developed countries (Arroyo 2006). In China, building energy consumption accounts for 19.74% of total energy consumption (Tsinghua University Building Energy Research Center 2013). Among building services, the air-conditioning systems accounts for 50% ~ 60% of energy use (Zhang et al. 2009), and they are the largest energy end use both in the residential and non-residential sector. In other words, the energy consumption of air-conditioning systems accounts for one tenth of the global energy use at least.

Developments of Air-Conditioning System and Major Issues

Proceedings of the 14th International Conference for Enhanced Building Operations, Beijing, China, September 14-17, 2014

In 1902, Willis Carrier, the "father of air conditioning", designed a humidity control to accompany a new air-cooling system. He pioneered modern air conditioning (Whitman et al. 2009). After that, to decades of competition with block ice and gas-powered absorption systems, to the rapid spread of space cooling in factories, theaters, offices, houses, automobiles, and entire shopping centers, to the contemporary emergence of cooling load problems and the discovery of the dangers of greenhouse gases and ozone-depleting CFC refrigerants (CFCs) (Kempton and Lutzenhiser 1992).

As a major environmental issue, scientists concluded that released CFCs were depleting the earth's protective ozone layer. Hence many countries, professional organizations and industrial associations made it against the laws or protocols to intentionally vent CFCs into the atmosphere and manufacture CFCs. With sorts of alternative refrigerants were developing and applying to air-conditioning systems, the situation was better with each passing day.

Nevertheless, the status of another major issue is serious enough. As mentioned previously, global warming stemming from the uncontrolled rate of greenhouse gas emissions from energy-consuming process is a major twofold issue, which is the combination of energy and environmental.

An overall objective of energy policy in buildings is to save energy without compromising comfort, health and productivity levels. In other words, the idea is to consume less energy while providing equal or improved building services, that is, being more energy efficient (Pérez-Lombard et al. 2009). Especially important has been the intensification of energy consumption in air-conditioning systems, which has now become almost essential in parallel to the spread in the demand for thermal comfort, considered a luxury not long ago.

Europe developed early building envelope regulations in the late 1970s to reduce heat transfer through envelope elements and to control vapour diffusion and air permeability. This was followed by regulations or best practice recommendations on design, calculation and maintenance of air-conditioning systems and other building thermal services. Significantly, air conditioning equipment was subject for the first time to minimum requirements of energy efficiency (Pérez-Lombard et al. 2009). Until recently, various theories, methods, processes, and practices concerning energy efficiency of air conditioning have continued to develop.

For instance, the principal electrical components of a typical all-air air-conditioning system include: (1) chiller electric driven, (2) condensing water pump, (3) chilled water pump, (4) supply air fan, (5) return air fan, and (6) cooling tower fan. Moreover, the six devices consume most of the energy (electricity) in the system, and soon became the center of a particular focus.

According to the standards, it must be based on the most disadvantageous situation (maximum cooling load) for air-conditioning system design. That is to say, an air-conditioning system must meet the maximum cooling requirements of air-conditioned spaces. Indeed, systems rarely operate with the extreme conditions (maximum cooling load), and most of the time they operate with the partial load. Research data show that the operating time of central air-conditioning systems at 70% load (and below) accounts for 97% of service life (Wang H. Z. et al. 2009; Yang et al. 2005). So it is very meaningful and valuable to study energy efficiency for air-conditioning systems.

STATUS OF ENERGY EFFICIENCY DIAGNOSIS FOR AIR-CONDITIONING SYSTEM

Energy efficiency diagnosis is able to save energy, reduce maintenance costs, extend equipment life, and improve air-conditioning system control and occupant comfort. In this section, we focus on another critical issue: the current energy efficiency diagnosis methods of air-conditioning system. Summarizing the methods and indicators, the paper also answer to why the introduction of the new route for air-conditioning system.

Energy Efficiency Evaluation and Index System

Most energy (or building) services companies use the energy performance index (EPI) (Pérez-Lombard et al. 2009) as a starting point in energy audits and assess saving opportunities by comparing with existing references. In their final reports or ads, we can find a multiplicity of terms and concepts such as energy performance, energy efficiency, energy rating, benchmarking, baseline, grade, and some acronyms like COP, EER, IPLV, etc. Some of them have nearly similar meaning and usage, which are hard to distinguish. This has frequently led to misleading interpretations by consumers, company staff, and even proficient technicians. Even so, these terms, concepts and indexes have a common purpose: to improve energy efficiency and reduce energy consumption.

Coefficient of performance (COP) (McQuistion et al. 2005) is widely used in the field of air conditioning, and almost is the most important indicator for energy efficiency of an air-conditioning system.

Almost all countries have standards about the limit values of COP, such as "For air-cooled chillers of all sizes, the minimum requirement is COP of 2.8."specified in the ANSI/ASHRAE/IESNA Standard 90.1, *Energy Standard for Buildings Except Low-Rise Residential Buildings* (2010) in the US and "A water-cooled chiller under 528kW has a required minimum COP of 4.0, and a chiller that is over 1163kW has a required COP of 4.2." specified in The People's Republic of China National Standard GB 19577-2004, *The Minimum Allowable Values of The Energy Efficiency and Energy Efficiency Grades for Water Chillers* (2004).

Obviously, COP does not cover various situations, like air-conditioning systems in different regions, orientations, building types, occupants' habits and so on. In order to improve the evaluation of energy efficiency for the systems, other indicators such as IPLV(Cao 2004), SEER (Gu et al. 2005), and DEER (Xiao and Fu 2007) are also used in evaluation of system efficiency. All these indicators provide the basis for energy efficiency diagnosis indeed.

Typical Energy Efficiency Diagnosis Method and Implementation

There is no single "right way" to undertake energy efficiency diagnosis. The practical process should shift, adjust, and adapt to each project's needs and context. Therefore many kinds of diagnosis methods are applied to actual projects, but from the overall method, they are similar to each other. Typically and ideally, an energy efficiency diagnosis method includes three phases. First, by looking at the as-built drawings and the energy consumption records of the air-conditioning system, researchers must get the basic information and current status of energy consumption of the system, as well as by learning about the existing problems or troubles and related complaints of the system from the operators. Then, it is necessary to test, analyze and calculate for each existing problem in more detail including the typical condition and various operating conditions of different seasons according to operation records. Finally, the solution and corresponding potential analysis are presented by the summary report. This typical process can be summarized as OTI method (Xue Z. F.2007a; Xue Z. F. 2007b), that is to say, "observation/question → test/calculation → identification/resolution," as shown in Figure 1.

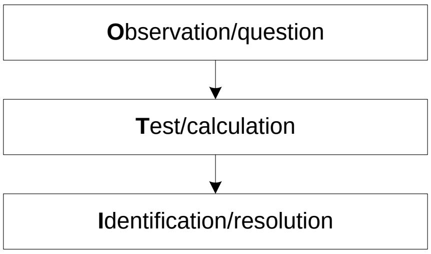

Figure 1 OTI method for energy efficiency diagnosis

This method involves building envelope, fresh air supply, cold and heat source, the transmission and distribution systems, etc., which can be divided into 15 steps for air-conditioning systems in detail, and 5 steps more for whole building. These steps connected to each other, but there is no fixed order. Normally, flexible work according to the specific situation of a system, different seasons, etc., will be carried out for an actual system.

But how to implement energy efficiency diagnosis by using this method? Let's take Step 6 of OTI method to illustrate it. The major contents of Step 6 are listed in Table 1.

Table 1 Diagnosis for COP of chiller (Step 6)

Phase	Job
O	—Whether the chilled water by-pass through a chiller, which is not running? —Is there a chiller correspond to two chilled water pumps? —Check the switch-status of by-pass valve between distributor and collector of chilled water.
T	—Test chilled water-flow of each chiller. —Test and calculate COP of chiller under typi-cal operating modes.
I	—To identify whether each chiller runs efficiently, by testing the typical operating modes and analyzing the operation records.

From the table, we can see that the second and third phase of the process are very important and difficult, especially the latter, and most of the time must be spent on it. At the same time, the indicators mentioned above as the basis of the final conclusion are exploited to evaluate the system. In addition, some methods like artificial neural network (ANN) (Zhao and Magoulès 2012) are used in the measurement and calculation for the value of parameters. That makes this process become more cumbersome and complex.

But generally speaking, although the OTI method is a self-contained system, it is limited to concrete analysis of a specific issue, without common applicable procedure. So to a large extent, the quality of the final result depends on the proficiency of the technicians and investigators in the work.

Summary

The above discussion reviews energy efficiency evaluation and index systems as the basis for diagnosis, presents the typical energy efficiency diagnosis method and implementation, The role of existing evaluation indicators and index systems is just at the level of guideline and standard, which is insufficient for tapping the potential capacity of energy conservation in an air-conditioning system. Present representative energy efficiency diagnosis methods are tedious and complex, and they require a heavy workload to acquire diagnosis results.

DATA MINING AND DATA QUALITY

This section describes the background and requirements of the new route presented in this paper.

Energy consumption of air-conditioning systems is one of the most direct and correct parameters, which reflects the dynamic characters of a system in actual running status. It is important to reduce system energy consumption by automatically monitoring, analyzing, and controlling. Nowadays, more and more air-conditioning systems have achieved energy consumption data acquisition, and more data are being accumulated in the operation process. These data will provide the foundation for analyzing the operation situation of the system. However, data collected in large data repositories become “data tombs”—data archives that are seldom visited (Han et al. 2012). Therefore, the traditional data analysis methods have some difficulties in dealing with the mass of data, resulting in more and more serious “data disasters” and making it difficult for operation staff to find the abnormal issues of a system efficiently. So it is more impossible to achieve the optimizing control.

Data Mining

Data mining is the application of specific algorithms for extracting patterns from data (Fayyad et al. 1996). Historically the notion of finding useful/interesting patterns in data has been given a variety of names including data mining, knowledge discovery from data (KDD), knowledge extraction, information discovery, knowledge mining from data, data/pattern analysis, data archaeology, data pattern processing, and data dredging. The knowledge discovery process as an iterative sequence comprises the following steps (Han et al. 2012): (1) data cleaning, (2) data integration, (3) data selection, (4) data transformation, (5) data mining, (6) pattern evaluation, and (7) knowledge presentation. In this sequence, steps (1) through (4) are different forms of data preprocessing, where data are prepared for mining.

Research (Kim et al. 2011; Ahmed et al. 2011; Wang Z. et al. 2011; Lian et al. 2006; Rakhshani et al. 2009; Liu et al. 2010; Li X. L. et al. 2010; Seem 2007) in the air-conditioning system or building energy consumption field by using the data mining approach has been yielding some results in recent years. By using this approach, this research has obtained results which are very difficult to be obtained by conventional methods. For example, energy modeling tasks are usually conducted later in the design process due to its time consuming data entry, but the case study in reference (Kim et al. 2011) revealed that data mining based energy modeling helps project teams discover useful patterns to improve the energy efficiency of building design during the design phase. A few researchers like (Ahmed et al. 2011) investigated the impact of connecting building characteristics and designs with their performance by data mining techniques. The derived results show the high accuracy and reliability of these techniques in predicting low-energy comfortable rooms. By using the approach and combining artificial intelligence algorithms, like ANN, a new modeling method for reducing the dimensions of acquired data, aiming to find a strong association, was developed in reference (Wang Z. et al. 2011), and some strategies (Lian et al. 2006, Rakhshani et al. 2009) were developed to detect and diagnose the faults of heating, ventilating, and air conditioning (HVAC) systems. At the same time, the detection/classification of abnormal energy consumption in buildings was researched by using the data mining approach. In Liu et al. (2010) and Li X. L. et al. (2010), the detection/classification method was proposed. Although it can even be used in conjunction with a building-management system to identify abnormal utility consumption and notify building operators in real time, we think, the outlier detection is only the first step that studies energy efficiency and energy conservation of the system. And in these, the causes of generating outlier and the optimization solutions of system operation require analysis and study by the operators and maintenance staff. The burden of this effort has been eased by the work of Seem (2007). The new method uses outlier detection to determine if the energy consumption for a particular day is significantly different than previous energy consumption. Operators should save time by not having to manually detect faults or diagnose false alarms; also, the new method will reduce operating costs by detecting problems that previously would have gone unnoticed (Seem 2007). But unfortunately, this method obtains the results (e.g., abnormal energy consumption) by comparing the actual energy consumption and normal energy consumption. Thus it does not completely satisfy the requirements of focus field of this paper, for instance, to ensure energy is being used most efficiently, to find out the potential capacity of energy conservation of the actual system.

Energy Consumption Data and Data Quality

Energy waste is often hidden, and some hidden problems that may not obviously affect comfort or environmental quality cause owners to pay much more for energy than necessary. In order to find problems and energy-saving potential and to optimize system operation more precisely and efficiently, it is necessary to put forward some requirements on the quality of the data. Hereafter, electricity-consumption data as the main research object will be discussed. Also, water, gas and other usage data of air-conditioning systems will be analyzed by using the method below.

(1) Data sources. Data can be derived from energy management systems (EMS), data loggers, or monitor sensors, etc.

(2) Accuracy. The data should reflect the real energy consumption of the monitoring components. Monitoring activities should not interfere with the air-conditioning system itself, for these monitoring tools/equipment are usually based on the Non-Intrusive Load Monitor (NILM) approach. More information about the NILM can be found in the references (Hart 1992; Norford and Leeb 1996; Fan et al. 2012).

(3) Completeness. A complete set of data should include the necessary parameters such as amperage, voltage, power factor or power of major energy-using devices for research or practical application. Data loss is not allowed. If any one of the parameters is missing, other parameters at the same time period shall not be used in the data.

(4) Consistency. The format of same type parameter in the system and the intervals of all parameters at the same monitoring time should be consistent.

NEW ROUTE

Overview of New Route

The flow chart is also the achievable functions of the new route as shown in Figure 2. And the whole process is aimed at detecting system problems and finding the energy-saving opportunities.

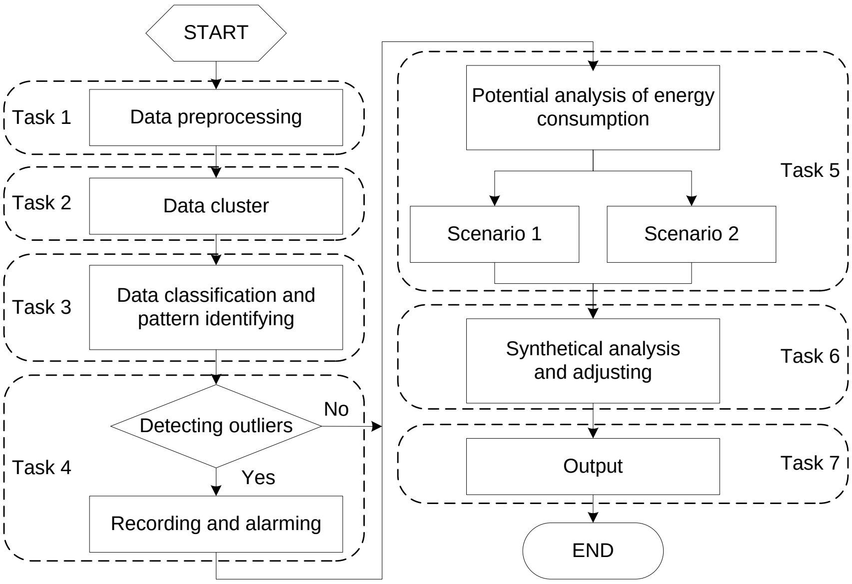

Figure 2 Flow chart of new route

The main tasks are as follows:

(1) Preprocess the data of energy consumption,

(2) Cluster the data and draw clustering figures,

(3) Classify the data of devices and system, and identify pattern of energy consumption (Seem 2005),

(4) Determine whether the energy consumption is abnormal, if "Yes", then take a record and alarm,

(5) In two scenarios, analyze the energy-saving potential of the system,

(6) Synthetically analyze and adjust the results in these two scenarios,

(7) Output the final result.

Data Preprocessing

Data preprocessing techniques can improve data quality, thereby helping to improve the accuracy and efficiency of the subsequent mining process. Data preprocessing is an important step in the knowledge discovery process, because quality decisions must be based on quality data. (Fayyad et al. 1996)

There are four major steps involved in data preprocessing as mentioned above, including many different methods in each step. For actual data preprocessing, every step should meet the requirements of data quality in the previous section and actual quality of the data, then an appropriate method should be chosen.

In the preprocessing step, the data are transformed or consolidated so that the resulting mining process may be more efficient, and the patterns found may be easier to understand (Han et al. 2012). Data normalization is one of strategies for data transformation, and many methods were adopted, such as min-max normalization, z-score normalization, normalization by decimal scaling, and so on. Because the values of energy consumption data are usually large, and exist with frequent variances at different times, so z-score normalization (Cai and Chen 2010) is useful and suitable for data normalization. In z-score normalization, the values for an attribute, A, are normalized based on the mean and standard deviation of A. And A value, v, of A is normalized to v' by computing

$$v' = \frac{v - \overline{A}}{\sigma_A} \tag{1}$$

where $\overline{A}$ and σ_A are the mean and standard deviation of attribute A, respectively.

The missing value of time-series data of energy consumption, for convenience, will be discarded in the preprocessing step.

Cluster Analysis

Cluster analysis (Han et al. 2012; He et al. 2007) can divide into several categories: partitioning clustering, hierarchical clustering, density-based clustering, and constraint clustering. Partitioning clustering (e.g., k-means, k-medoids, etc.) and hierarchical clustering (e.g., AGNES, DIANA, BIRCH, Chameleon, etc.) are designed to find spherical-shaped clusters, which have difficulties in finding clusters of arbitrary shape. Density-based clustering (e.g., DBSCAN, OPTICS, DENCLUE, etc.) can discover clusters of arbitrary shape, and handle anomalous data effectively. Constraint clustering is commonly used to handle certain requirements in the specific application areas.

Due to the potentially large dataset of energy consumption, there may be noisy data, and we have no prior knowledge on its shape, so DBSCAN (Density-Based Spatial Clustering of Applications with Noise) can be suitable for it.

DBSCAN algorithm steps (Nasibov and Ulutagay 2009) are explained as follows:

(1) Specify ε and *MinPts*,

(2) Find an unclassified core-point p with parameters ε and *MinPts*, mark the point p to be classified, start a new empty cluster C_t and assign p to this cluster,

(3) Find all the unclassified points in the ε-neighborhood of p and call them a set of seeds,

(4) Get a point q in the set of seeds, mark q to be classified, assign q to the cluster C_t, and remove q from the set of seeds,

(5) Check if q is a core-point with parameters ε and *MinPts*, if so, add all the unclassified points in the ε-neighborhood of q to the set of seeds,

(6) Repeat steps (4) and (5) until the set of seeds is empty,

(7) Repeat steps (2)-(6) until no more core points can be found.

where ε is the radius parameter, and *MinPts* is the neighborhood density threshold.

From the results analyzed by DBSCAN algorithm, we can get the preliminary patterns of energy consumption of a system, for example, time-series patterns, “weekday business hours,” “weekday off hours,” “weekend business hours,” and “weekend off hours.” Although these results cannot be used directly for energy efficiency diagnosis and potential analysis of energy consumption, that is the necessary basis of the next step.

Data Classification and Pattern Identifying

There are some major classification methods (Kotsiantis 2013): decision tree classifiers, Bayesian classifiers, and support vector machine (SVM), etc. Bayesian classifiers are statistical classifiers, which assume that the effect of an attribute value on a given class is independent of the values of the other attributes. However, this assumption tends to be invalid in many real systems, including air-conditioning systems. The SVM is a highly accurate classification method. Nevertheless, even the fastest SVM still suffers from slow processing, especially when training with massive datasets.

The construction of decision tree classifiers does not require any domain knowledge or parameter setting. The representation of acquired knowledge in tree form is intuitive and generally easy for people to assimilate. In addition, the decision tree classifiers have good accuracy, therefore, the study choose this method to construct the decision tree for patterns of energy consumption.

As a classic algorithm, C4.5 (Polat and Güneş 2009) is widely used to build classification models. C4.5 algorithm steps (Taherkhani 2010) are shown as follows:

(1) Compute gain ratio for each attribute by

$$GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)} \quad (2)$$

where *Gain* is information gain; *SplitInfo* is split information, and $SplitInfo(A) = -\sum_{i=1}^{n} p_i \log_2(p_i)$.

(2) Select the attribute with the maximum gain ratio as the splitting attribute of given set and create a node for the splitting attribute and mark the attribute, and then create a branch for each value in the attribute, and partition sample accordingly.

Through steps above, the patterns of energy consumption will be classified and identified for a given dataset. It should be noted that the patterns are more colorful and richer than those in previous subsection. Taking an example for time-series, the patterns were classified by hours, or even minutes, which are selected and adjusted in conformity to the character of practical data, whereas in previous subsection the patterns are classified simply by days.

Outliers Analysis and Energy Consumption

Anomalies Detection

Outlier detection methods (Xue A. R. et al. 2008) can divide into several categories, based on distribution, distance, density, etc. Distribution-based outlier detection methods (e.g., ESD, GESR, etc.) make assumptions of data normality. They assume that normal data objects are generated by a statistical (stochastic) model, and that data not following the model are outliers (Han et al. 2012). Distance-based outlier detection (e.g., k-NN, etc.) takes a global view of the dataset, and such outliers can be regarded as “global outliers.” Density-based outlier detection (e.g., LOF, etc.) is useful for global outliers and local outliers.

Given that there may be multiple clusters with different densities, and varieties of position relations between clusters in the dataset of energy consumption, in order to improve the detection accuracy, density-based outlier detection is a good choice. As a classic algorithm, LOF (Breunig et al. 2000) has some nice properties and wide applicability in a real world application.

LOF algorithm steps (Li J.et al. 2008) are expressed as follows:

(1) Search neighborhoods for each object in the dataset, calculate the *MinPts*-neighborhood, and store the distances between the object and its neighbors,

(2) Calculate the local reachability density and local outlier factor (LOF) for each object,

(3) Check every object to determine if it is an outlier according to the pre-set threshold for LOF.

Outliers are not only great energy-saving opportunities, but also the reflection of troubles and faults of a system. So this method will take a record and give an alarm to operators when it finds any outliers.

Potential Analysis and Results

Two scenarios for analysis

This paper analyses the potential capacity of energy conservation under the following two scenarios:

Scenario 1: Study the best running state of system under various operating conditions, without considering the reasonableness of design and device selection;

Scenario 2: Study the best running state of system under various operating conditions, with energy-saving revamping measures.

As one can see, Scenario 1 is the most basic way, which hardly requires money and tunes the system to get it running as well as possible. In addition, the potentials are also very limited, and they are largely affected by the existing system itself. In Scenario 2, it will take the necessary reform measures to improve problem areas, for instance, install the frequency convertor on pump or fan motors, replace the air-cooled chiller with oil-free chiller or water-cooled chiller, etc. which are more energy-efficient than existing facilities. In a word, it costs something in Scenario 2, but saves much more than Scenario 1.

Theoretical basis

Take pumps as an example, η, efficiency parameter, is often used to evaluate the economy of energy consumption. But in fact, it cannot directly reflect the economy level of pump or transport system by itself. Just like the efficiency of auto engines or transmission systems, the values of η cannot be measured directly, and they are not the users' concern. What is more interesting to us is that which can directly and accurately reflect the economic performance of devices or systems, such as fuel consumption per 100 km for cars, and fuel consumption per ton-kilometer for trucks. Some research (Lu 2007; Dong et al. 2008; Lu and Wang 2009) has used engineering attributes to deal with the energy analysis and provided some new concept like K_E, the "energy transport consumption per unit volume (i.e. energy consumption coefficient)" (Dong et al. 2008). The definition (Lu and Wang 2009) is presented as:

$$K_E = \frac{P}{q_V} = \frac{(\rho g H q_V)/\eta}{q_V} = \frac{\rho g H}{\eta} \quad (3)$$

where P is pump shaft power, q_V is transport volume, ρ is density of liquid, g is the gravitational acceleration, H is pump head, η is pump efficiency. And the unit of K_E is J/m^3.

For real-world systems, power consumption of pump system (P_E) is easier to measure than P, the above can be adapted as (Lu and Wang 2009):

$$K_{E,E} = \frac{P_E}{q_V} = \frac{(\rho g H q_V)/(\eta \cdot \eta_M)}{q_V} = \frac{\rho g H}{\eta \cdot \eta_M} \quad (4)$$

where η_M is pump motor efficiency, and the unit of $K_{E,E}$ is kWh/m^3.

As an index for the economic estimation, this concept will be used in the analysis.

Analysis thread and results

Analysis thread for Scenario 1 as follows:

(1) Analyze the data of each pattern of energy consumption, and get the operating characteristic curve,

(2) Analyze the data of each pattern of energy consumption by time-series, and get the attenuation of system and each device,

(3) Study the time and space distribution of data of each pattern of energy consumption, and get the logical relationship between data of each device,

(4) Determine the optimal operating point or range for each pattern of energy consumption based on above works, and as measured by $K_{E,E}$,

(5) Calculate the difference between actual operating data and the optimal data of corresponding patterns, and get the energy-saving potential of devices and systems.

The analysis thread for Scenario 2 is similar to above, except that it will call the data of revamping measure database in step (4).

Before the final result comes out, we need to synthetically analyze and adjust the results of two scenarios accounting for variations in weather, space use, or other variables from year to year.

METHOD VALIDATIONS

It is known from analysis in the second section that we still lack a systematic method for energy efficiency diagnosis of air-conditioning systems, so we proposed the new route before that. This section selected one case to investigate the application of the new route mentioned at the previous section, in order to validate its feasibility and effectiveness.

Case Description

As described in the above sections, the electricity of pumps was a part of air-conditioning system energy consumption, which was shown in Figure 3. We can see that the pumps are intermittent, running in 0:00 to 06:00 and 22:00 to 0:00, the rest of the time running continuously. Note that this pump electricity was simulated by using "RefBldgLargeOfficeNew2004_Chicago.idf" on July 21 in the Energyplus V8.0.0 (http://apps1.eere.energy.gov). The "timestep" was set to 60 times per hour. This study selected

SPSS Clementine 12.0 (http://www.SPSS.com) to preprocess and classify the data. In addition, ELKI 0.6.0 (http://elki.dbs.ifi.lmu.de) was used to do cluster analysis and outlier analysis for all the data.

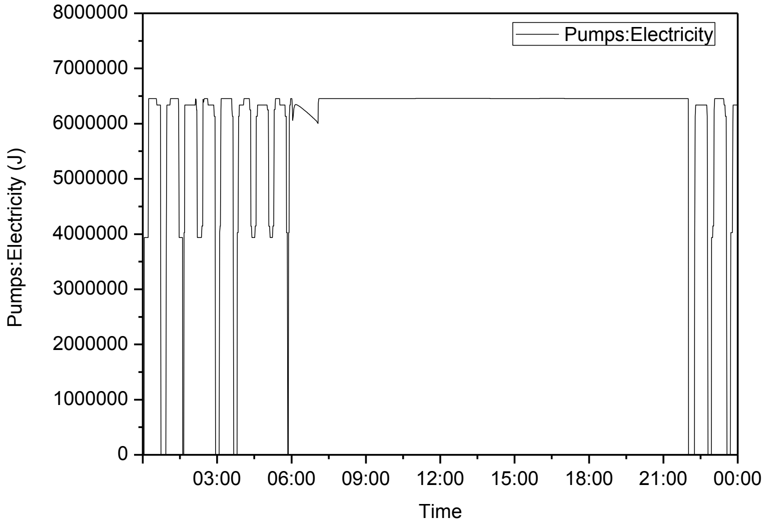

Figure 3 Electricity consumption of pumps in July 21st

Validation Process and Results

According to the method shown in the previous section, the energy consumption data were first preprocessed before the cluster analysis. To find the potential energy consumption pattern for an air-conditioning system, the cluster analysis needs to extract the feature vector from the time series data of energy consumption. This important feature vector shows the system energy consumption and was represented as:

$$\boldsymbol{C}=(C_{avg},\ C_{max}) \tag{5}$$

where C_{avg} and C_{max} are the averaged and maximum energy consumption per quarter of an hour, respectively.

Figure 4 shows the results of cluster analysis utilizing the method presented in the previous section. From this figure, we can find that there are three patterns for system energy consumption, which are system halt and underload condition and constant load condition. Here these two conditions were named Cluster A and Cluster B, respectively.

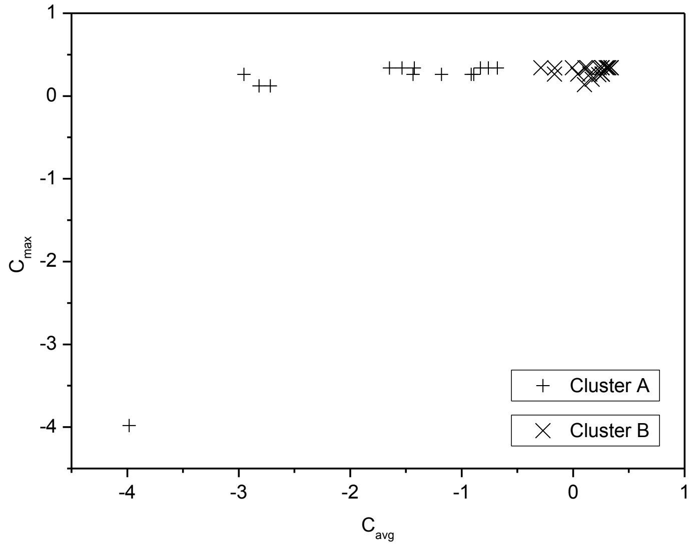

Figure 4 Cluster analysis results

To easily analyze the criterion of energy consumption pattern, this study constructed one new attribute: energy consumption label, and brought results of the cluster analysis into the case data. These case data were then classified by using the method mentioned in the previous section and applied to obtain the decision tree (Figure 5) of energy consumption pattern.

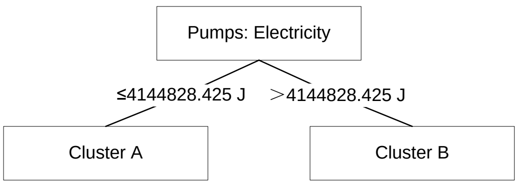

Figure 5 Decision tree of energy consumption pattern

The conclusion was presented as:

(1) If “Pumps: Electricity”≤4144828.425 J (based on 109 training data), this energy consumption pattern is Cluster A with confidence level of 100%.

(2) If “Pumps: Electricity”＞4144828.425 J (based on 891 training data), this energy consumption pattern is Cluster B with confidence level of 100%.

To validate the effect of outlier analysis, we assumed one abnormal point at one moment between 8:01 to 22:00, such as the decrease of motor efficiency. According to the decision tree of energy consumption pattern, the entire patterns are Cluster B at this moment. In addition, to estimate the abnormal moment as detailed as possible, the energy consumption data at abnormal moment and normal moment with the same pattern (Cluster B) were analyzed by using the method mentioned in the previous section. Figure 6 shows the local outlier factor for all the data between 8:01 to 22:00. From this figure, we can find that the majority of LOF values were around 1, while the LOF value at 16:33 was 35.25. In fact, the motor efficiency at this moment decreased from 90% to 89%, resulting in abnormal energy consumption. Therefore, this special point was exactly the outlier which was characterized by the abnormal energy consumption.

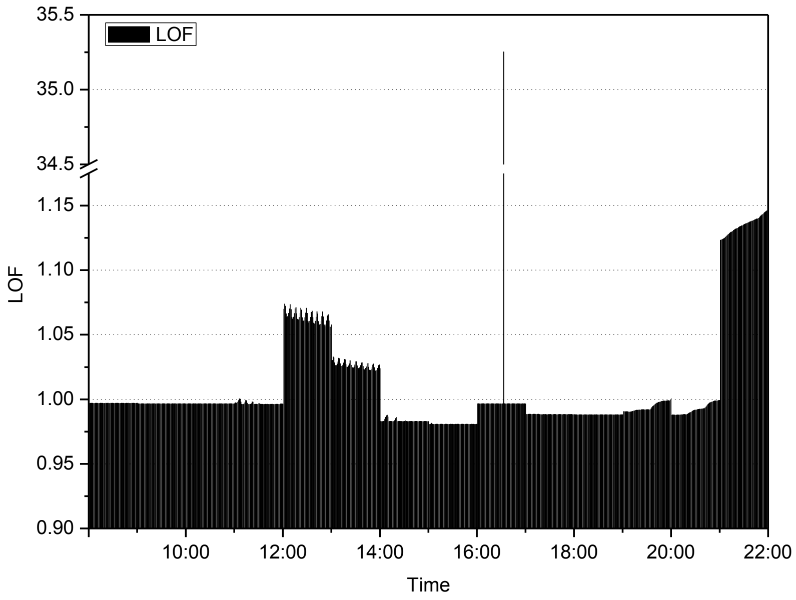

Figure 6 The LOF values of energy consumption data between 8:01 to 22:00

Discussion

For the potential analysis of energy consumption conservation, the more variable and more substantial data are required, such as data warehouse. It should be noted that the energy consumption difference between the motor efficiency with the value of 89% and 90% at 16:33 was considered as basic energy-saving potential. Due to lack of more data and collision with the main object of this research, this paper will not investigate the validation more deeply. Further study will focus on more detailed application research.

CONCLUSIONS

In this paper, we reviewed the major issues and status of energy efficiency diagnosis for air-conditioning systems. The role of existing evaluation indicators and index systems is just at the level of guideline and standard, and it is insufficient for tapping the potential capacity of energy conservation in air-conditioning systems. Although efforts have been made to develop energy efficiency diagnosis methods, including expert diagnosis systems and knowledge-based technologies, these typically rely on users or domain experts to manually input knowledge into knowledge bases. Unfortunately, the manual knowledge input procedure is prone to biases and errors and is extremely costly and time consuming.

Although many buildings use a sophisticated system (e.g., EMS) to manage a wide and varied range of building systems, which can collect and store massive quantities of energy consumption data, facility operators can be overwhelmed with the quantity of data. For many operators, it is not possible to detect equipment, design, or operation problems because of data overload. Data mining, which has benefited from the development of computer and information technology, provides a basis for using the data.

Based on these understandings, we presented a new route for solving the energy efficiency diagnosis and potential analysis of energy consumption using the data mining approach, and introduced the main tasks, implementation methods and some requirements. We then selected one case to investigate the application of the new route, and the results show that the route is feasible and applicable in air-conditioning systems.

At last, for this route, we recognize that there is still a long way to investigate, improve, and practically apply. Even so, we still believe that this approach advances a brand-new research method and is of great project application value in bringing about a leap of energy efficiency of air-conditioning systems in a real sense.

ACKNOWLEDGMENTS

Authors would like to thank Dr. Ji-Ying Liu for his advice and constructive discussions.

CONFLICT OF INTEREST

Authors declare that we do not have any commercial or associative interest that represents a conflict of interest regarding the publication of this article.

REFERENCES

Ahmed, A., Korres, N. E., Ploennigs, J., Elhadi, H., & Menzel, K. (2011). Mining building performance data for energy-efficient operation. *Advanced Engineering Informatics, 25*(2), 341-354, doi:10.1016/j.aei.2010.10.002.

ANSI/ASHRAE/IESNA Standard 90.1, Energy standard for buildings except low-rise residential buildings (2010). (Vol. 90.1). Atlanta, USA: ASHRAE, Inc.

Arroyo, V. (2006). Agenda for climate action. Arlington, VA, USA: Pew Center on Global Climate Change.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: identifying density-based local outliers. *SIGMOD Rec., 29*(2), 93-104, doi:10.1145/335191.335388.

Cai, W. L., & Chen, D. X. (2010). Influence of date normalization methods on k-nearest neighbor classifier. *Computer engineering*(22), 175-177.

Cao, Q. (2004). Disccussion on meanings of IPLV. *Construction Machinery for Hydraulic Engineering & Power Station*(2), 9-10.

Croci, E., Melandri, S., & Molteni, T. (2011). Determinants of cities' GHG emissions: a comparison of seven global cities. *International Journal of Climate Change Strategies and Management, 3*(3), 275-301.

Dong, Y., Yan, G. J., & Lu, Z. D. (2008). Energy efficiency analysis and evaluation for the pump fluid system. *China mechanical engineering, 19*(14), 1687-1690.

Fan, Y. C., Liu, X. J., Lee, W. C., & Chen, A. L. P. Efficient Time Series Disaggregation for Non-intrusive Appliance Load Monitoring. In *Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on, 4-7 Sept. 2012 2012* (pp. 248-255). doi:10.1109/UIC-ATC.2012.122.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*: AAAI/MIT Press.

Gu, B., Wang, P., & Xi, D. M. (2005). *SEER standard calculation based on weather parameters for air-conditioner in China.* Paper presented at the Chinese Association of Refrigeration (CAR) Kunming, China.

Gupta, R., & Chandiwala, S. (2009). *A critical and comparative evaluation of approaches and policies to measure, benchmark, reduce and manage CO_2 emissions from energy use in the existing building stock of developed and rapidly-developing countries - case studies of UK, USA, and India.* Paper presented at the 5th Urban Research Symposium: Cities and Climate Change - Responding to an Urgent Agenda, Marseille, France, 28-30 June

Han, J., Kamber, M., & Pei, J. (2012). *Data mining: concepts and techniques*. Singapore: Elsevier Inc.

Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proceedings of the IEEE, 80*(12), 1870-1891, doi:10.1109/5.192069.

He, L., Wu, L. D., & Cai, Y. C. (2007). Survey of clustering algorithms in data mining. *Application research of computers, 24*(01), 10-13.

Kempton, W., & Lutzenhiser, L. (1992). Air-Conditioning - The Interplay of Technology, Culture and Comfort - Introduction. *Energy and Buildings, 18*(3-4), 171-176, doi:10.1016/0378-7788(92)90011-5.

Kim, H., Stumpf, A., & Kim, W. (2011). Analysis of an energy efficient building design through data mining approach. *Automation in Construction, 20*(1), 37-43, doi:10.1016/j.autcon.2010.07.006.

Kotsiantis, S. B. (2013). Decision trees: a recent overview. [Article]. *Artificial Intelligence Review, 39*(4), 261-283, doi:10.1007/s10462-011-9272-4.

Li, J., Yan, B. P., & Li, J. (2008). Memory-effect-based local outlier detection algorithm. *Computer engineering, 34*(12), 4-6.

Li, X. L., Bowers, C. P., & Schnier, T. (2010). Classification of Energy Consumption in Buildings With Outlier Detection. *Industrial Electronics, IEEE Transactions on, 57*(11), 3639-3644, doi:10.1109/TIE.2009.2027926.

Lian, Z. W., Hou, Z. J., Yao, Y., & Yuan, X. J. (2006). Data mining based sensor fault diagnosis and validation for building air conditioning system. *Energy Conversion and Management, 47*(15-16), 2479-2490, doi:10.1016/j.enconman.2005.11.010.

Liu, D., Chen, Q., Mori, K., & Kida, Y. (2010). A Method for Detecting Abnormal Electricity Energy Consumption in Buildings. *Journal of Computational Information Systems, 6*(14), 4887-4895.

Lu, Z. D. (2007). New energy efficiency analysis and evaluation for the design and implementation of pump and fan systems. *Energy conservation technology, 25*(02), 182-189.

Lu, Z. D., & Wang, L. W. (2009). *Energetics and economy analysis of pump and fan systems*. Beijing, China: National defense of industry press.

Ma, R. J., Yu, N. Y., & Hu, J. Y. (2013). Application of Particle Swarm Optimization Algorithm in the Heating System Planning Problem. [Article]. *Scientific World Journal*, doi:10.1155/2013/718345.

McQuistion, F. C., Parker, J. D., & Spitler, J. D. (2005). *Heating, Ventilating, and Air Conditioning Analysis and Design* (6ed.). NJ, USA: John Wiley & Sons, Inc.

Nasibov, E. N., & Ulutagay, G. (2009). Robustness of density-based clustering methods with various neighborhood relations. *Fuzzy Sets and Systems, 160*(24), 3601-3615, doi:10.1016/j.fss.2009.06.012.

Norford, L. K., & Leeb, S. B. (1996). Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. *Energy and Buildings, 24*(1), 51-64, doi:10.1016/0378-7788(95)00958-2.

The People's Republic of China National Standard GB 19577-2004, The minimum allowable values of the energy efficiency and energy efficiency grades for water chillers (2004). Beijing, China: China Standards Press.

Pérez-Lombard, L., Ortiz, J., González, R., & Maestre, I. R. (2009). A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes. *Energy and Buildings, 41*(3), 272-278, doi:10.1016/j.enbuild.2008.10.004.

Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings, 40*(3), 394-398, doi:10.1016/j.enbuild.2007.03.007.

Polat, K., & Güneş, S. (2009). A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multi-class classification problems. *Expert Systems with Applications, 36*(2, Part 1), 1587-1592, doi:10.1016/j.eswa.2007.11.051.

Rakhshani, E., Sariri, I., & Rouzbehi, K. Application of data mining on fault detection and prediction in Boiler of power plant using artificial neural network. In *Power Engineering, Energy and Electrical Drives, 2009. POWERENG '09. International Conference on, 18-20 March 2009 2009* (pp. 473-478). doi:10.1109/POWERENG.2009.4915186.

Seem, J. E. (2005). Pattern recognition algorithm for determining days of the week with similar energy consumption profiles. *Energy and Buildings, 37*(2), 127-139, doi:10.1016/j.enbuild.2004.04.004.

Seem, J. E. (2007). Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy and Buildings, 39*(1), 52-58, doi:10.1016/j.enbuild.2006.03.033.

Taherkhani, A. Recognizing Sorting Algorithms with the C4.5 Decision Tree Classifier. In *Program Comprehension (ICPC), 2010 IEEE 18th International Conference on, June 30 2010-July 2 2010 2010* (pp. 72-75). doi:10.1109/ICPC.2010.11.

Tsinghua University Building Energy Research Center. (2013). *2013 Annual report on China building energy efficiency*. Beijing, China: China Architecture & Building Press.

Wang, H. Z., Wang, H. M., & Lin, W. (2009). Energy-saving research on central air-conditioning system. *Energy Engineering*(03), 61-64.

Wang, Z., Han, N., & Wang, Y. (2011). *Studies on Neural Network Modeling for Air Conditioning System by Using Data Mining with Association Analysis.* Paper presented at the Proceedings of the 2011 International Conference on Internet Computing and Information Services (ICICIS 2011), 2011

Whitman, W. C., Johnson, W. M., Tomczyk, J. A., & Silberstein, E. (2009). *Refrigeration & air conditioning technology* (sixth ed.). Delmar: Cengage Learning.

Xiao, Y. M., & Fu, X. Z. (2007). *Limit value on design energy efficiency ratio of central air conditioning project in public buildings.* Paper presented at the Academic annual conference for HVAC and thermal power of southwestern China, Chengdu, China,

Xue, A. R., Yao, L., Ju, S. G., Chen, W. H., & Ma, H. D. (2008). Survey of outlier mining. *Computer science, 35*(11), 13-18+27.

Xue, Z. F. (2007a). *Building energy efficiency technology and application*. Beijing, China: China Architecture & Building Press.

Xue, Z. F. (2007b). *Energy efficiency diagnosis and refit for existing buildings*. Beijing, China: China Architecture & Building Press.

Yang, C. Z., Liu, G. D., & Li, N. P. (2005). *Engineering design method and system analysis of HVAC.* Beijing, China: China Architecture & Building Press.

Zhang, Y. B., Bai, X. L., Tian, Z. H., & Wu, L. J. (2009). Study on operating dynamic energy efficiency of HVAC systems. *Refrigeration & Air Conditioning, 23*(4), 16-19+23.

Zhao, H.-x., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews, 16*(6), 3586-3592, doi:10.1016/j.rser.2012.02.049.