Information Sciences 523 (2020) 1–13

Contents lists available at ScienceDirect

Information Sciences

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

TSI: Time series to imaging based model for detecting anomalous energy consumption in smart buildings

Muhammad Fahim*, Khadija Fraz, Alberto Sillitti

Innopolis University, Innopolis, Russia

ARTICLE INFO

Article history:
Received 9 August 2019
Revised 19 February 2020
Accepted 25 February 2020
Available online 26 February 2020

Keywords:
Smart meter
Data analysis
One class support vector machine
Anomaly detection

ABSTRACT

In smart buildings, efficient energy consumption is one of the biggest challenges to solve, which can contribute to reduce the global warming of our planet, due to its relevance. In this paper, a *time series to image (TSI)* based model is introduced to identify anomalous energy consumption in residential buildings. It has a novel encoding scheme to transform univariate time series data into images for extracting the useful information and *one-class support vector machine (OCSVM)* for the classification task. The TSI extracts descriptive and representative feature spaces from the data and encode them into images using *Markov Transition Function (MTF)*. We empirically evaluate the proposed model over publicly available real world dataset and compared the results with the state-of-the-art method. The obtained results are competitive and confirm the applicability of TSI model in real-world scenarios.

1. Introduction

Global warming is gradually increasing the average temperature of our planet. The main known cause is the greenhouse effect. Over the past 50 years, an increased volume of carbon dioxide (CO_2) and other greenhouse gasses have been observed to be released by fossil fuels, land clearing, and other human activities [1] [2]. This change causes extreme weather events such as more frequent wildfires, droughts, heat waves, melting of glaciers, and floods threatening our planet ecosystem [3]. According to Gul et al. [4], developed countries consume a huge amount of energy in buildings, where 50% ~ 65% is consumed by electricity. This consumption is related to various activities including electrical appliances, lightning, heating, ventilation and cooling (HVAC) systems. Moreover, an ample amount of energy is wasted in both residential and commercial buildings due to inefficient management or unawareness [5]. If we know the day-to-day energy usage patterns, it is possible to identify anomalous behaviors by processing the data logs of smart meters. In this way, energy consumption can be optimized by providing information to energy providers and occupants to improve the energy usage as it happens in mobile devices [6–8]. This optimization may contribute to reduce the emission of CO_2 in atmosphere. Furthermore, it can reduce the overhead costs as well as building's energy consumption.

Smart meter is one of the best solutions to monitor energy consumption [9]. It can be easily adopted in both commercial and residential settings because it does not require any additional setup inside the building. Many countries successfully deployed this technology including UK, USA, Germany, China, Switzerland, Denmark, and Brazil [10–12]. Smart meter has the capability to measure the energy usage from few seconds to minutes and transmit this information to smart grid in the

* Corresponding author.
E-mail address: m.fahim@innopolis.ru (M. Fahim).

https://doi.org/10.1016/j.ins.2020.02.069
0020-0255/

form of univariate time-series data [13,14]. The collected logs contain only two columns: a time stamp and an integer value (i.e., energy consumption in KWh). Smart meter logs the energy usage 24 × 7 every few seconds. For instance, the smart meters deployed in UK can log the energy consumption every 6 ~ 8 seconds. In this way, an enormous amount of data is collected. To analyze this data for recognizing anomalous energy consumption is a challenge due to limited information and computationally expensive. The analysis can be organized at four different levels:

1. Long-term analysis (with a timeframe of years)
2. Medium-term analysis (with a timeframe of months)
3. Short-term analysis (with a timeframe of weeks)
4. Very short-term analysis (with a timeframe of days or even hours)

The aim of this research is to analyze the univariate time-series energy data at the fourth level (i.e., very short-term analysis) to detect the anomalous energy consumption in residential buildings. At present, the most common strategy to process the massive data logs is based on the extraction of energy consumption features with the support of domain expert knowledge [15]. The experts elicit useful features for a specific context over a large amount of empirical studies. However, domain experts cannot provide the optimal set of features since environments can become very complex when socio-economics characteristics (e.g., number of people, number of appliances, employment status, etc.) of residential buildings varies from one building to another.

Inspired by recent success in reorganizing time-series data as images, allowing machine learning models to learn classification and recognition task more accurately and efficiently [16,17]. Such approaches can be one of the potential solutions to process the smart meters logs for recognizing the anomalous energy consumption patterns in residential buildings. The TSI model automatically extract multiple features (more details in Section 4.2) from univariate time-series data and encode it into 2D images using MTF. The generated images represent the Markov transition probability with one dimension and temporal dependency with other dimensions. During the classification task, we adopted one-class support vector machine for the following reasons:

1. It does not require the availability of anomaly data patterns.
2. It has the ability to differentiate the normal energy consumption patterns from all other possible complex anomalous cases. Thus, during the prediction phase, it can identify the normal and abnormal patterns in the day-to-day energy consumption.

2. Paper contribution and outline

Our contributions in this work are three-fold. First, we designed a mechanism to automatically extract descriptive and representative features from univariate time series and compact this sparse information into 2D images using Markov transition functions. This procedure does not require to elicit hand crafted features from domain experts for individual residential buildings. Consequently, this image representation further helps the classifier to detect the anomalous behaviour more efficiently. Second, it has a better performance as compared to the existing state-of-the-art method such as Principal Component Analysis (PCA). Third, we reduced the storage cost by converting the feature space of each day into an image.

The rest of this paper is organized as follows: we briefly describe the related work in Section 3. Section 4 provides the details about the proposed TSI-based anomaly detection model. In Section 5, we presents the obtained experimental results followed by comparison to show our improvements. Finally, Section 6 provides conclusions and possible future directions.

3. Related work

Anomalous behavior or outlier detection is a well-known research area that exists in many application domains. A wide range of methods have been developed from simple threshold check to complex artificial neural network models [18–24]. Moreover, computational intelligence-based techniques such as fuzzy models [25], fuzzy clustering [26] and particle swarm optimization using genetic algorithm are also developed for multivariate time-series analysis [27].

In residential building scenarios, the research community is contributing to find anomalous energy consumption. The most common detection techniques are based on regression analysis. In regression-based techniques, the basic idea is to use the actual energy consumption as a baseline. If the real consumption is far from the baseline, then it is classified as an anomaly. In regression-based methods, results are sensitive to the chosen threshold [28].

Jakkula et al. [29] built one of the pioneer model based on a statistical method (i.e., t-distribution) and the k-nearest neighbour clustering algorithm to detect abnormal energy consumption. The energy consumption information is recorded hourly in a home settings while anomalous data have been generated as random data. They mentioned that more robust methods can be constructed using machine learning methods such as Support Vector Machine (SVM), Markov models, or artificial neural networks. Similarly, an early work of Wrinch et al. [30] presents a method to detect anomalies by transforming time series data into the frequency domain and analyzing the energy consumption patterns in buildings. During their analysis, they detect inappropriately configured thermostat as an anomaly and performed experiments over a dataset collected over the five months. However, this dataset is not publicly available.

Chou et al. [31] developed a hybrid model using a neural network and an auto regressive integrated moving average (ARIMA) for daily power consumption. The experiments were performed over the collected dataset of an office building.

Furthermore, they performed prediction of energy consumption. A detailed analysis and in-depth understanding is required to find the optimal hyper parameters.

Guo et al. [32] developed an optimized neural network-based fault diagnosis strategy for variable refrigerant flow air conditioning system to detect anomalies. They achieved good results and proved the effectiveness of their approach. However, the applicability of such approaches are limited since obtaining the high-quality data required is time consuming and obtaining the anomalous data samples is costly and often unfeasible.

Capozzoli et al. [33] developed a methodology based on statistics, artificial neural network, and the clustering technique. They improved fault detection processes by reducing the number of false positives. For identifying anomalies in buildings, they utilized unsupervised density-based spatial clustering of applications with noise (DBSCAN). This method can group all the outliers in one cluster. Unsupervised anomaly detection methods are practical solutions for real applications but computationally expensive in case of massive data.

A more recent work by Fan et al. [15] developed an autoencoder based model that adopts a neural network to perform unsupervised learning. They also highlight the limitations of handcrafted feature extraction since their construction is based on domain experts or simple statistics (e.g., the mean and standard deviation). They also highlight the need to automate the feature generation process for generalization since unsupervised learning heavily depends on the extracted features.

In previous works, a number of approaches have been developed from simple statistical methods to complex unsupervised learning models. However, computationally expensive processes and the need of domain expertise limit their applicability in real life scenarios. To overcome such limitations, we propose an alternative approach to transform time-series data into 2D images without involvement of domain expert. It has the ability to retain the characteristics of time series to distinguish between the normal and abnormal energy consumption patterns. Furthermore, the proposed model reduced the storage costs to store this information. For these reasons, our TSI model has the potential to work with real-world applications.

4. Proposed model

The block diagram of the proposed model for the detection of anomalous energy consumption is presented in Fig. 1. It contains four major components:

1. Data preprocessing
2. Feature extraction
3. Imaging
4. Classifier

Each component is described in detail in the following sections.

4.1. Data preprocessing

Smart metering data is stored in log files as time series (i.e., time-stamp, energy consumption value). The sampling period of the series is about 8 seconds and we define a non-overlapping window of 30 minutes to segment the continuous time-series data. This is a well-proven approach to process energy consumption patterns for many application scenarios [34]. The Fig. 2, presents one week of energy consumption data segmented over a 30 minutes window and average values are retained for further processing.

In Fig. 2, it can be seen that the identification of anomalous energy consumption is not a straight forward process. For instance, it is not possible to set a threshold to classify the segmented windows as a normal or anomalous. A meaningful presentation is required to extract this information, as described in the following section.

4.2. Feature extraction

In this phase, the segmented time-series is transformed to a new space where relevant features are extracted. One of the possible ways to extract these features is calculating basic statistics, performing spectral analysis, or applying signal processing techniques. However, there are approaches to extract features automatically from time-series data [35,36]. We adopted the latter approach that automatically extracts features utilizing the time series feature extraction library based on the scalable hypothesis tests (also known as TSFRESH). It accelerates this process by combining 63-time-series characterization methods, which compute a total of 794 descriptive time-series features from simple mean to power spectral density. The primary features extracted are listed in Table 1.

The complete list of extracted features is available in [36]. Furthermore, this library is computationally efficient in applying distributed and parallel processing on data streams.

4.3. Imaging

To obtain the feature space snapshot, we formulate the imaging procedure by mapping the feature space into image as shown below in Fig. 3:

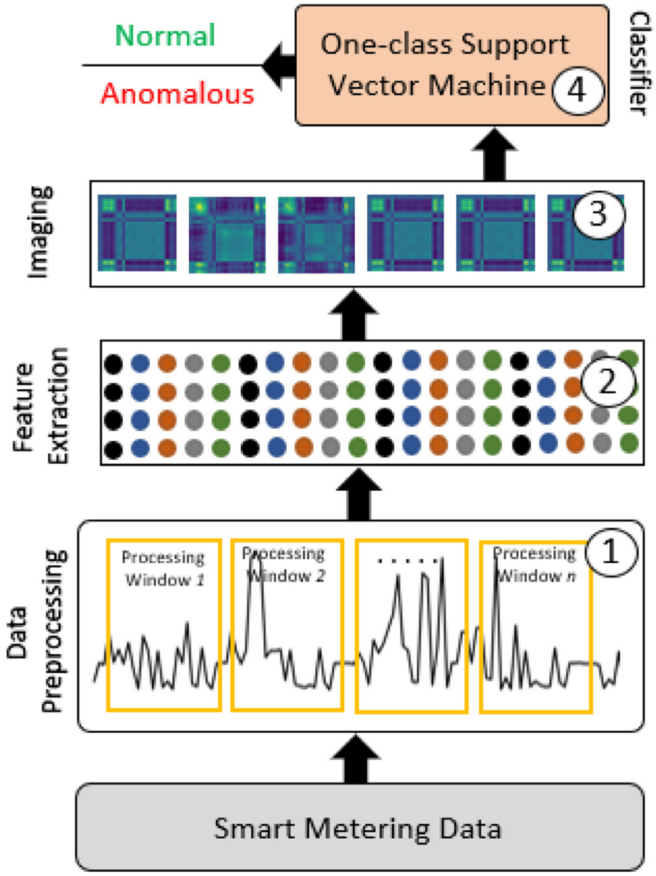

Fig. 1. The proposed model for detection of anomalous energy consumption. The data stream flows from the smart building to data preprocessing block and segmented. Later, features are extracted and transform into images. Finally, OCSVM classify the energy consumption behavior as normal or anomalous.

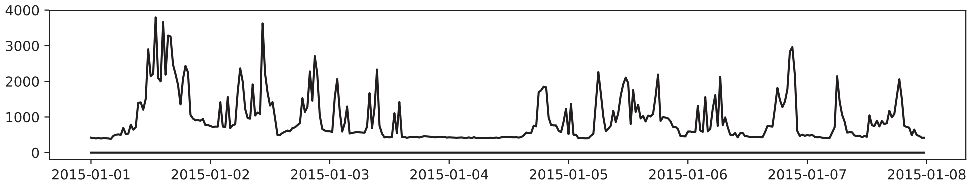

Fig. 2. One week of energy consumption.

Table 1
A subset of the extracted primary features.

Area	Measures
Energy	Absolute energy, energy ratio, FFT coefficients
Statistical	Auto-correlation, entropy, mean absolute change
Linear trend intercept	Standard deviation, max, min, mean and variance
Wavelet	Continuous wavelet coefficients and peaks
Spectral	Spectral density estimation, signal symmetry

Let D is a given time-series data logs that is segmented into $S = \{s_1, s_2 \cdots s_n\}$ segments. In previous section, for every s_i there is a total number of f_m features are extracted. We identify Q quantile bins using Gaussian quantiles and quantify each f_i to the corresponding bins using symbolic aggregation approximation [37]. Every feature vector is guaranteed to non-zero numerical values. This procedure constructs $Q \times Q$ adjacency matrix W by counting transitions among quantile bins in

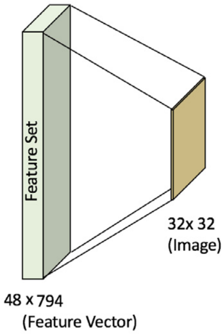

Fig. 3. The transformation of feature space into image.

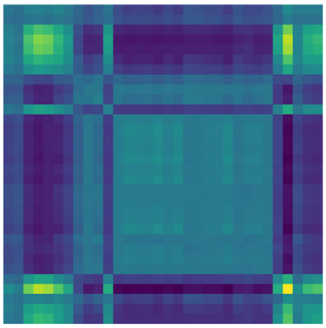

Fig. 4. The one day snapshot based on Markov transition field.

first-order Markov along the time axis. After the calculation of W, we can build the Markov transition field as follows:

$$\mathrm{M} = \begin{bmatrix} m_{1,1} & m_{2,1} & m_{1,3} & \cdots\cdots\cdots & m_{1,m} \\ m_{2,1} & m_{2,2} & m_{2,3} & \cdots\cdots\cdots & m_{2,m} \\ \vdots & \vdots & \vdots & \cdots\cdots\cdots & \vdots \\ m_{n,1} & m_{n,2} & m_{n,3} & \cdots\cdots\cdots & m_{n,m} \end{bmatrix}$$

In addition, we reduce the image size from 48 $\times$ 48 to 32 $\times$ 32 by averaging the pixels in each 2 $\times$ 2 patch using a non-overlapping window method. This approach provides a suitable path to learn complex temporal correlations through visual encoding. The idea is to preserve the inter-relationships of features and compact information through Markov transition probabilities. It is implemented by following steps:

1. The first step deals with the quantization of the feature space by symbolic aggregation approximation and computing the bins.
2. The second step computes the Markov transition matrix considering the quantized features space as a Markov chain.
3. Finally, the Markov transition field is computed from the transition matrix.

This approach reduces the storage cost compared to the direct storage of the logs and extracts features automatically. After encoding the feature space by imaging procedure, we obtain the feature space snapshot that preserve temporal information. The single day image is presented in Fig. 4 and a further discussion is presented in Section 5.

4.4. One-class support vector machine

In the proposed model, the classification of energy consumption into normal or abnormal is based on a one-class support vector machine [38]. The OCSVM is supervised learning model that has the ability to learn over a normal class information. It transforms the energy consumption data into a high dimensional feature space through a kernel function. It iteratively finds the maximal margin hyperplane which best separates the training data from the origin using relaxation parameters. Where origin of the hyperplane is treated as anomalous data points of the considered data. It has ability to construct boundaries around the daily energy consumption. During the test phase, any instance that is outside of defined boundary is considered as an anomalous pattern because the abnormal point is quantitatively deviate from the normal energy consumption observations. Thus, the objective function is trying to maximize the separation between the origin and all positive observations.

$$f(x)\begin{cases}+1 & if\ x\ \in\ normal\ pattern\\ -1 & if\ x\ \in\ abnormal\ pattern\end{cases} \tag{1}$$

In our case, the positive observations are normal energy consumption snapshots. Consider $x_i \in \Re^n$, where $i = 1, 2, 3, ..., m$ and x_i is associated with class $y_i \in \{1\}$. To build the classifier, data points X are mapped to a new feature space H:

$$\phi = X \to H$$

Thus, it follows the general scheme of the support vector machine and it is written as follows:

$$min\ \ \frac{1}{2}\|w\|^2 \tag{2}$$

$$Subject\ to\ \ \{w\cdot\phi(x_i) + b \geq 0\ \ for\ all\ i.$$

Where w is a weight vector and b a bias. In feature space, linear separation becomes much easier than input space by defining a kernel function. The role of kernel function is to maximize the separation of the origin and the normal energy consumption. However, not all datasets are linearly separable and such cases lead to infinite solution [39]. In order to make an acceptable solution, error limits ν are set with slack variable ξ_i, that allow the the process to incur in errors [40]. The value of ν is set to 0.1 in our solution. The modified optimization problem can be rewritten as follows:

$$min_{w,\xi_i,b}\ \ \frac{1}{2}\|w\|^2 + \frac{1}{\nu m}\sum_{i=1}^{m}\xi_i - b \tag{3}$$

$$where\ \ \begin{cases} w\cdot\phi(x_i) - b + \xi_i \geq 0,\\ \xi_i \geq 0,\\ \nu \in (0,1]. \end{cases}$$

This optimization problem can be solved using quadratic programming and well known tools are available for solving it [38]. We use linear kernel function $K(x, y) = (x.y)$ to construct the boundary around the normal energy patterns. In this case, the decision function f(x) becomes:

$$f(x) = sign(w\cdot\phi(x)) - b) \tag{4}$$

In Eq. 4, the decision function will be positive for most of the data points x_i (i.e., obtained from imaging procedure). Training and testing of the classifier are described in the following subsections.

4.4.1. Train model

In this phase, the model is trained to learn the normal energy consumption patterns. Fig. 5 shows the training process of the model.

In Fig. 5, the univariate time-series data is preprocessed and the model is saved after training over the energy consumption patterns. During the training phase, we have only normal energy consumption patterns, so we split the data into training and validation set by a ratio of 70% and 30%, respectively. During the training of OCSVM, we select linear kernel which requires one parameter ν and its value should be in the interval of (0,1]. In our training procedure, it is set to 0.1. This corresponds to an upper bound on the fraction of training errors and a lower bound of the fraction of support vectors. We obtain the value of ν experimentally and it may vary in different problems. It is expected to classify the normal and abnormal energy consumption patterns during the testing phase.

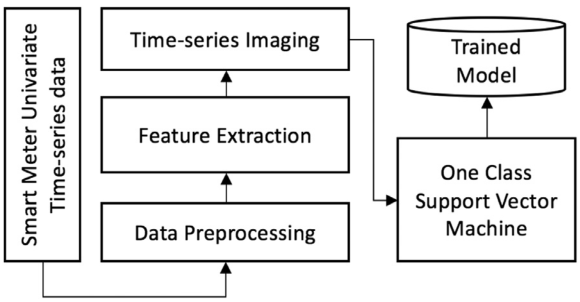

Fig. 5. The model training mechanism to learn the normal energy consumption patterns.

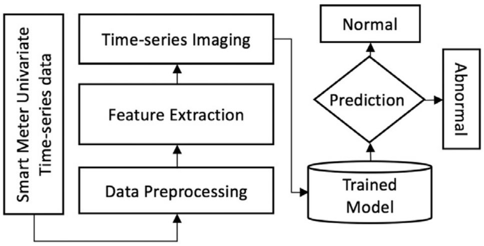

Fig. 6. The model testing mechanism to predict the normal or abnormal energy consumption.

4.4.2. Test model

In testing phase, unseen energy consumption patterns are presented to predict the class labels as normal or abnormal. Testing process of the classifier is presented in Fig. 6.

During the test phase, energy consumption pattern is processed through the designed pipeline and classified by OCSVM trained model. Fig. 6 shows that trained model is invoked for the prediction of the energy consumption.

5. Results and discussion

In this section, we present the results of the proposed model to measure the accuracy level and validate the feasibility of the time-series imaging technique in the energy consumption domain. We process the univariate time-series data over a day because it simplifies the identification of anomalous behaviors in time avoiding issues related to the seasonality patterns. In each day, the data is segmented over a period of 30 minutes and a huge number of features are extracted (i.e., 794 features). These features are further processed to generate the images of feature spaces – as described in Section 4.3. To learn the energy consumption patterns, we trained a one-class support vector machine over the normal energy consumption. In the training step, we have only information about the normal class, hence, the classifier boundaries are set around it. Whatever is outside of these boundaries is considered as an anomalous energy consumption.

5.1. Dataset

The experiments are performed over the REFIT electrical load measurement dataset [41]. This dataset contains smart meter readings and the individual appliances energy usage. The dataset was collected over a period of two years from 20 houses in the UK. During the data collection, the occupants were conducting their usual routines. The energy consumption depends on the number of occupants, the electrical appliances, and the number of bedrooms. Therefore, generalizing the model can mislead the abnormalities because each house has different energy consumption patterns. The experiments are performed by considering the different socio-economics characteristics (Table 2).

Furthermore, there is no information about the presence of anomalies in the dataset. However, it does not mean that they do not exist. In our experiments, we generate the augmented anomalies to evaluate the performance of the model. The following section describes the details to generate the anomalous data.

5.2. Synthetic anomalies generation

In anomaly detection task, mostly we have the training examples of normal behavior of the system. In real-world, it is difficult to generate the anomalous behaviors because of the waste of resources or even sometimes it is not possible due to system constrains. From the model learning prospective, it is good to construct the boundaries around the normal class and consider everything that does not belong to it as abnormal. For such a task, research community propose synthetic anomaly generation [29,42]. A simple approach to generate the anomalies is to choose the feature values randomly in the domain. In this case, the generated anomalies can be easily identified because it has no correlation with time scale. The alternative approach is distribution-based anomaly generation. In our case, we did not assume a prior distribution of the data but generate anomalous data that is very close to the existing distribution. To achieve this, we generate the Imposter

Table 2
Characteristics of dwellings.

	House 1	House 2	House 3
No. of Occupants	2	4	4
No. of Appliances	35	44	28
Size	4-bedrooms	3-bedrooms	4-bedrooms
Dwelling Type	Detached	Semi-detached	Detached

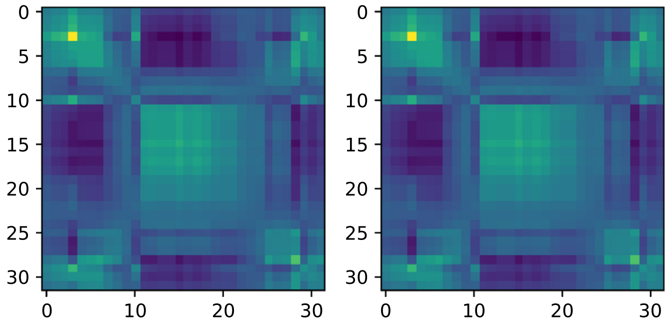

Fig. 7. The normal and abnormal energy consumption patterns after injecting synthetic anomalies.

day_1	day_2	day_8	day_{10}	day_{16}	day_{17}	day_{19}	day_{23}

Fig. 8. The randomly selected days (injected anomalies) House 1.

Data (*ID*) and added back to the original data as follows:

$$ID \sim UD(0, 1) \tag{5}$$

$$Anomaly = X + ID \tag{6}$$

where, *UD* is uniform distribution and we generate imposter data by setting the limits of variance to low = 0 and high = 1. In the next step, we are adding this one to the original time-series (i.e., *X*) to generate the anomalies. To evaluate the model, we inject eight days (about 1/4 of the whole month) of synthetic anomalies for *House 1, House 2*, and *House 3* and analyze the model performance.

In Fig. 7, the left-side image is the normal energy consumption projection for one day, while the right side is anomalous energy consumption. Both images look similar but they are technically different and this small change needs to be detected in real life scenarios. The developed model needs to find these patterns as an anomaly energy consumption.

5.3. Performance measures

Evaluating the model has entailed using three standard metrics precision, recall, and F1-score are applied as performance measures [43]. In anomaly detection system, high precision and high recall are desired to build a good system. In such a situation, F-measure is used to give an equal importance to precision and recall. The precision, recall and F1-score metrics are calculated by Eqs. 7, 8 and 9 respectively.

$$Precision = \frac{TP}{(TP + FP)} \tag{7}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{8}$$

$$F1 - Score = \frac{2 * Precision * Recall}{(Precision + Recall)} \tag{9}$$

5.4. Experimental results of House 1

In this experimental setup, we consider the residential house with 2 occupants (i.e., couple), 4-bed rooms, 35 electrical appliances, and one person is employed. The days are randomly selected from a month and synthetically made them anomalous. These selected days are listed in Fig. 8. The results of TSI model is presented in Fig. 9.

Fig. 9 shows that the model miss-classifies only two normal days as anomalous ones (i.e., Day 8 and 23), while all others are classified correctly.

5.5. Experimental results of House 2

For the second building, we repeat the experimental setup and train the model to detect anomalous energy consumption. The residential building have different numbers of electrical appliances and inhabitants compared to the house 1 (ref.

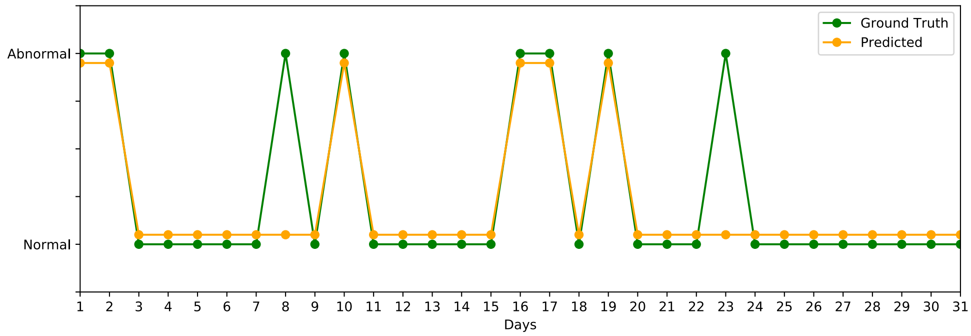

Fig. 9. The TSI model prediction for anomalous energy consumption for House 1.

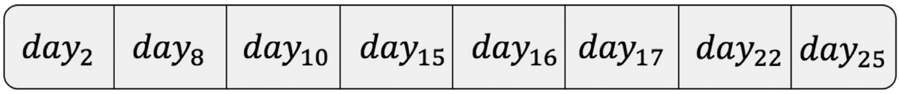

Fig. 10. The randomly selected days (injected anomalies) House 2.

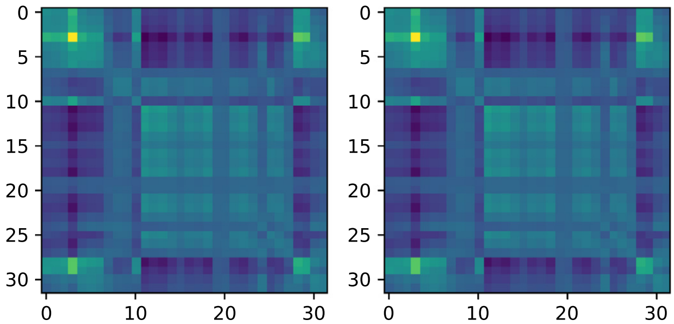

Fig. 11. Imaging of feature spaces of a day (House 2).

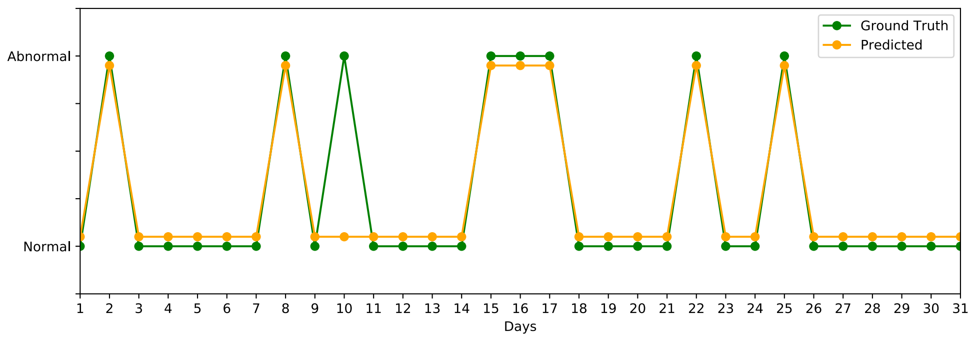

Fig. 12. TSI model prediction for anomalous energy consumption in House 2.

Table 2). As discussed, the collected data do not have any anomalous energy consumption information. Therefore, we generate eight days of data for the second building to evaluate TSI model. The randomly selected days are listed in Fig. 10. The visualization of normal and generated abnormal energy consumption is presented in Fig. 11.

The obtained results of TSI model is presented in Fig. 12. In case of House 2, model predict each day correctly, while only confuse day 10 with normal energy consumption pattern.

5.6. Experimental results of House 3

The characteristics of house 3 is presented in Table 2. In Fig. 13, we present the randomly selected days and prepared them as an anomalous energy consumption. The obtained results are presented in Fig. 14, only three days are confused with normal days while all other days are predicted as abnormal.

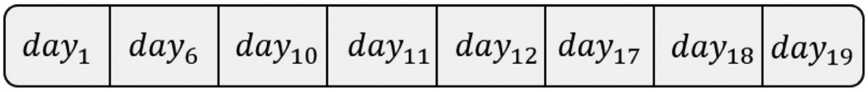

Fig. 13. The randomly selected days (injected anomalies) House 3.

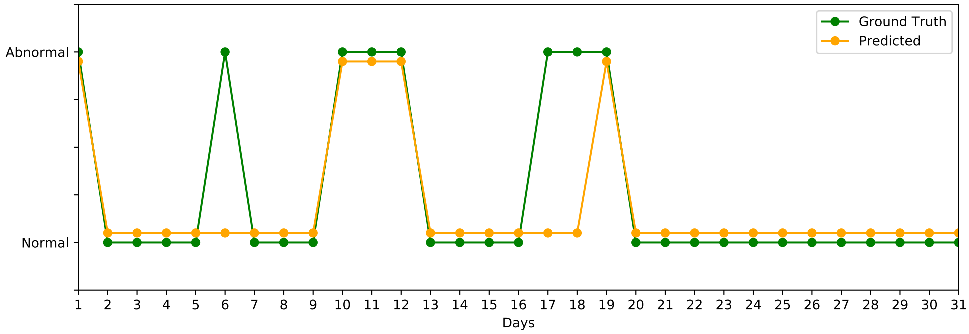

Fig. 14. TSI model prediction for anomalous energy consumption in House 3.

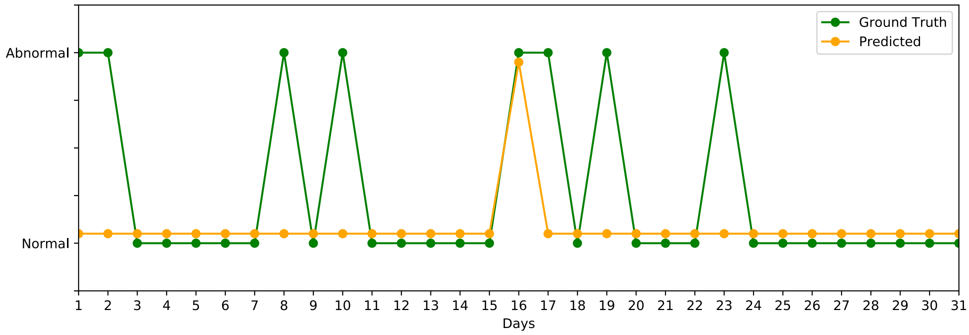

Fig. 15. The prediction results in House 1 (feature Space + OCSVM).

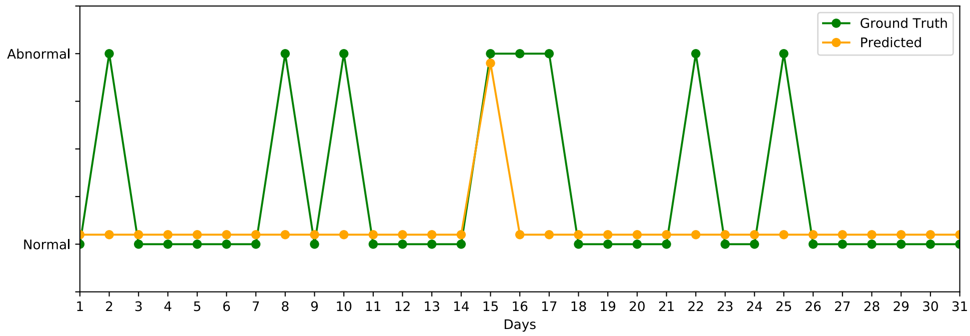

Fig. 16. The prediction results in House 2 (feature Space + OCSVM).

5.7. Comparison

We compare the obtained results produced by the proposed model (Fig. 1). The following two procedures are designed to compare our TSI model results.

5.7.1. Without imaging

In this comparison procedure, the main goal is to analyze the impact of imaging and find if there is a significant difference between the obtained results. We fed the feature spaces directly to OCSVM for anomaly detection instead of imaging. The results are presented in Figs. 15, 16 and 17 for House 1, 2 and 3 respectively.

In Fig. 15, only one day is predicted anomalous while all other seven days are considered normal. In Fig. 16, only one day is predicted correctly, while all other days are considered as normal. Similarly, Fig. 17 demonstrates that only two days

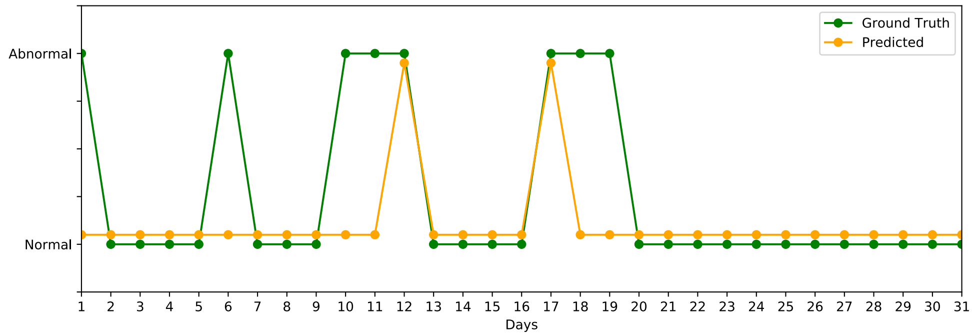

Fig. 17. The prediction results in House 3 (feature Space + OCSVM).

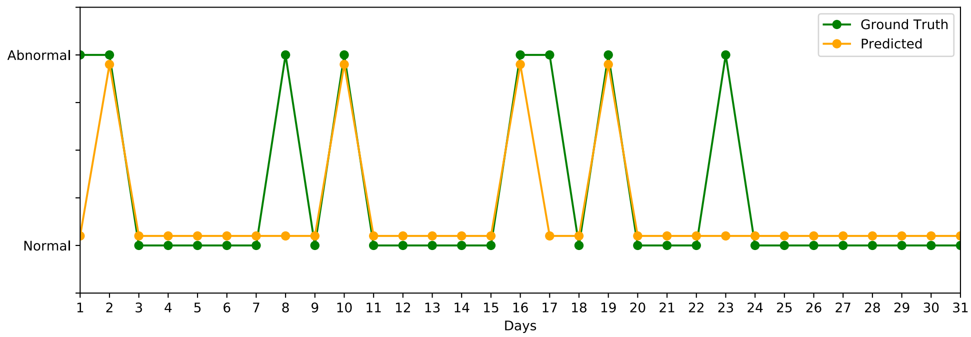

Fig. 18. The prediction results in House 1 (PCA + OCSVM).

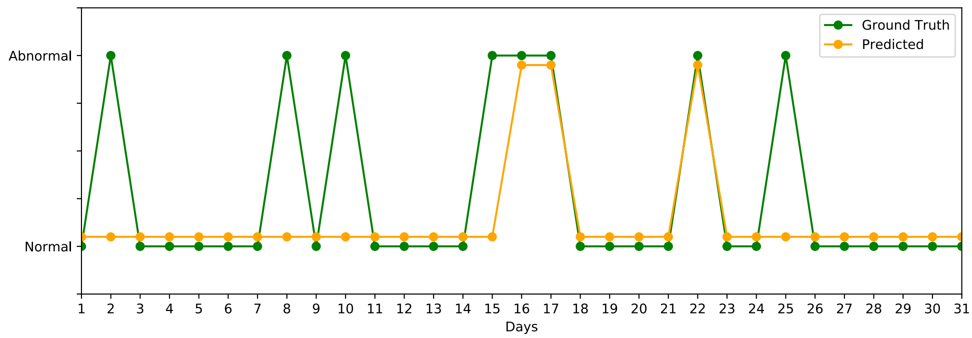

Fig. 19. The prediction results in House 2 (PCA + OCSVM).

are predicted correctly while six days are considered normal. In all theses cases, it does not mean that classifier fails to recognize anomalies, but the input space to the learning model is too sparse that leads to these inaccurate results.

5.7.2. Principal component analysis (PCA) for dimensionality reduction

This second comparison procedure is based on the principal component analysis to show the effectiveness of the proposed model. We reduce the sparsity of the feature space by applying PCA over the obtained feature spaces of all analyzed houses. We present the obtained results in Figs. 18, 19 and 20 for House 1, 2 and 3 respectively.

In Fig. 18, the classifier classifies 50% of days as normal and 50% as abnormal. For the above analysis, it can be seen that feature extraction technique requires a further dimensionality reduction procedure that is able to retail the temporal characteristics. Fig. 19 illustrates that 60% of the anomalous days are considered normal. Only day 16, 17, and 22 are predicted correctly while the others are misclassified as normal days. Similarly, Fig. 20 presents the three correct predictions while all others as normal days. In all cases, we can see that PCA helps to improve the prediction. It still suffers to present the true representation of feature space as compared to TSI model. We presented Precision, Recall and F1-Score results in Fig. 21 for House 1, 2 and 3 respectively. The obtained results are highly accurate and stable in all experimental setups.

In order to make the proposed model applicable in real life, we also calculate the storage requirement. The storage requirements of each approach is reported in Table 3. If the only purpose of the data collection is the detection of the

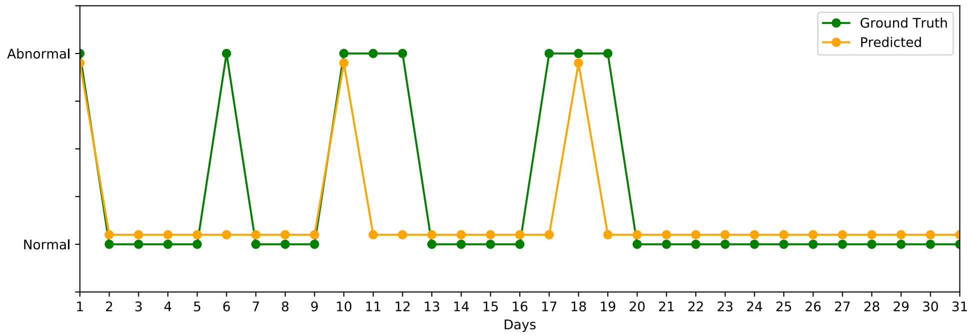

Fig. 20. The prediction results in House 3 (PCA + OCSVM).

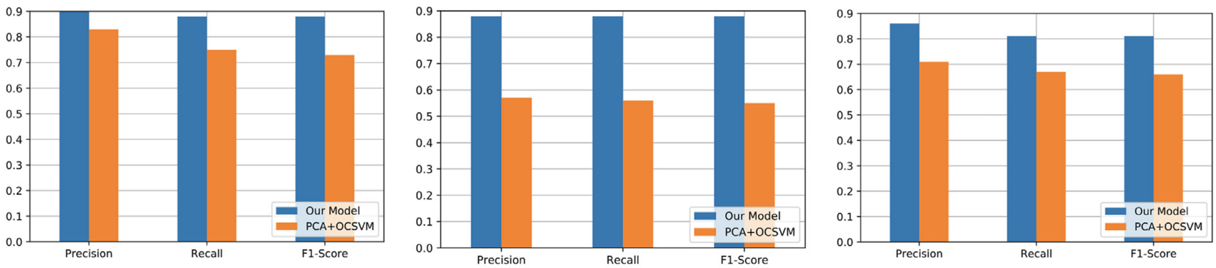

Fig. 21. The comparison results of House 1, 2 and 3 respectively.

Table 3
Required storage for 2 years of data.

	Storage Space
Log Streams	400.8 MB
Feature Spaces	26.4 MB
Imaging	1.46 MB

anomalous energy consumption patterns, then our solution can be considered, due to its storage efficiency and suitable processing model for mobile devices.

6. Conclusion

Smart meters provide an excellent opportunity to collect high resolution electricity data and its analysis for diverse potential applications, including the detection of anomalous energy consumption. In this paper, a TSI model is presented to process the data stream of smart meters to classify each day of consumption as normal or abnormal. Our approach relaxes the domain expert constrain to elicit the hand crafted feature by extracting the descriptive and representative features automatically. We compact this information into 2D images. Our classification model is based on OCSVM to learn the normal behaviour of energy consumption and any deviation is considered as anomalous. The experiments are performed on real-world publicly available datasets to demonstrate the applicability of TSI model. We obtained 88% F1-score in our experimental settings, which is 24% higher – average case of F1-scores – compared to state-of-the-art methods.

In our future work, we planned to detect the abnormal behaviour at appliance level for each day. We will extend this model to obtain the time-series image of different electrical appliances inside the building. We are also planning to make this model as a service to handheld devices for consumers to see the energy usage patterns.

Declaration of Competing Interest

The authors confirm that they don't have any conflict of interest with anyone about this research.

References

[1] T.R. Anderson, E. Hawkins, P.D. Jones, CO2, the greenhouse effect and global warming: from the pioneering work of arrhenius and callendar to today's earth system models, Endeavour 40 (3) (2016) 178–187.
[2] J.G. Canadell, C. Le Quéré, M.R. Raupach, C.B. Field, E.T. Buitenhuis, P. Ciais, T.J. Conway, N.P. Gillett, R. Houghton, G. Marland, Contributions to accelerating atmospheric co2 growth from economic activity, carbon intensity, and efficiency of natural sinks, Proc. Nat. Acad. Sci. 104 (47) (2007) 18866–18870.

[3] R. Swart, J. Robinson, S. Cohen, Climate change and sustainable development: expanding the options, Climate Policy 3 (sup1) (2003) S19–S40.
[4] M.S. Gul, S. Patidar, Understanding the energy consumption and occupancy of a multi-purpose academic building, Energy Build. 87 (2015) 155–165.
[5] T.A. Nguyen, M. Aiello, Energy intelligent buildings based on user activity: a survey, EnergyBuild. 56 (2013) 244–257.
[6] L. Corral, A.B. Georgiev, A. Sillitti, S. G., A method for characterizing energy consumption in android smartphones, 2nd International Workshop on Green and Sustainable Software (GREENS 2013), 2013.
[7] L. Corral, A.B. Georgiev, A. Sillitti, S. G., Method reallocation to reduce energy consumption: an implementation in android os, 29th ACM Symposium on Applied Computing (SAC 2014), ACM, 2014.
[8] L. Corral, A.B. Georgiev, A. Sillitti, S. G., A study of energy-aware implementation techniques: redistribution of computational jobs in mobile apps, Sustain. Comput. Informat. Syst. 7 (2015).
[9] M. Fahim, A. Sillitti, Anomaly detection, analysis and prediction techniques in iot environment: a systematic literature review, Access 7 (2019).
[10] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: applications, methodologies, and challenges, IEEE Trans. Smart Grid (2018).
[11] T. Zufferey, A. Ulbig, S. Koch, G. Hug, Forecasting of smart meter time series based on neural networks, in: International Workshop on Data Analytics for Renewable Energy Integration, Springer, 2016, pp. 10–21.
[12] A. Abuadbba, I. Khalil, X. Yu, Gaussian approximation-based lossless compression of smart meter readings, IEEE Trans. Smart Grid 9 (5) (2018) 5047–5056.
[13] S.S.S.R. Depuru, L. Wang, V. Devabhaktuni, N. Gudi, Smart meters for power gridâ;;challenges, issues, advantages and status, in: 2011 IEEE/PES Power Systems Conference and Exposition, IEEE, 2011, pp. 1–7.
[14] M. Carratù, M. Ferro, A. Pietrosanto, V. Paciello, Smart power meter for the iot, in: 2018 IEEE 16th International Conference on Industrial Informatics (INDIN), IEEE, 2018, pp. 514–519.
[15] C. Fan, F. Xiao, Y. Zhao, J. Wang, Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data, Appl. Energy 211 (2018) 1123–1135.
[16] Z. Wang, T. Oates, Imaging time-series to improve classification and imputation, in: Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.
[17] T. Hur, J. Bang, J. Lee, J.-I. Kim, S. Lee, et al., Iss2image: a novel signal-encoding technique for cnn-based human activity recognition, Sensors 18 (11) (2018) 3910.
[18] Q. Liu, S. Li, X. Liu, N. Linge, A method for electric load data verification and repair in home environment, Int. J. Embed. Syst. 10 (3) (2018) 248–256.
[19] M. Fahim, A. Sillitti, Analyzing load profiles of energy consumption to infer household characteristics using smart meters, Energies 12 (5) (2019).
[20] L. Martí, N. Sanchez-Pi, J.M. Molina, A.C.B. Garcia, Anomaly detection based on sensor data in petroleum industry applications, Sensors 15 (2) (2015) 2774–2797.
[21] M. Goldstein, S. Uchida, A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data, PloS one 11 (4) (2016) e0152173.
[22] C. Hegedus, P. Ciancarini, A. Franko, A. Kancilija, I. Moldovan, G. Papa, S. Poklukar, M. Riccardi, A. Sillitti, P. Varga, Proactive maintenance of railway switches, in: 5th International Conference on Control, Decisions and Information Technologies (CoDIT 2018), 2018.
[23] A. Capozzoli, M.S. Piscitelli, S. Brandi, D. Grassi, G. Chicco, Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings, Energy 157 (2018) 336–352.
[24] S. Mahadevan, S.L. Shah, Fault detection and diagnosis in process data using one-class support vector machines, J. Process Control 19 (10) (2009) 1627–1639.
[25] J. Soto, O. Castillo, P. Melin, W. Pedrycz, A new approach to multiple time series prediction using mimo fuzzy aggregation models with modular neural networks, Int. J. Fuzzy Syst. 21 (5) (2019) 1629–1648.
[26] C. Gupta, A. Jain, D.K. Tayal, O. Castillo, Clusfude: forecasting low dimensional numerical data using an improved method based on automatic clustering, fuzzy relationships and differential evolution, Eng. Appl. Artif.Intell. 71 (2018) 175–189.
[27] F. Gaxiola, P. Melin, F. Valdez, J.R. Castro, O. Castillo, Optimization of type-2 fuzzy weights in backpropagation learning for neural networks using gas and pso, Appl. Soft Comput. 38 (2016) 860–871.
[28] Y. Zhang, W. Chen, J. Black, Anomaly detection in premise energy consumption data, in: Power and energy society general meeting, 2011 ieee, IEEE, 2011, pp. 1–8.
[29] V. Jakkula, D. Cook, Outlier detection in smart environment structured power datasets, in: Intelligent Environments (IE), 2010 Sixth International Conference on, IEEE, 2010, pp. 29–33.
[30] M. Wrinch, T.H. El-Fouly, S. Wong, Anomaly detection of building systems using energy demand frequency domain analysis, in: Power and Energy Society General Meeting, 2012 IEEE, IEEE, 2012, pp. 1–6.
[31] J.-S. Chou, A.S. Telaga, Real-time detection of anomalous power consumption, Renew. Sustain. Energy Rev. 33 (2014) 400–411.
[32] Y. Guo, G. Li, H. Chen, J. Wang, M. Guo, S. Sun, W. Hu, Optimized neural network-based fault diagnosis strategy for vrf system in heating mode using data mining, Appl. Therm. Eng. 125 (2017) 1402–1413.
[33] A. Capozzoli, F. Lauro, I. Khan, Fault detection analysis using data mining techniques for a cluster of smart office buildings, Expert Syst. Appl. 42 (9) (2015) 4324–4338.
[34] C. Beckel, L. Sadamori, T. Staake, S. Santini, Revealing household characteristics from smart meter data, Energy 78 (2014) 397–410.
[35] D.M. Burns, C.M. Whyne, Seglearn: a python package for learning sequences and time series, J. Mach. Learn. Res. 19 (1) (2018) 3238–3244.
[36] M. Christ, N. Braun, J. Neuffer, A.W. Kempa-Liehr, Time series feature extraction on basis of scalable hypothesis tests (tsfresh–a python package), Neurocomputing (2018).
[37] J. Lin, E. Keogh, L. Wei, S. Lonardi, Experiencing sax: a novel symbolic representation of time series, Data MiningKnowl. Discov. 15 (2) (2007) 107–144.
[38] B. Schölkopf, J.C. Platt, J. Shawe-Taylor, A.J. Smola, R.C. Williamson, Estimating the support of a high-dimensional distribution, Neural Comput. 13 (7) (2001) 1443–1471.
[39] H.J. Shin, D.-H. Eom, S.-S. Kim, One-class support vector machinesâ;;an application in machine fault detection and classification, Comp. Indust. Eng. 48 (2) (2005) 395–408.
[40] B. Schölkopf, A.J. Smola, F. Bach, et al., Learning with kernels: support vector machines, regularization, optimization, and beyond, MIT press, 2002.
[41] D. Murray, L. Stankovic, V. Stankovic, An electrical load measurements dataset of united kingdom households from a two-year longitudinal study, Scientific Data 4 (2017) 160122.
[42] W. Fan, M. Miller, S. Stolfo, W. Lee, P. Chan, Using artificial anomalies to detect unknown and known network intrusions, Knowl. Inform. Syst. 6 (5) (2004) 507–527.
[43] E. Alpaydin, Introduction to machine learning, 2004.