Information Fusion 91 (2023) 93–102

Contents lists available at ScienceDirect

Information Fusion

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

Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges

Gen Li, Jason J. Jung *

Department of Computer Engineering, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, 06974, Republic of Korea

ARTICLE INFO

Keywords:
Anomaly detection
Multivariate time series
Research challenge

ABSTRACT

Anomaly detection has recently been applied to various areas, and several techniques based on deep learning have been proposed for the analysis of multivariate time series. In this study, we classify the anomalies into three types, namely abnormal time points, time intervals, and time series, and review the state-of-the-art deep learning techniques for the detection of each of these types. Long short-term memory and autoencoders are the most commonly used methods for detecting abnormal time points and time intervals. In addition, some studies have implemented dynamic graphs to examine relational features between the time series and detect abnormal time intervals. However, anomaly detection still faces some limitations and challenges, such as the explainability of anomalies. Many studies have focused only on anomaly detection methods but failed to consider the reasons for the anomalies. Therefore, increasing the explainability of anomalies is an important research topic in anomaly detection.

1. Introduction

Anomaly detection is a data mining task that aims to identify data that significantly differ from most other data. Many applications of anomaly detection in different contexts have been developed, including financial card fraud, industrial intrusion, and medical seizure detection. A large number of time series are recorded from a wide range of fields, including medicine, industry, and the natural sciences. The analysis of time series with the aim of extracting useful information has become a major field of study. The extracted information can be used to solve a variety of problems. Anomaly detection on time series, as one of a problem, aims to discover the abnormal event from the time series recorded from variety areas. For example, Dwivedi et al. The paper [1] used an ensemble feature selection from a computer network and the grasshopper optimization approach to improve the detection of network intrusions through support vector machines. Siniosoglou et al. [2] used an autoencoder-based generative adversarial network to identify anomalies in a smart electrical grid system.

The tradition methods for anomaly detection are based on the statistical indices and density of the dataset. However, there some limitations for the tradition methods. For example, the performance of traditional methods in anomaly detection on medical images and sequential datasets is terrible because they cannot capture complex structures in the data. In addition, it is impossible for tradition methods to extent to large-scale data to find anomalies. Also, the boundary between the abnormal and normal behavior is often imprecisely defined in time series data and is constantly evolving. This lack of well-defined representative normal boundaries poses a challenge to traditional method. Therefore, the deep learning-based methods are proposed. Deep anomaly detection technology learns hierarchical discrimination features from the time series data. This automatic feature learning capability eliminates the need for domain experts to manually develop features. So that is advocated to solve the anomaly detection problem on time series.

There are lots of survey papers have introduced the deep learning methods for anomaly detection on time series. Such as the paper [3] reviewed the methods and applications related detecting abnormal time series, and summarized the challenges for the methods. However, this paper only considers the one dimensional time series, and there are lots of multivariate time series are recorded from the real-world. In addition, the existing survey papers related to the time series anomaly detection did not introduce the commonly used datasets and the research challenge. Therefore, to fill these gaps, we reviewed the papers related to the methods, applications, research challenge, and datasets for anomaly detection on multivariate time series.

Because the input data and types of anomalies involved in the analysis of time series can vary considerably, distinct anomaly detection technologies are required. As shown in Fig. 1, there are three main types of anomalies, which are abnormal time point, time interval, and

* Corresponding author.
E-mail address: j3ung@cau.ac.kr (J.J. Jung).

https://doi.org/10.1016/j.inffus.2022.10.008
Received 3 April 2022; Received in revised form 16 September 2022; Accepted 9 October 2022
Available online 17 October 2022

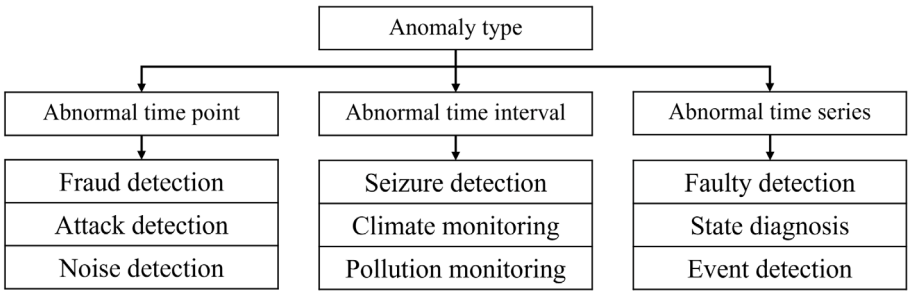

Fig. 1. Types of anomalies and applications.

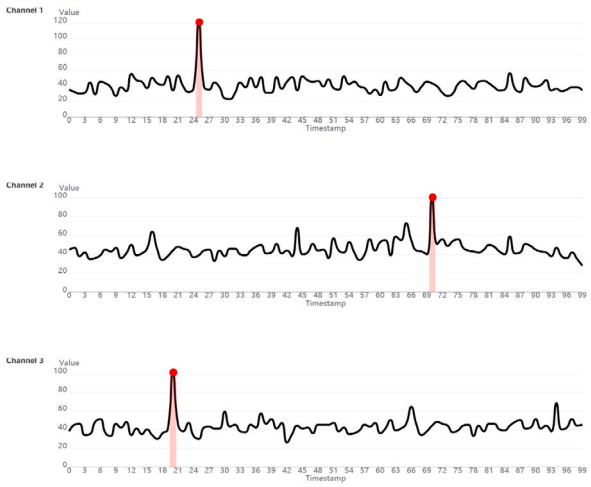

Fig. 2. Example of an abnormal time point.

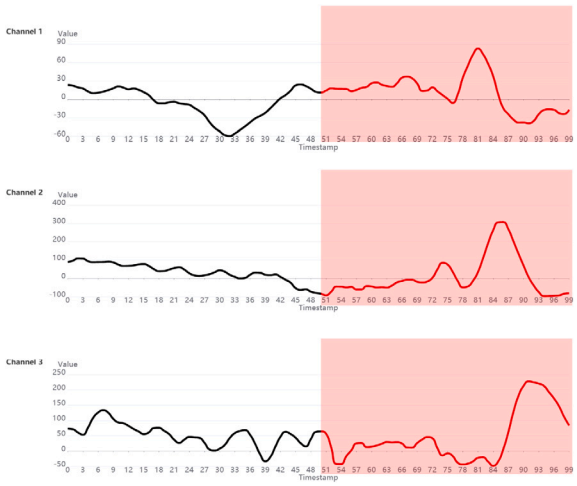

Fig. 3. Example of an abnormal time interval.

time series. The mainly applications based on the abnormal time point detection are fraud detection, attack detection, and noise detection. The abnormal time interval detection is commonly used for seizure detection, climate monitoring, and pollution monitoring. The faulty detection, state diagnosis, and event detection are based on the abnormal time series detection.

An abnormal time point is a certain time point t_i at which the value of the monitored variable is significantly different from those at any other time points. This is formulated as $A(t_i) < \theta$ or $A(t_i) > \theta$, where t_i is the ith time point, $A(t)$ is the abnormal score for time t, and θ is the threshold for detecting the anomaly. An abnormal time point in a time series generally corresponds to an extreme value. In some cases, the anomalies correspond to noise in the data. Because in such cases the anomalies affect the data analysis results, they must be detected and removed from the given time series. In other cases, this type of anomalies may represent events of interest. For example, Fig. 2 shows three recorded electroencephalograms (EEGs) in which three extreme values are displayed. These abnormal points indicate either noise or a relevant event at that time point. Therefore, detecting this type of anomalies is useful for time series analyses. Ranjan et al. [4] improved the traditional prediction method based on time window sliding to detect anomalies in a given time series and increase forecasting accuracy for the planning and operation of a power system. Similarly, Wang et al. [5] applied a convolutional neural network (CNN) to detect anomalies in a time series collected from a wind power system to improve short-term prediction.

Abnormal time interval. An abnormal time interval is a certain time interval t_i, during which the variable behaves significantly differently than during the rest of the time series. This is formulated as $A(t_i) < \theta$ or $A(t_i) > \theta$, where t_i is the ith time interval, $A(t)$ is the abnormal score for time t, and θ is the threshold for detecting the anomaly, typically indicating the occurrence of an unusual event. Analysts may focus on these anomalies to explore the meaning of the detected events. Fig. 3 presents an example of an abnormal time interval in three EEG signals, where the red areas correspond to an epileptic episode suffered by a patient. Li and Jung [6] proposed an anomalous time interval detection method to identify seizures from multivariate EEG signals. Because brain discharge is quite different with and without a seizure, the authors considered the seizure interval as an abnormal time interval and proposed a graph-based detection approach. In this case, detecting this type of anomalies can help experts analyze and understand the patterns of abnormal events. Chen et al. [7] presented an approach for detecting network attacks using Internet of Things (IoT) time series. This study helped experts recognize and analyze the patterns of network attacks.

Abnormal time series. An abnormal time series is a certain time series s_i, which is significantly different from other time series. This is formulated as $A(s_i) < \theta$ or $A(s_i) > \theta$, where s_i is the ith time series, $A(s)$ is the abnormal score for time s, and θ is the threshold for detecting the anomaly. Fig. 4 shows three time series where two of the signals exhibit high correlation, in contrast with the one in red, which is thus detected as an abnormal time series [8]. Such as Li and Jung [9] discover the relationship between the stock and financial indies. Then, a dynamic graph is used for modeling these relationships. Finally, a abnormal stock is detected by proposing a graph embedding model.

The remainder of this paper is organized as follows. In Sections 2, 3, and 4 technologies for abnormal point, abnormal time interval and abnormal time series detection are discussed, respectively. In Section 5, the approaches for selecting the threshold are discussed. In Section 6, applications based on anomaly detection in multivariate time series are reviewed. In Section 7, available open-access time series datasets for anomaly detection are listed. In Section 8, limitations and

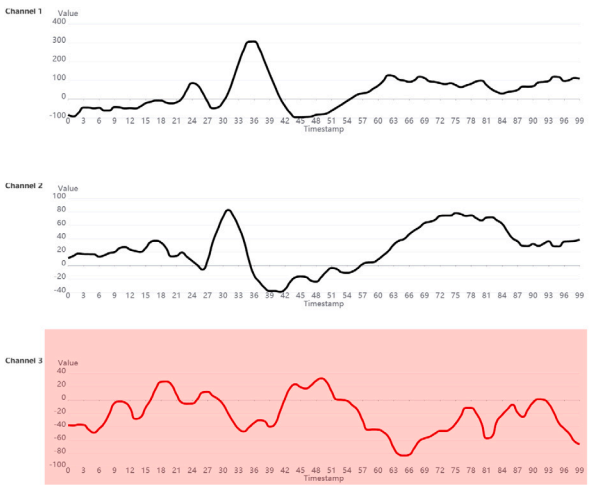

Fig. 4. Example of an abnormal time series.

research challenges in anomaly detection in multivariate time series are presented. Finally, in Section 9, conclusions are discussed.

2. Abnormal time points

An abnormal time point is commonly found in multivariate time series, representing noise or an event that occurs at a certain instant. This section reviews the technologies available for detecting abnormal time points in multivariate time series. Fig. 2 provides an example of the detection of abnormal time points. Three time series corresponding to the EEG signals of a patient recorded from three channels are shown. Each signal exhibits an anomaly, but because the three anomalies do not occur at the same instant, there are no correlations between them.

Because an abnormal time point in a multivariate time series consists of an extreme value, the basic procedure for detecting this type of anomaly is to predict or estimate the value at a certain timestamp, and subsequently compute the difference between such value and the ground truth. If the difference between the predicted or estimated value and the ground truth value is larger than a set threshold, the value is detected as an anomaly. The basic approach is formulated as follows.

$$|x_t - \hat{x}_t| \rangle \tau, \tag{1}$$

where τ is the threshold, x_t is the ground truth value at the tth timestamp, and $\hat{x}_t$ is the predicted or estimated value for x_t.

The most commonly used approaches for detecting abnormal time points in multivariate time series are based on prediction models. With the development of deep learning technologies, numerous models have been proposed to predict time series. For example, a time series is formulated as $X = \{\langle x_t \rightarrow x_{t+1}\rangle | x_i \in X, t \in [0, T]\}$, where x_t is the value at time stamp t, and x_{t+1} is the value at the next timestamp after x_t. To detect whether the value x_t is an anomaly, the prediction model utilizes past values x_{t-i} to predict the value $\hat{x}_t$ at the current x_t. Nguyen et al. [10] applied the LSTM model to predict the values in a time series and utilized an autoencoder (AE) to extract features from it. The authors first input past values x_{t-i} into the LSTM model and predict the current value $\hat{x}'_t$. Subsequently, assuming it as independent, they input the predicted value to an autoencoder to obtain a new predicted value $\hat{x}_t$. Finally, they compute the error between the predicted and true values using a loss function $L = \sum_i^N \|x_i - \hat{x}_i\|$, where N is the number of timestamps in the time series.

For example, Girish and Rao [11] applied an LSTM model to a cloud environment to detect abnormal events in multivariate time series. However, this study only applied a simplistic LSTM architecture to detect anomalies in the multivariate time series collected from the cloud environment, making the model unsuitable for other applications. Ullah et al. [12] detected anomalies in surveillance networks by first extracting spatial features from the multivariate time series through convolutional neural networks and then inputting the extracted features into a bidirectional LSTM (Bi-LSTM) model to predict future data. Finally, a statistical method was applied to detect whether the predicted data were anomalous. Qin et al. [13] input the controller area network bus into an LSTM model to predict the future values of the network and then identified network attacks by obtaining the loss of the LSTM. If the loss was greater than a threshold, the corresponding timestamp was detected as an anomaly.

However, several forms of anomalies exist, and the time series does not designate them. Therefore, several papers have been published to address this issue, most of them attempting to employ unsupervised learning models [14]. For example, Kieu et al. [15] reconstructed an input using two autoencoder models. In general, regular data is more abundant than anomalous data. Therefore, when the autoencoder receives anomalous inputs, the reconstruction loss becomes significant and, if a threshold for the autoencoder loss is set, outliers can be detected. Their model outperformed the baseline models in experiments. Other researchers have examined the possibility of merging a convolutional neural network with an autoencoder to identify an abnormality. This includes the work by Yin et al. [16] who proposed the use of a recurrent neural network to record temporal information to enhance the performance of the integration model for anomaly identification. However, the integration model, cannot demonstrate enhanced performance.

Predicting multivariate time series is difficult because the changes in the variables involved are affected by many factors. Therefore, the effectiveness of prediction-based models in detecting anomalies in multivariate time series is limited. To address this, some studies have aimed to combine LSTM and autoencoders to detect abnormal time points. The core idea in LSTM-autoencoder (LSTM-AE) models is to utilize the LSTM unit as a hidden layer of the autoencoder to reconstruct the input data [17] and identify the anomaly by obtaining the difference between the input x_t and reconstructed $\hat{x}_t$ values. A threshold τ is also set to detect anomalies using Eq. (1).

Other approaches utilize supervised learning models to detect anomalies in multivariate time series. For instance, Yeung et al. [18] treated anomaly detection as a class-imbalance problem, labeling the classes for each timestamp and applying supervised learning models to learn the features of the labeled data. Similarly, Su et al. [19] proposed studying the patterns of normal data in multivariate time series using a supervised learning model to identify anomalies. Because the model learns the patterns of the normal data, the accuracy of the model is affected if the anomaly is included in the learning stage. Therefore, an anomaly is discovered based on the loss of the supervised model.

For example, Maleki et al. [20] applied one such hybrid model to detect anomalies in a multivariate time series. Chen et al. [21] also applied an LSTM-AE model to detect abnormal time points of motor failure in wind turbine systems. LSTM-AE models have been applied for anomaly detection in a variety of areas, such as industrial IoT and sports [22,23]. Additional technologies for detecting abnormal time points are summarized in Table 1.

3. Abnormal time intervals

An abnormal time interval in multivariate time series corresponds to a period during which an atypical event occurs. This section reviews approaches for detecting abnormal time intervals in multivariate time series.

Table 1
Abnormal time point detection methods.

Reference	Dataset	Architecture	Threshold	Feature extraction	Evaluation metric
Nguyen et al. [10]	C-MAPSS dataset	LSTM and Autoencoder	Optimal threshold	No	AUC, MSE, and F1-score
Girish and Rao [11]	Cloud environment dataset	LSTM	Fixed threshold	No	Accuracy
Ullah et al. [12]	UCF-Crim	CNN and LSTM	No threshold	Yes	AUC and ROC
Qin et al. [13]	Vehicle dataset	LSTM	Fixed threshold	Yes	AUC and ROC
Kieu et al. [15]	ECG dataset	Recurrent autoencoder	Fixed threshold	No	PR-AUC
Reunanen et al. [14]	Sensor dataset	Autoencoder	Optimal threshold	No	Recall, AUC, and ROC
Yin et al. [16]	IoT dataset	Autoencoder	Optimal threshold	No	F1-score
Yeung et al. [18]	Financial dataset	LSTM	Specific threshold	No	F1-score
Chen et al. [21]	Wind turbine dataset	LSTM and autoencoder	Adaptive threshold	No	MSE and MAE
Liu et al. [22]	ECG	LSTM and autoencoder	Optimal threshold	Yes	Sensitivity
Maleki et al. [20]	Amazon web services	LSTM and autoencoder	Optimal threshold	No	F1-score and accuracy
Homayouni et al. [23]	Covid-19 dataset	LSTM and autoencoder	Specific threshold	Yes	MSE
Munir et al. [24]	IoT	CNN	Adaptive threshold	Yes	F1-score
Ding et al. [25]	Numenta anomaly benchmark	LSTM	Optimal threshold	Yes	F1-score
Kim and Cho [26]	Web traffic dataset	LSTM	Optimal threshold	Yes	F1-score

3.1. LSTM-based approaches

An LSTM-based autoencoder model utilizes an LSTM unit to extract temporal features from the multivariate time series, which are then used by the autoencoder to reconstruct the input time interval. The idea behind the use of an autoencoder model for anomaly detection is to obtain the loss between input and output [27–31]. For example, Lin et al. [32] combined a variational autoencoder model (VAE) with LSTM model to detect anomalies in multivariate time series of ambient and machine temperatures. First, the autoencoder model was utilized to extract local features from the multivariate time series. Subsequently, LSTM units captured temporal information from the local features. Finally, an anomaly detection score was defined to be equal to the loss between the input and output time intervals.

Considering that generative adversarial networks (GANs) have been rapidly developed, Niu et al. [33] improved the LSTM-based VAE model by adding a GAN. The study indicated that although GANs have been used for anomaly detection in many areas, they still require a considerable amount of time for training. The proposed LSTM-based VAE-GAN model attempts to address this challenge. With such an objective in the study, first, all the hidden layers of the encoder, decoder, generator, and discriminator were established using LSTM units. Next, the time interval in the multivariate time series was input into a VAE model to compute low-dimensional embedding vectors that were subsequently input to the generator. Subsequently, the output of the generator and ground truth were input to a discriminator for training. Finally, the test data was fed into the combination unit to calculate the average anomaly score for each input time interval and, if the score was greater than a set threshold τ, the input time interval was identified as an anomaly. However, this study had some limitations. First, the authors utilized LSTM and VAE to construct a GAN model but reported that the proposed model could reduce the training time. This is because the LSTM-based VAE-GAN model has a more complex structure than the LSTM-based VAE model and, hence, requires more training time than the latter. Second, the proposed model was a supervised learning model, which cannot perform better nor be applied in more cases than an unsupervised learning model.

3.2. Dynamic graph-based approaches

The use of dynamic graphs to represent time series is being widely used to analyze anomalies. The core idea of anomaly detection using dynamic graphs is to utilize the relationships between multivariate time series to construct a single graph for each time interval. The advantage of these methods is that they identify relationship patterns that can be used as a basis for anomaly detection. For illustration, Fig. 5 presents the time series recorded from three different channels, Ch_{01}, Ch_{02}, and Ch_{03}, where T_i is the ith time interval and the red area corresponds to a time interval anomaly. Graphs are constructed by computing the correlation coefficients between the three time series and assigning them as the weights of the edges in the graphs, with $e(G_i)$ being the entropy of graph G_i. Because the relationships between the multivariate time series at time interval with red area are quite different from those at other time intervals, the time interval with red area is detected as an abnormal time interval.

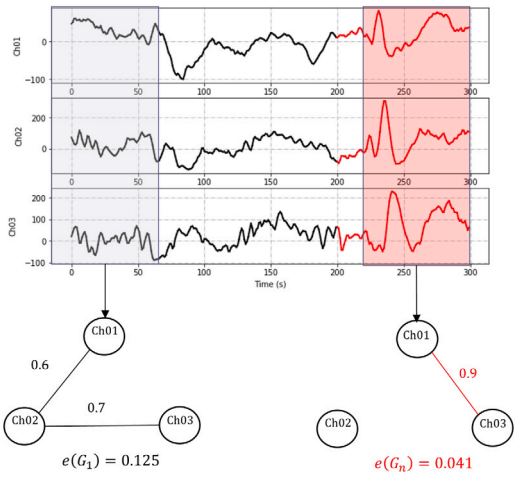

Fig. 5. Dynamic graph-based approach.

Because graphs are useful to model objects and their relationships, several studies have applied dynamic graphs to represent multivariate time series and their correlations in continuous time intervals. For example, Deng and Hooi [34] attempted to combine a deep learning structure and a graph neural network (GNN) to detect time interval anomalies. In this study, the multivariate time series were divided into several time intervals, which were, in turn, transformed into static graphs where each vertex corresponded to one of the multivariate time series and each edge represented the relationship between a pair of time series. A graph attention model was then proposed to forecast future multivariate time intervals and these were used for comparison to detect the anomalies. In another study, Zhao et al. [35] utilized 1D convolutional kernels to extract features from multivariate time series, modeled these features and their relationships through graphs. Then they applied a graph convolutional network (GCN) for detecting anomalies through graph prediction. Chen et al. [7] applied series-to-graph technology to IoT data series collected from three sensors. Therefore, the constructed graph comprised three nodes, each representing the time series collected from a single sensor. To detect whether there was a connection between two nodes, they applied the Gumbel-Softmax sampling strategy to sample a vector. Finally, a graph learning model was used to find the anomaly in the multivariate IoT time series.

Table 2
Abnormal time interval detection methods.

Reference	Dataset	Architecture	Threshold	Feature extraction	Evaluation metric
Lin et al. [32]	Amazon Web Services	LSTM-AE	Fixed threshold	No	F1-score
Provotar et al. [27]	Various signal dataset	LSTM-AE	Optimal threshold	No	Accuracy
Niu et al. [33]	Yahoo dataset	VAE-GAN	Optimal threshold	Yes	F1-score
Liu et al. [28]	MIT-BIH arrhythmia dataset	LSTM-CNN	Adaptive threshold	Yes	Sensitivity and specificity
Wu et al. [29]	IIoT dataset	LSTM	Optimal threshold	Yes	AUC, ROC, and F1-score
Vos et al. [30]	Gearbox dataset	LSTM	Specific threshold	No	F1-score
Chang et al. [31]	Fall detection dataset	LSTM-CNN	Specific threshold	No	Sensitivity and specificity
Deng and Hooi [34]	Sensor dataset	GNN	Optimal threshold	Yes	F1-score
Zhao et al. [35]	NASA dataset	GCN	Non-parametric decision threshold	Yes	F1-score
Chen et al. [7]	Secure water treatment dataset	GNN	Adaptive threshold	Yes	F1-score
Li and Jung [36]	Meteorological dataset	GNN	Specific threshold	Yes	F1-score
Li and Jung [38]	Meteorological dataset	GNN	Specific threshold	Yes	F1-score
Li et al. [39]	Traffic dataset	GNN	Specific threshold	Yes	F1-score

Relationship patterns are important for anomaly detection. The most commonly used relationship is the correlation. However, correlation ignores the causality among the multivariate time series. To address this problem, in a previous study, we proposed identifying spurious relationships as patterns [9,36]. A spurious relationship is defined as that between two correlated time series that do not have a causal relationship. One such relationship was presented in a study from the Netherlands in which a correlation was found between the number of storks nesting during the spring and the number of human children born during the same period, despite there being no causal connection because both variables were due only to the meteorological conditions nine months prior to the observations. Similarly, Hofer et al. [37] examined post-reunification Germany and found no correlation between clinical deliveries and stork populations but found an association between out-of-hospital births and a rise in stork populations. A summary of the approaches for abnormal time interval detection is presented in Table 2.

4. Abnormal time series

The detection of abnormal time points and intervals can reveal events of interest and useful information in multivariate time series, which explains why many technologies have been proposed to detect them. The goal of abnormal time series detection is to identify time series whose relationships with other time series in the dataset are significantly different from those observed between the remaining series. Abnormal time series detection also is important since it indicates that a time series exhibits a long-term abnormal behavior. This abnormal behavior can lead to a high risk or be a reason for the risk. Therefore, detecting abnormal time series can help people analyze the reason for the risk and avoid it. For example, during analyzing climate change, we have to record several time series, such as pressure, temperature. In a long-term period, the weather exhibits a high temperature. We can think the long-term high temperature is a reason of the climate change.

However, the approaches for detecting abnormal time series are less sensitive than those for detecting abnormal time points and intervals. The main step in the detection of abnormal time series is to compare the similarity between all the time series involved. This implies that the computing required to detect the anomalies increases considerably with the scale of the time series dataset. Beggel et al. [40] used them to describe the shape of normal variables and detect outlier time-dependent variables in a multivariate time series. The shapelets are extracted from each time series and a deep neural network is applied to learn the patterns of the shapelets. Then, the abnormal time series are detected based on the patterns of the shapelets. Cheong et al. [41] proposed a spatiotemporal convolutional network-based relational network (STCNN-RN) to learn the correlations among the financial companies. Then, they applied the genetic algorithms to discover the abnormal company by fitting the STCNN-RN model. In our previous study [9], we identified the relational features between the stock and financial indices. Then, a dynamic graph is created for modeling these features. The dynamic graph embedding model is proposed to map the dynamic graph. Finally, we detect an abnormal stock in the embedding space. Although the anomaly in our study is defined as a certain time interval, the goal of the study is to identify an abnormal stock at a certain time interval, which is an abnormal time series detection. Also, Li et al. [42] establish a GAN model by using the LSTM and autoencoder models. The reconstruction loss and discrimination loss are defined to calculate the anomaly score. Finally, the threshold is selected by optimizing the GAN. In that study, LSTM unit is used to extracted feature from the multivariate time series and create the generator and discriminator. Autoencoder is used for establishing an anomaly detector. Audibert et al. [43] utilized an autoencoder to reconstruct the multivariate time series. The abnormal time series are detected by obtaining the loss of the autoencoder. A summary of the approaches for abnormal time interval detection is presented in Table 3.

Table 3
Abnormal time series detection methods.

Reference	Dataset	Architecture	Threshold
Beggel et al. [40]	UCR dataset	LSTM	Specific threshold
Cheong et al. [41]	Stock data	STCNN-RN	Specific threshold
Li and Jung [9]	Stock data	GCN	Specific threshold
Li et al. [42]	Secure water treatment	GAN	Optimal threshold
Audibert et al. [43]	Secure water treatment	Autoencoder	Optimal threshold

5. Threshold selection

The threshold is an important parameter, which decides the final performance for the anomaly detection on multivariate time series. The most typical method for anomaly detection is based on the box-plot. In our previous study [9], we tried to applying the dynamic graph embedding method to establish an embedding space for mapping the time intervals. hen, the box-plot method is used to detect the anomaly in the embedding space. The threshold for the box-plot method is based on two values that are lower and upper quartiles formulated as $Q1$ and $Q3$, respectively. The interquartile range (I) are calculated by using these values formulated as $I = Q3 - Q1$. It is deemed abnormal if the data is more than $Q3 + 1.5I$ or less than $Q1 - 1.5I$.

The other threshold selection method is non-parametric thresholding method. The method supports that the detecting value is denoted as y_t the truth value is denoted as $\hat{y}_t$. In this case, the error between the detecting and truth values can be formulated as $e(t) = |y_t - \hat{y}_t|$. Then the exponentially weighted averages is calculated for the errors formulated as $e_s = \{e_s^t | t \in [0, t]\}$. The threshold α is calculated by using the function $\alpha = \mu(e_s) + \delta(e_s)$. The studies [44,45] utilized the LSTM model to predict the time series and applied the non-parametric dynamic threshold to detect the anomaly based on the prediction errors.

Su et al. [46] proposed an adjusted peak over threshold method for automated anomaly threshold selection that outperforms the non-parametric dynamic thresholding method. Li et al. [47] proposed a FluxEV method that is a fast and effective unsupervised framework for time series anomaly detection. The paper also applied the peak over threshold method to automatically decided a threshold for anomaly detection.

In addition, many studies selected the threshold by optimizing the deep learning model. The optimal threshold is trained based on the loss function of the deep learning model [10,14,16]. Firstly, these studies computed the loss for each sample. Then, the distribution of the losses can be obtained. Finally, the threshold is decided based on the distribution of the losses. Such as Yin et al. [16] utilized the autoencoder to reconstruct the multivariate time series. Then, the loss between the reconstructed results and the inputted time series is computed. They computed the loss of each time interval and obtain the distribution of the losses. Finally, they selected the threshold for detecting the abnormal time interval based on the statistical indices.

The adaptive threshold also is popular on the anomaly detection, which can select the threshold automatically by using the specific algorithm [21,28]. Such as Chen et al. [21] provided an adaptive threshold based on the support vector regression. Firstly, they discover a mapping relationship between the parameters P in the deep learning model and the final performance value a. The mapping function is denoted as $\hat{f} = P \rightarrow a$. When the time series S^t is inputted to the model, the corresponding parameters P^t and final performance value a^t can be computed, which is formulated as $\hat{f}(S^t)$. Then the threshold is formulated as $\theta = \hat{f}(S^t) + \gamma$ where γ is a expected performance index value. In this case, if the performance a^t is larger than the threshold, the time interval S^t is detected as an anomaly.

Further, some studies selected the specific threshold for detecting the anomaly. Such as the paper [6] selected several thresholds to conduct the experiments. Then, the threshold is decided based on the final performance. Also, the threshold in some areas is already decided based on the experiences. Such as in the financial area [18], the threshold of the jump detection is set as 4.60001 based on the study [48].

6. Applications

6.1. Abnormal time point-based applications

Abnormal point on the time series commonly is an extreme value caused by some reasons. This point in some case represent a noise or an abnormal activity. Such as a certain time point when the Distributed denial of service (DDoS) attacks in the industrial IoT (IIoT) environment. For example, Huong et al. [49] proposed a VAE-LSTM model to detect attacks to IIoT data. The basic idea was to compute the loss of the VAE model to detect abnormal activity. Distributed denial of service (DDoS) attacks are typical anomalies in the IoT environment, which lead to a server being inoperable for an extended period, causing services to breakdown under severe load. Sharma et al. [50] applied PCA to reduce the dimensions of IoT time series and utilized a statistical method to detect DDoS attacks in the reduced IoT time series. Zhan et al. [51] proposed a novel method that represents a time series with different hierarchical layers allowing the extraction of temporal and significant amplitude and detecting time points with abnormal features. They applied their method to the UCR time series datasets [52], and the experimental results showed that their method performed better than the baselines.

Although, abnormal point detection plays an important role in financial area, such as fraud detection. With the rise of technology and the continuous growth of the modern social economy, fraud has become more common in the financial industry, causing hundreds of billions of dollars in losses to institutions and consumers every year. Fraudsters continuously come up with new ways to take advantage of the loopholes found in the current safety measures implemented by the financial sector. The challenge in fraud detection is the insufficient labeled data in financial time series, which prevents the existing supervised learning models from effectively detecting fraud. To address this problem, Xiao and Jiao [53] proposed utilizing a multivariate instance learning model to detect fraud when the availability of labeled data is limited. The theory of the proposed model is based on bag-space learning, which constructs several bags using a clustering method. The label of a cluster depends on the labeled data in the cluster. If the data are from a time point that was part of a fraud cluster, they are marked as fraud. Some studies have detected fraud by predicting user behavior. For example, Benchaji et al. [54] collected transaction data for 180 days from several customers. They then input these data into an LSTM model to predict customer behavior. The given dataset was labeled as fraudulent or normal. The proposed method utilizes supervised learning to detect credit card fraud by detecting abnormal behavior.

In addition, abnormal point detection on smart home and smart city areas also is valuable. In the context of smart cities, IoT can help recognize anomalous road conditions. Moreover, route-finding apps that utilize data from smart machines to detect high-congestion locations and provide alternative routes to customers are becoming commonplace. For example, a user's location data may be used to discover and aggregate odd movement patterns, which may help them avoid congested regions and lessen the environmental effect of their travel [55]. Small IoT networks are being deployed in business and residential buildings at an increasing rate, which generate data that may be utilized to examine and enhance energy efficiency [56–58].

6.2. Abnormal time interval-based applications

Abnormal time interval detection are commonly applied on the biosignals such as seizure detection in EEG signals. Seizure is the brain exhibit an abnormal discharge at a certain time interval. Therefore, most of seizure detection methods are based on the abnormal time interval detection [59–61]. For example, Martini et al. [62] applied a self-supervised learning model to detect seizures from EEG signals in real-time. Furthermore, according to statistical records, safety accidents caused by driver fatigue lead to more than 1.3 million deaths each year in the worldwide. Considering this situation, Chaabene et al. [63] utilized a deep CNN model to detect drowsiness from EEG signals. The experimental results showed that the proposed method achieved an accuracy of 90.42%. The goal of this study was to monitor the diver's brain states to reduce the losses caused by fatigue from driving. Electrocardiography (ECG) is another important tool for monitoring human health. Many artificial intelligence technologies have been applied to detect anomalous changes in ECG signals and allow for timely measures to be taken during the onset stage of illnesses [64,65].

Levels of airborne pollution in urban areas are another major concern worldwide. Several studies have demonstrated that networked sensors can be used to detect and monitor pollution levels in cities [66]. Activity monitoring base on smart home data has been proposed to support assisted living circumstances, whereby an individual's usual routines are learned and large deviations are flagged as abnormal, providing improved awareness to caregivers or health services.

6.3. Abnormal time series-based applications

As a consequence of increased monitoring and sensing within the power network, there has been a change in energy regulation. Smart meters are currently being installed across networks in several countries. These devices can monitor the power they use at different times and automatically send information to the network operator. Monitoring electricity consumption has the benefit of allowing power companies to proactively detect flaws in the local distribution network rather than relying on customers to warn them of outages.

Table 4
Summary of anomaly detection-based application and their approaches.

Reference	Anomaly type	Model	Threshold	Application
Huong et al. [49]	Point	VAE-LSTM	Kernel quantile estimator	Attack detection
Sharma et al. [50]	Point	LSTM	Non-parametric decision threshold	Attack detection
Benchaji et al. [54]	Point	LSTM	Non-parametric decision threshold	Fraud detection
Xiao and Jiao [53]	Point	LSTM	Non-parametric decision threshold	Fraud detection
D'Andrea and Marcelloni [55]	Point	LSTM	Different thresholds	Incident detection
Wijayasekara et al. [56]	Point	Fuzzy rule-based model	Sensitivity threshold	State detection
Araya et al. [58]	Point	Neural network	Pattern recognizer engine	State detection
Chou et al. [57]	Point	Neural network	Specific threshold	Fraud detection
Zhan et al. [51]	Point	LSTM	Specific threshold	State detection
You et al. [59]	Interval	VAE	Optimal threshold	Seizure detection
Rajinikanth et al. [60]	Interval	Neural network	Specific threshold	Seizure detection
Varone et al. [61]	Interval	Neural network	Optimal threshold	Seizure detection
Martini et al. [62]	Interval	LSTM	Adaptive threshold	Seizure detection
Chaabene et al. [63]	Interval	CNN model	Specific threshold	State detection
Forooghifar et al. [65]	Interval	Neural network	Optimal threshold	Diagnosis
Erhan et al. [64]	Interval	Neural network	Optimal threshold	Diagnosis
Jain and Shah [66]	Interval	Neural network	Specific threshold	Pollution monitoring
Cheong et al. [41]	Series	CNN	Specific threshold	Stock data
Madurawe et al. [67]	Series	Graph embedding	Specific threshold	Stock data

Furthermore, investors face significant risks of financial market instability, which can include a market crash caused by systematic hazards and extreme stock price volatility generated by a false hype. Detecting anomalies in the stock market may enable investors to reduce losses due to unstable stock market conditions. Cheong et al. [41] proposed a relation-network-based model to detect stock market anomalies. First, they utilized a CNN to capture the spatial and temporal features of companies. A graph was then used to model the relationships between companies. Finally, an anomaly was detected by applying a genetic algorithm.

Madurawe et al. [67] also constructed a graph to detect stock market anomalies. In their graph, the vertices corresponded the investor and the edges between two vertices represented the interactions. First, they used a sliding time window to divide the stock data into several time intervals. Next, for each time interval, they constructed a graph to model investors and their interactions. Subsequently, they applied a clustering method to group the graphs. In this case, the density of the abnormal graph is low. Finally, they applied the local outlier factor method to detect anomalies in clustered graphs. The anomaly detection-based application and their approaches are summarized in Table 4.

7. Datasets

Because tagging the label of the anomaly in the multivariate time series is time-consuming, the datasets within the labels are small. For example, the segment of the seizure in EEG signals is short, and doctors need to spend a certain amount of time diagnosing and identifying the incidence interval. In this study, we provided multivariate time-series datasets from different fields.

Four datasets were available for anomaly detection in the industrial IoT field. The first is the power system attack dataset published by [68], which represents the power system attack dataset. The dataset consisted of 128 time series acquired from the electric power grid. There were 37 different situations to choose from, including natural disasters, no disasters, and attack disasters. The situations are separated into five types, which are short-circuit fault, line maintenance, remote tripping command injection, relay setting modification, and data injection. Short-circuit faults are the most common scenario. The attack events are labeled in the provided dataset.

The studies [69–71] provide three different gas pipeline datasets. The new gas pipeline dataset consists of 20 separate elements divided into three categories, which are network information, payload information, and labeling information. The network information contains the station address of the gas pipeline network and the conditions of the timestamps corresponding to the station address and circumstances. Information regarding the system control and state of the gas pipeline control system, such as its mode of operation, pressure measurement, and control scheme, is recorded in the payload information. The label contains the current state of the gas pipeline, which indicates whether the pipeline is experiencing an abnormality on the current date. In the presented dataset, status was classified as either normal or anomalous.

Table 5
Summary of commonly used time series dataset for anomaly detection.

Reference	Dataset	Number of series	Label
Pan et al. [68]	Industrial IoT	56	No
Beaver et al. [69]	Gas pipeline	128	Yes
Morries and Gao [70]	Gas pipeline	12	Yes
Morries and Thornton [71]	Gas pipeline	7	Yes
CHB-MIT [72]	EEG dataset	23	Yes
Schulze-Bonhage [75]	EEG dataset	23	Yes
Kramer [74]	ECoG dataset	64	Yes
Andrzejak [76]	EEG dataset	128	Yes

The most frequently used dataset in the medical field is the CHB-MIT scalp EEG database [72]. It contains data from 23 children who had epilepsy, and the Children's Hospital of Boston used international 10–20 sensors to record all EEG signals from their brains. In addition, the study [73] also provided other EEG datasets. The dataset gathered by the epilepsy center at the University of California in Berkeley [74] collects electrocorticogram from a patient. The dataset [75] was gathered by the Epilepsy Center at the University Hospital of Freiburg. The final dataset [76] consisted of EEG recordings obtained from five patients.

In addition, in other areas, there are some public time-series graph datasets for event detection and a time-series point dataset on multivariate time series. The Yahoo dataset [77] comprises actual and synthetic time series containing labeled anomalous points. The dataset assesses the accuracy with which different anomaly categories, such as outliers and change points, may be detected. The synthetic dataset is composed of time series with variable degrees of trend, noise, and seasonality. The commonly used time series dataset for anomaly detection are summarized in Table 5.

8. Research issues and challenges

Although approaches and applications for anomaly detection in multivariate time series are developing rapidly, there are some research issues and challenges in this area. Here, we list some open issues for anomaly detection in multivariate time series that detect anomalies at an early stage, understand the meaning of the anomaly, and explain the reason for the anomaly.

The first is detecting anomalies early in a multivariate time series. The most common strategy for finding anomalies early is to forecast future changes in a multivariate time series. For example, because proactive virtual machine migration is frequently lengthy and time-consuming, it is vital to detect potential future faults as soon as feasible. The detection of network congestion is another significant case. Transportation system congestion is a serious problem [78]. It is vital to execute the proactive detection of an anomaly as soon as possible to reduce the possibility of future detrimental consequences. Ref. [79] provides a proactive anomaly detection ensemble (ADE) technique for dealing with the aforementioned issues, with a focus on early anomaly detection. To establish its accuracy, each strategy was examined over time for a particular dataset type. ADE detects any anomaly in the incoming data based on previous learning for data measuring the same parameters as the training dataset. This method provides a weighted anomalous window based on historical data used to train the model.

It is difficult to identify anomalies before they occur by predicting multivariate time series. For certain non-periodic time series, the prediction performance is influenced by a variety of factors. Therefore, maintaining a high level of accuracy in early anomaly detection is an important research topic. Herein, we provide a new method for detecting anomalies at an early stage. We assumed that the anomaly in the multivariate model was caused by a small change in the historical time point. For example, in traffic systems, a high vehicle speed is one of the reasons for accidents. If the speed of a vehicle rapidly increases, the probability of an accident is high. The change in vehicle speed can be extracted to detect accidents early. There are also many problems that need to be solved using this idea. High speed is not the only reason for the accident, and the accident depends not only on speed. Therefore, discovering the relationship between anomalies and small changes is a research challenge.

The second issue is to discover the meaning of these anomalies. Existing studies have mainly focused on detecting anomalies in a given dataset. It is important to understand the meaning of these anomalies. The types of abnormal patterns varied in the multivariate time series. Some rare patterns are often ignored, but these anomalies are likely to lead to severe disasters. For example, in a financial time series, people mainly focus on fraud detection. However, in some cases, the financial time series includes abnormal patterns related to the financial crisis. These patterns are not the purpose of fraud detection and are easily ignored. However, the risks of these abnormal patterns are large.

Many studies have attempted to define the meaning of anomalies, such as epileptic signals in EEG, financial fraud in financial time series, and energy theft in meter datasets. However, it is difficult to understand abnormal events in most datasets. For example, there are many abnormal discharges in EEG signals that harbor any disease. If we focus only on epilepsy detection and ignore other abnormal patterns, it will pose a large risk to patients. Therefore, the classification of these anomalies is challenging. Existing classification methods are based on supervised learning models, but we cannot label all anomalies in the EEG signals. Therefore, a general method is required to understand the meaning of a variety of abnormal patterns.

The third issue is to explain the anomalies. Understanding the reason for an anomaly can help people analyze it and avoid risk, and explaining the anomalies is an important issue. In the multivariate time series, there was a high correlation between variables. A small change in one of the time series can lead to a large change in the other time series in the future time interval. Therefore, this small change can be considered as a reason for this anomaly. In addition, if a small change can be discovered, an anomaly can be detected before it occurs. In this case, we can avoid risks before an abnormal event occurs.

Recently, attribution has been widely studied in the field of climate-change detection. Researchers generally detect climate change from multivariate meteorological time series and attempt to determine the reasons for climate change. Ribes et al. [80] applied a statistical model to the global mean temperature and found that global warming is primarily attributable to human factors and that there is a limited contribution from natural factors.

There are some challenges in the explainability of anomalies. The first is how to discover patterns that are related to anomalies. Because these patterns can be small changes in the time series, discovering the relationship between them and anomalies is a research challenge. There are many different abnormal patterns. Therefore, the development of a general model for explaining anomalies is also a challenge. In addition, the anomalies in the multivariate time series were affected by a variety of factors. The scale of the time series affects the explainability of anomalies. If the scale of the multivariate time series is large, a large amount of information is included and the accuracy of explaining the anomalies is high. However, anomaly detection in large-scale multivariate time series is time-consuming. Therefore, developing a model for anomaly detection in large-scale multivariate time series is a research challenge.

9. Conclusions

In this study, we defined three types of anomalies found in multivariate time series, namely abnormal time points, time intervals, and time series. Furthermore, we reviewed the corresponding technologies for detecting each type of anomalies in multivariate time series, as well as applications of anomaly detection in various fields and several open access datasets. Finally, we discussed open issues and challenges still faced by the research in anomaly detection in multivariate time series.

Abnormal time points commonly represent noise or events at a certain timestamp. To detect this type of anomalies, the basic idea is to apply a regression model that predicts future values of the time series and then compute the loss between the predicted and true values, identifying an abnormal time point when a significant loss is obtained.

Abnormal time intervals in multivariate time series commonly indicate an event that occurred during a certain period. The main approach to detect an abnormal time interval is using an LSTM-AE model to reconstruct each time interval and determine the loss between the input and reconstructed intervals. If the loss is greater than a certain threshold, the corresponding time interval is detected as an anomaly. However, this approach fails to consider the relationship patterns among the multivariate time series. Therefore, approaches based on dynamic graphs have been proposed. The idea behind these methods is to use a dynamic graph to represent the relationship between multivariate time series at each time interval and then apply anomaly detection methods to determine whether the graph corresponds to an abnormal time interval.

The most commonly used method for abnormal time series detection is dimensionality reduction. Because an abnormal time series exhibits significantly different patterns when compared with other time series in the dataset, the abnormal time series lies far from the others in a low-dimensional feature space. Therefore, the PCA has become a popular method for detecting abnormal time series.

Abnormal time interval detection is more critical than time point and time series detection because abnormal time intervals typically represent an event. By detecting abnormal time intervals, we can create a wide range of applications in various fields. Therefore, we also summarize some applications of anomaly detection on multivariate time series in different fields, such as medicine and the Internet of Things.

In this study, we provide several open datasets for anomaly detection. These datasets contain the industrial IoT, medical, meteorological, and smart city areas. Each dataset contained a different number of time series, and the anomalies were labeled in most of the given datasets. For example, all the given EEG datasets label the seizure interval. The proposed anomaly detection approaches can be evaluated using labeled data.

Three open issues are presented in this survey for anomaly detection in a multivariate time series. First, we propose that detecting anomalies at an early stage is a research challenge in multivariate time series. The difficulty is to discover patterns that are helpful for detecting anomalies at an early stage. Second, understanding the meaning of the anomaly in a time series also plays an important role. In such cases, we can understand the types of abnormal events that occur at a certain time point. The challenge is to develop a model for understanding these abnormal events. Finally, explaining the anomalies is important in the field of anomaly detection and is used to explain the reason for the anomaly. The challenge is to discover the relationship between the abnormal time point and the time point that leads to an anomaly.

CRediT authorship contribution statement

Gen Li: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Jason J. Jung:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

This study was supported by a grant from the National Research Foundation of Korea (NRF) funded by the Korean government (MSIP) (NRF-2020R1A2B5B01002207). All authors approved the version of the manuscript to be published.

References

[1] S. Dwivedi, M. Vardhan, S. Tripathi, Building an efficient intrusion detection system using grasshopper optimization algorithm for anomaly detection, Cluster Comput. (2021) 1–20.
[2] I. Siniosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Fouliras, P. Sarigiannidis, A unified deep learning anomaly detection and classification approach for smart grid environments, IEEE Trans. Netw. Serv. Manag. 18 (2) (2021) 1137–1151.
[3] K. Choi, J. Yi, C. Park, S. Yoon, Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines, IEEE Access 9 (2021) 120043–120065.
[4] K.G. Ranjan, D.S. Tripathy, B.R. Prusty, D. Jena, An improved sliding window prediction-based outlier detection and correction for volatile time-series, Int. J. Numer. Modelling, Electron. Netw. Devices Fields 34 (1) (2021) e2816.
[5] S. Wang, B. Li, G. Li, B. Yao, J. Wu, Short-term wind power prediction based on multidimensional data cleaning and feature reconfiguration, Appl. Energy 292 (2021) 116851.
[6] G. Li, J.J. Jung, Seizure detection from multi-channel EEG using entropy-based dynamic graph embedding, Artif. Intell. Med. 122 (2021) 102201.
[7] Z. Chen, D. Chen, X. Zhang, Z. Yuan, X. Cheng, Learning graph structures with transformer for multivariate time-series anomaly detection in IoT, IEEE Internet Things J. 9 (12) (2022) 9179–9189.
[8] F. Passerini, A.M. Tonello, Smart grid monitoring using power line modems: Anomaly detection and localization, IEEE Trans. Smart Grid 10 (6) (2019) 6178–6186.
[9] G. Li, J.J. Jung, Dynamic relationship identification for abnormality detection on financial time series, Pattern Recognit. Lett. 145 (2021) 194–199.
[10] H. Nguyen, K.P. Tran, S. Thomassey, M. Hamad, Forecasting and anomaly detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management, Int. J. Inf. Manage. 57 (2021) 102282.
[11] L. Girish, S.K. Rao, Anomaly detection in cloud environment using artificial intelligence techniques, Computing (2021) 1–14.
[12] W. Ullah, A. Ullah, I.U. Haq, K. Muhammad, M. Sajjad, S.W. Baik, CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks, Multimedia Tools Appl. 80 (11) (2021) 16979–16995.
[13] H. Qin, M. Yan, H. Ji, Application of controller area network (CAN) bus anomaly detection based on time series prediction, Veh. Commun. 27 (2021) 100291.
[14] N. Reunanen, T. Räty, J.J. Jokinen, T. Hoyt, D. Culler, Unsupervised online detection and prediction of outliers in streams of sensor data, Int. J. Data Sci. Anal. (2019) 1–30.
[15] T. Kieu, B. Yang, C. Guo, C.S. Jensen, Outlier detection for time series with recurrent autoencoder ensembles, in: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19, International Joint Conferences on Artificial Intelligence Organization, Vienna, Austria, 2019, pp. 2725–2732.
[16] C. Yin, S. Zhang, J. Wang, N.N. Xiong, Anomaly detection based on convolutional recurrent autoencoder for IoT time series, IEEE Trans. Syst. Man Cybern. 52 (1) (2022) 112–122.
[17] C.Y. Priyanto, Hendry, H.D. Purnomo, Combination of isolation forest and LSTM autoencoder for anomaly detection, in: 2021 2nd International Conference on Innovative and Creative Information Technology (ICITech), Salatiga, Indonesia, 2021, pp. 35–38.
[18] J.F.A. Yeung, Z.-k. Wei, K.Y. Chan, H.Y. Lau, K.-F.C. Yiu, Jump detection in financial time series using machine learning algorithms, Soft Comput. 24 (3) (2020) 1789–1801.
[19] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun, D. Pei, Robust anomaly detection for multivariate time series through stochastic recurrent neural network, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 2828–2837.
[20] S. Maleki, S. Maleki, N.R. Jennings, Unsupervised anomaly detection with LSTM autoencoders using statistical data-filtering, Appl. Soft Comput. 108 (2021) 107443.
[21] H. Chen, H. Liu, X. Chu, Q. Liu, D. Xue, Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network, Renew. Energy 172 (2021) 829–840.
[22] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, C. Wu, Arrhythmia classification of LSTM autoencoder based on time series anomaly detection, Biomed. Signal Process. Control 71 (2022) 103228.
[23] H. Homayouni, I. Ray, S. Ghosh, S. Gondalia, M.G. Kahn, Anomaly detection in COVID-19 time-series data, SN Comput. Sci. 2 (4) (2021) 1–17.
[24] M. Munir, S.A. Siddiqui, A. Dengel, S. Ahmed, Deepant: A deep learning approach for unsupervised anomaly detection in time series, Ieee Access 7 (2018) 1991–2005.
[25] N. Ding, H. Ma, H. Gao, Y. Ma, G. Tan, Real-time anomaly detection based on long short-term memory and Gaussian mixture model, Comput. Electr. Eng. 79 (2019) 106458.
[26] T.-Y. Kim, S.-B. Cho, Web traffic anomaly detection using C-LSTM neural networks, Expert Syst. Appl. 106 (2018) 66–76.
[27] O.I. Provotar, Y.M. Linder, M.M. Veres, Unsupervised anomaly detection in time series using LSTM-based autoencoders, in: 2019 IEEE International Conference on Advanced Trends in Information Theory, ATIT, Kyiv, Ukraine, 2019, pp. 513–517.
[28] F. Liu, X. Zhou, J. Cao, Z. Wang, T. Wang, H. Wang, Y. Zhang, Anomaly detection in quasi-periodic time series based on automatic data segmentation and attentional LSTM-CNN, IEEE Trans. Knowl. Data Eng. 34 (6) (2022) 2626–2640.
[29] D. Wu, Z. Jiang, X. Xie, X. Wei, W. Yu, R. Li, LSTM learning with Bayesian and Gaussian processing for anomaly detection in industrial IoT, IEEE Trans. Ind. Inf. 16 (8) (2019) 5244–5253.
[30] K. Vos, Z. Peng, C. Jenkins, M.R. Shahriar, P. Borghesani, W. Wang, Vibration-based anomaly detection using LSTM/SVM approaches, Mech. Syst. Signal Process. 169 (2022) 108752.
[31] C.-W. Chang, C.-Y. Chang, Y.-Y. Lin, A hybrid CNN and LSTM-based deep learning model for abnormal behavior detection, Multimedia Tools Appl. (2022) 1–19.
[32] S. Lin, R. Clark, R. Birke, S. Schönborn, N. Trigoni, S. Roberts, Anomaly detection for time series using VAE-LSTM hybrid model, in: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, Virtual, Barcelona, 2020, pp. 4322–4326.
[33] Z. Niu, K. Yu, X. Wu, LSTM-based VAE-GAN for time-series anomaly detection, Sensors 20 (13) (2020) 3738.
[34] A. Deng, B. Hooi, Graph neural network-based anomaly detection in multivariate time series, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35, (5) 2021, pp. 4027–4035.

[35] H. Zhao, Y. Wang, J. Duan, C. Huang, D. Cao, Y. Tong, B. Xu, J. Bai, J. Tong, Q. Zhang, Multivariate time-series anomaly detection via graph attention network, in: 2020 IEEE International Conference on Data Mining, ICDM, IEEE, 2020, pp. 841–850.
[36] G. Li, J.J. Jung, Dynamic graph embedding for outlier detection on multiple meteorological time series, Plos One 16 (2) (2021) e0247119.
[37] T. Höfer, H. Przyrembel, S. Verleger, New evidence for the theory of the stork, Paediatr. Perinat. Epidemiol. 18 (1) (2004) 88–92.
[38] G. Li, J.J. Jung, Entropy-based dynamic graph embedding for anomaly detection on multiple climate time series, Sci. Rep. 11 (1) (2021) 1–10.
[39] G. Li, T.-H. Nguyen, J.J. Jung, Traffic incident detection based on dynamic graph embedding in vehicular edge computing, Appl. Sci. 11 (13) (2021) 5861.
[40] L. Beggel, B.X. Kausler, M. Schiegg, M. Pfeiffer, B. Bischl, Time series anomaly detection based on shapelet learning, Comput. Statist. 34 (3) (2019) 945–976.
[41] M.-S. Cheong, M.-C. Wu, S.-H. Huang, Interpretable stock anomaly detection based on spatio-temporal relation networks with genetic algorithm, IEEE Access 9 (2021) 68302–68319.
[42] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, S.-K. Ng, MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks, in: I.V. Tetko, V. Kůrková, P. Karpov, F. Theis (Eds.), Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series, Springer International Publishing, Cham, 2019, pp. 703–716.
[43] J. Audibert, P. Michiardi, F. Guyard, S. Marti, M.A. Zuluaga, USAD: unsupervised anomaly detection on multivariate time series, in: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '20, Association for Computing Machinery, New York, NY, USA, 2020, pp. 3395–3404.
[44] G. Li, X. Zhou, J. Sun, X. Yu, Y. Han, L. Jin, W. Li, T. Wang, S. Li, OpenGauss: An autonomous database system, Proc. VLDB Endow. 14 (12) (2021) 3028–3042.
[45] K. Hundman, V. Constantinou, C. Laporte, I. Colwell, T. Soderstrom, Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 387–395.
[46] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun, D. Pei, Robust anomaly detection for multivariate time series through stochastic recurrent neural network, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 2828–2837.
[47] J. Li, S. Di, Y. Shen, L. Chen, Fluxev: A fast and effective unsupervised framework for time-series anomaly detection, in: Proceedings of the 14th ACM International Conference on Web Search and Data Mining, WSDM '21, Association for Computing Machinery, New York, NY, USA, 2021, pp. 824–832.
[48] S.S. Lee, P.A. Mykland, Jumps in financial markets: A new nonparametric test and jump dynamics, Rev. Financ. Stud. 21 (6) (2008) 2535–2563.
[49] T.T. Huong, T.P. Bac, D.M. Long, T.D. Luong, N.M. Dan, B.D. Thang, K.P. Tran, et al., Detecting cyberattacks using anomaly detection in industrial control systems: A federated learning approach, Comput. Ind. 132 (2021) 103509.
[50] D.K. Sharma, T. Dhankhar, G. Agrawal, S.K. Singh, D. Gupta, J. Nebhen, I. Razzak, Anomaly detection framework to prevent DDoS attack in fog empowered IoT networks, Ad Hoc Netw. 121 (2021) 102603.
[51] P. Zhan, S. Wang, J. Wang, L. Qu, K. Wang, Y. Hu, X. Li, Temporal anomaly detection on IIoT-enabled manufacturing, J. Intell. Manuf. 32 (6) (2021) 1669–1678.
[52] H.A. Dau, A. Bagnall, K. Kamgar, C.-C.M. Yeh, Y. Zhu, S. Gharghabi, C.A. Ratanamahatana, E. Keogh, The UCR time series archive, IEEE/CAA J. Autom. Sin. 6 (6) (2019) 1293–1305.
[53] Z. Xiao, J. Jiao, Explainable fraud detection for few labeled time series data, Secur. Commun. Netw. 2021 (2021) 9941464.
[54] I. Benchaji, S. Douzi, B. El Ouahidi, Credit card fraud detection model based on LSTM recurrent neural networks, J. Adv. Inf. Technol. 12 (2) 113–118.
[55] E. D'Andrea, F. Marcelloni, Detection of traffic congestion and incidents from GPS trace analysis, Expert Syst. Appl. 73 (2017) 43–56.
[56] D. Wijayasekara, O. Linda, M. Manic, C. Rieger, Mining building energy management system data using fuzzy anomaly detection and linguistic descriptions, IEEE Trans. Ind. Inf. 10 (3) (2014) 1829–1840.
[57] J.-S. Chou, A.S. Telaga, Real-time detection of anomalous power consumption, Renew. Sustain. Energy Rev. 33 (2014) 400–411.
[58] D.B. Araya, K. Grolinger, H.F. ElYamany, M.A.M. Capretz, G. Bitsuamlak, Collective contextual anomaly detection framework for smart buildings, in: 2016 International Joint Conference on Neural Networks, IJCNN, Vancouver, Canada, 2016, pp. 511–518.
[59] S. You, B.H. Cho, Y.-M. Shon, D.-W. Seo, I.Y. Kim, Semi-supervised automatic seizure detection using personalized anomaly detecting variational autoencoder with behind-the-ear EEG, Comput. Methods Programs Biomed. 213 (2022) 106542.
[60] V. Rajinikanth, S. Kadry, D. Taniar, K. Kamalanand, M.A. Elaziz, K.P. Thanaraj, Detecting epilepsy in EEG signals using synchro-extracting-transform (SET) supported classification technique, J. Ambient Intell. Humaniz. Comput. (2022) 1–19.
[61] G. Varone, W. Boulila, M. Lo Giudice, B. Benjdira, N. Mammone, C. Ieracitano, K. Dashtipour, S. Neri, S. Gasparini, F.C. Morabito, et al., A machine learning approach involving functional connectivity features to classify rest-EEG psychogenic non-epileptic seizures from healthy controls, Sensors 22 (1) (2022) 129.
[62] M.L. Martini, A.A. Valliani, C. Sun, A.B. Costa, S. Zhao, F. Panov, S. Ghatan, K. Rajan, E.K. Oermann, Deep anomaly detection of seizures with paired stereoelectroencephalography and video recordings, Sci. Rep. 11 (1) (2021) 1–11.
[63] S. Chaabene, B. Bouaziz, A. Boudaya, A. Hökelmann, A. Ammar, L. Chaari, Convolutional neural network for drowsiness detection using EEG signals, Sensors 21 (5) (2021) 1734.
[64] L. Erhan, M. Ndubuaku, M. Di Mauro, W. Song, M. Chen, G. Fortino, O. Bagdasar, A. Liotta, Smart anomaly detection in sensor systems: A multi-perspective review, Inf. Fusion 67 (2021) 64–79.
[65] F. Forooghifar, A. Aminifar, T. Teijeiro, A. Aminifar, J. Jeppesen, S. Beniczky, D. Atienza, Self-aware anomaly-detection for epilepsy monitoring on low-power wearable electrocardiographic devices, in: 2021 IEEE 3rd International Conference on Artificial Intelligence Circuits and Systems, AICAS, IEEE, 2021, pp. 1–4.
[66] R. Jain, H. Shah, An anomaly detection in smart cities modeled as wireless sensor network, in: 2016 International Conference on Signal and Information Processing (IConSIP), Nanded, India, 2016, pp. 1–5.
[67] R.N. Madurawe, B.I. Jayaweera, T.D. Jayawickrama, I. Perera, R. Withanawasam, Collusion set detection within the stock market using graph clustering amp; anomaly detection, in: 2021 Moratuwa Engineering Research Conference (MERCon), Sri Lanka, 2021, pp. 450–455.
[68] S. Pan, T. Morris, U. Adhikari, Developing a hybrid intrusion detection system using data mining for power systems, IEEE Trans. Smart Grid 6 (6) (2015) 3104–3113.
[69] J.M. Beaver, R.C. Borges-Hink, M.A. Buckner, An evaluation of machine learning methods to detect malicious SCADA communications, in: 2013 12th International Conference on Machine Learning and Applications, Vol. 2, Washington, USA, 2013, pp. 54–59.
[70] T. Morris, W. Gao, Industrial control system traffic data sets for intrusion detection research, in: J. Butts, S. Shenoi (Eds.), Critical Infrastructure Protection VIII, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 65–78.
[71] T.H. Morris, Z. Thornton, I. Turnipseed, Industrial control system simulation and data logging for intrusion detection system research, in: 7th Annual Southeastern Cyber Security Summit, 2015, pp. 3–4.
[72] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S.T. Treves, J. Guttag, Patient-specific seizure onset detection, Epilepsy Behav. 5 (4) (2004) 483–498.
[73] M.K. Siddiqui, R. Morales-Menendez, X. Huang, N. Hussain, A review of epileptic seizure detection using machine learning classifiers, Brain Inform. 7 (1) (2020) 1–18.
[74] M.A. Kramer, E.D. Kolaczyk, H.E. Kirsch, Emergent network topology at seizure onset in humans, Epilepsy Res. 79 (2–3) (2008) 173–186.
[75] A. Schulze-Bonhage, F. Sales, K. Wagner, R. Teotonio, A. Carius, A. Schelle, M. Ihle, Views of patients with epilepsy on seizure prediction devices, Epilepsy Behav. 18 (4) (2010) 388–396.
[76] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907.
[77] K. Yoshihara, K. Takahashi, A simple method for unsupervised anomaly detection: An application to web time series data, PLoS One 17 (1) (2022) e0262463.
[78] X. Ma, H. Yu, Y. Wang, Y. Wang, Large-scale transportation network congestion evolution prediction using deep learning theory, PLoS One 10 (3) (2015) e0119044.
[79] T.S. Buda, H. Assem, L. Xu, ADE: An ensemble approach for early anomaly detection, in: 2017 IFIP/IEEE Symposium on Integrated Network and Service Management, IM, Lisbon, Portugal, 2017, pp. 442–448.
[80] A. Ribes, F.W. Zwiers, J.-M. Azais, P. Naveau, A new statistical approach to climate change detection and attribution, Clim. Dynam. 48 (1) (2017) 367–386.