Applied Energy 211 (2018) 1123–1135

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data

Cheng Fan[a], Fu Xiao[b], Yang Zhao[c,*], Jiayuan Wang[a]

[a] Department of Construction Management and Real Estate, Shenzhen University, Shenzhen, China
[b] Department of Building Services Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China
[c] Institute of Refrigeration and Cryogenics, Zhejiang University, Hangzhou, China

HIGHLIGHTS

- An autoencoder-based ensemble method is developed for anomaly detection.
- Autoencoders can capture the intrinsic characteristics in building energy data.
- The performance of various autoencoder types and training schemes is compared.
- Methods are developed for performance evaluation without using anomaly labels.
- This study provides data-driven solutions to unsupervised anomaly detection.

ARTICLE INFO

Keywords:
Autoencoder
Unsupervised data analytics
Anomaly detection
Building operational performance
Building energy management

ABSTRACT

Practical building operations usually deviate from the designed building operational performance due to the wide existence of operating faults and improper control strategies. Great energy saving potential can be realized if inefficient or faulty operations are detected and amended in time. The vast amounts of building operational data collected by the Building Automation System have made it feasible to develop data-driven approaches to anomaly detection. Compared with supervised analytics, unsupervised anomaly detection is more practical in analyzing real-world building operational data, as anomaly labels are typically not available. Autoencoder is a very powerful method for the unsupervised learning of high-level data representations. Recent development in deep learning has endowed autoencoders with even greater capability in analyzing complex, high-dimensional and large-scale data. This study investigates the potential of autoencoders in detecting anomalies in building energy data. An autoencoder-based ensemble method is proposed while providing a comprehensive comparison on different autoencoder types and training schemes. Considering the unique learning mechanism of autoencoders, specific methods have been designed to evaluate the autoencoder performance. The research results can be used as foundation for building professionals to develop advanced tools for anomaly detection and performance benchmarking.

1. Introduction

Building operational performance has become one of the top concerns in achieving global sustainability. On the one hand, building operations are energy-intensive and contribute to approximately one-third of the world final energy consumption [1]. On the other hand, building operations have substantial energy saving potential considering the wide existence of equipment faults, energy-wasting occupant behaviors and improper control strategies. It is estimated that 16% of the energy consumed during building operations can be conserved through currently available energy management techniques [2]. One of the most promising solutions to tackling energy wastes during building operations is anomaly detection. Anomaly detection refers to the process of identifying rare observations in a data set [3]. It has been widely used in various industries, such as intrusion detection in network systems, fraud detection in financial transactions, and patient health monitoring in medical treatment [4]. In the building field, anomaly detection focuses on the detection and diagnostics of abnormal building operational patterns, which can be results of atypical operating behaviors, errors in sensing and transmission systems, equipment faults and energy-inefficient operating strategies [5]. It has been applied to three different levels, i.e., whole building level, subsystem level and

* Corresponding author.
E-mail address: youngzhao@zju.edu.cn (Y. Zhao).

https://doi.org/10.1016/j.apenergy.2017.12.005
Received 15 September 2017; Received in revised form 13 November 2017; Accepted 1 December 2017

component level [6]. The timely discovery of anomalies in building operations can be very helpful for building operation staff to understand building operating conditions, perform energy performance assessment and develop actionable measures for energy conservation.

Thanks to the development in information technologies, the real-time building operational performance can be well monitored and controlled through the building energy management systems (BEMS) or building automation systems (BAS). Massive amounts of building operational data are being collected and available for data analysis. It is therefore very promising to develop data-driven approaches to achieving reliable and robust anomaly detection. Based on the types of data analytics used, existing anomaly detection methods in the building field can be classified into two groups, i.e., supervised and unsupervised methods. The former adopts a model-based approach to anomaly detection. Supervised learning algorithms, such as support vector machines and multi-layer perceptions, are adopted for developing predictive models [7]. The model output is typically defined in two ways. The first is to use a variable to describe operating conditions as model output. In such a case, anomaly detection is performed based on the difference between predicted and actual values. Du et al. adopted neural networks to predict the supplied air temperature of the HVAC system given the supplied and returned chilled water temperature, the chilled water flow rates and etc. [8]. The difference between predicted and actual supplied air temperature was used for anomaly detection and various anomalies due to sensor biases, water valve stuck and controller faults were successfully discovered. Magoules et al. set the aggregated electricity consumption of building facilities, interior equipment, chillers, fans and pumps as the output for neural network models [9]. The method was successfully used to perform building-level anomaly detection. The other way of defining model output is to directly use a label stating whether an observation is an anomaly or not. In such a case, anomaly detection is transformed into a binary or multi-class classification problem. Zhao et al. developed a support vector machine-based method to detect anomalies in chiller operations [10]. A number of chiller operating parameters were used as model inputs and the model was trained using normal data alone. It was reported that anomalies at various severity levels could be successfully identified. Guo et al. applied a back-propagation neural network to identify anomalies in a variable refrigerant flow air conditioning system [11]. The model output was set as labels stating whether observations were normal or faulty. Despite of the effectiveness of model-based methods, their practical values are usually limited. The reasons are: (1) Obtaining high-quality training data can be time-consuming; (2) Obtaining the label (or ground truth) on whether an observation is abnormal or not can be costly and sometimes infeasible.

By contrast, unsupervised anomaly detection methods are more promising for practical applications, as they do not require anomaly labels. Existing methods can be further classified into two types, i.e., statistical methods and unsupervised data mining methods. Statistical methods make statistical assumptions on the underlying data distribution (e.g., Gaussian normal distribution), based on which scores are calculated for anomaly detection. One prominent example is the generalized extreme studentized deviate (GESD)-based method. Previous studies have demonstrated the usefulness of GESD-based methods in identifying anomalies in building energy consumption profiles [12,13]. The method typically involves a feature extraction step to represent high-dimensional data as low-dimensional features. The GESD algorithm is then applied on features for anomaly ranking [14]. Other statistical methods include the nearest neighbors and principal component analysis-based methods [15,16]. Recently, unsupervised data mining techniques have gained increasing interests in anomaly detection due to their excellence in handling massive and complicated data sets [17,18]. The most widely used unsupervised data mining techniques in the building field include clustering analysis and association rule mining (ARM) [18]. The original intention of clustering analysis is to group similar observations into one cluster. Advanced clustering algorithms have been designed for anomaly identification. As illustrated in [19], the density-based spatial clustering of applications with noise (DBSCAN) algorithm could effectively identify anomalies in building energy consumption data. Association rule mining (ARM) aims to extract significant associations among data variables. The knowledge discovered can therefore be applied for anomaly detection, e.g., directly identifying rules on energy waste patterns or detecting observations not fulfilling normal associations [20,21]. Cabrera and Zareipour used ARM to identify abnormal energy waste patterns in lighting systems of educational buildings [22]. Fan et al. developed an anomaly detection engine based on normal associations discovered from massive operational data [23].

To summarize, the requirement of labeled high-quality training data has imposed great constraints on the applicability of supervised anomaly detection methods. Unsupervised anomaly detection is more flexible for practical applications. The main limitations of existing unsupervised anomaly detection methods are: (1) The anomaly detection performance and computational efficiency can be degraded dramatically when applying to big data. For instance, statistical methods are not scalable to large-scale data and they are subject to stringent mathematic assumptions, which may not be fulfilled by real-world high-dimensional data. Some unsupervised data mining techniques have been used to enhance the effectiveness and efficiency in analyzing big data. Nevertheless, the associated post-mining workload can be overwhelming, e.g., selecting useful and non-redundant association rules describing normal or faulty working conditions can be very time-consuming [24,25]. (2) The performance of existing unsupervised methods relies heavily on features used. Currently, features for anomaly detection are selected or constructed based on domain expertise or simple statistics (e.g., the mean and standard deviation of a numeric variable). There is a lack of data-driven methods to automate the feature generation process for generalization purposes. More advanced methods are desired to enhance the performance and applicability of unsupervised anomaly detection in the building field.

One promising solution to these limitations is the autoencoder. An autoencoder adopts the neural network architecture to perform unsupervised learning, where the model input and output are set identical. The rapid development in the deep learning community has provided various techniques for analyzing different types of data (e.g., cross-sectional or temporal data) and training models with advanced architectures (e.g., deep convolutional autoencoders) [26]. More importantly, autoencoders enable a data-driven approach to high-level feature extraction, which can be used to tackle the most challenging task in unsupervised anomaly detection, i.e., feature engineering [27].

To the best of the authors' knowledge, there is a lack of studies to systematically examine the potential of different types of autoencoders in the unsupervised anomaly detection of building operational data. This study is performed to fill this knowledge gap. More specifically, an autoencoder-based ensemble method is proposed for detecting anomalies in building energy data. The autoencoder ensemble is developed considering different autoencoder architectures and training schemes. In addition, novel methods are developed to evaluate the autoencoder performance without the use of anomaly labels. The paper is organized as follows: Section 2 introduces the basics on autoencoders. Section 3 describes the research methodology, including the research outline, the development of autoencoder-based ensembles for anomaly detection and the proposed performance evaluation methods. Section 4 presents a case study. The research results are shown and discussed in Section 5. Conclusions are drawn in Section 6.

2. Basics on autoencoders

2.1. The general autoencoder architectures

Autoencoders can be regarded as a special form of neural network designed for unsupervised learning [28]. The learning process is