ISSN 2278-3091

Volume 10, No.3, May - June 2021

International Journal of Advanced Trends in Computer Science and Engineering

Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse571032021.pdf

https://doi.org/10.30534/ijatcse/2021/571032021

A Review on Anomaly Detection in Time Series

Syed Hassan Ali Shah [1]*, Muhammad Junaid Akbar [2], Usman Ahmed Raza [3]

[1] Department of Computer Science, Lahore Leads University, Lahore, Pakistan, hasssan7864@gmail.com

[2] Department of Computer Science, Lahore Leads University, Lahore, Pakistan, junaidakbar527@gmail.com

[3] Department of Computer Science, Lahore Leads University, Lahore, Pakistan, usmanahmedraza@gmail.com

ABSTRACT

Time series is a very common class of data sets. Among others, it is very simple to obtain time series data from a variety of various science and finance applications and an anomaly detection technique for time series is becoming a very prominent research topic nowadays. Anomaly identification covers intrusion detection, detection of theft, mistake detection, machine health monitoring, network sensor event detection or habitat disturbance. It is also used for removing suspicious data from the data set before production. This review aims to provide a detailed and organized overview of the Anomaly detection investigation. In this article we will first define what an anomaly in time series is, and then describe quickly some of the methods suggested in the past two or three years for detection of anomaly in time series.

Key words : Anomaly Detection, Time Series, Outlier, Review

1. INTRODUCTION

The growth of data has increased with growing use of information technology. The recent study, conducted in at University of Southern California, has shown that after 1980s, the scale of data has risen dramatically and doubled per year or even a few months. Both the global data and the regional data reached 295EB until 2007, and then they hit about 1.8 ZB until 2011, and more than 40 ZB by 2020. Big data time has come to university these days. In addition, it offers solutions to figure out how to summarize potential knowledge from the rich data [1]. A major part of these details is the time series. The so-called time series is a sequence of chronologically sequential occurrences. With the use of time series, scientists can accumulate and interpret a vast volume of data and understand more about these fields than they did with simplistic analysis with less numbers.

Although there are relatively few situations in which the standard abnormality seems to be off for time series results, this does not inherently indicate that the abnormalities are not important. Under the tiny dataset behind the unusual values, the actual piece of data which be hidden. In the field of medicine an abnormal heart beat is identifiable enough that the ECG observation is able to reveal the disorder in time. In addition, time-series anomaly detection software may be used to map engine state, network violation detection, money-laundering, tracking of public opinion networks, credit card theft, stock market analysis, unlawful enforcement of tax acts, major inspections of construction sites, possibly any IT device, and several other IT scenarios. And though it could be fascinating to research this phenomena from a scientific viewpoint, it may make your life simpler.

2. ANOMALY DETECTION DEFINITION OF TIME SERIES

Obtained Time Series are commonly available in many major databases involving finance, medicine, engineering, and social science. It is important to understand that there are two different features of time-series data: time-series and other data types. First, we need to group the records by time and/or day of the year, and before including the records we must compare the values of the attributes effectively and in a logical order. Secondly, in certain places and times, the sequence attribute, the record values, are a long flowing sequence. To determine whether time series are univariate or multivariate, they can be broken up into univariate time series and univariate time series.

Anomaly detection challenge for time series is simply achieved by finding outlier signal relative to some regular signal, like unpredicted spikes, drops, changes in trends and changes in level.

3. METHODS FOR ANOMALY DETECTION

Several foreign researchers have penetrated into the analysis of time series outlier detection after Barnett wrote the first book Outliers of observational data on anomaly detection in the 1980s, such as M Breunig, E M Knorr, E Keogh, Portnoy, J Takeuchi, M Agyemang, M Markou, V Chandola and so on. Domestic research begins very late, but advances rapidly. Related research are being carried out by Tsinghua University, Xi'an Jiao Tong University, Tianjin University, Fudan University, Hong Kong University of Science and Technology, etc. Because of the scientific relevance and deployment prospect of time series outlier identification, a large number of scientists have joined in its study. Many high-quality UTS achievements have been reached and published in journals such as IEEE TKDE Neural Computing Numerical Statistics and Data Processing at a renowned international conference such as PAKDD PKDD SIGKDD VLDB over the past 10 years.

Anomaly detection is attracting even more recognition and analysis as an essential sub-branch of data mining. Most approaches of anomaly detection have been suggested by domestic and international researchers, which can be classified into five categories: abnormal statistical-based detection, abnormal clustering-based detection, abnormal distance-based detection, abnormal density-based detection, etc. [2].

3.1. Deep Neural Network for Unsupervised Anomaly Detection

Multi-turn data are currently gradually gathered from different real-time applications, such as power plants, wearable devices, etc. The multivariate pathological identifying and assessing sequence identifies and separates at some stages the root causes of sporadic illness. But it is important to design this mechanism, not just to document time dependence and time series, but also to encrypt interrelationships between different time series pairs. The device may also be noise tolerant and give operators numerous anomalies depending on many collisions. Although many unexpected anomaly detection algorithms have been developed, few can jointly solve these difficulties. We give a multi-variable time series for the identification of deviations, a CRD in the article (MSCRED). The multi-scaled (resolution) signature matrices MscrED originally produces to characterize system status levels in various timescales. The signature measurements then encode the associations between the (time series) sensor and the attention-based CTM networks for transient model capture by using a convolutionary encoder (ConvLSTM). Finally, an input signature matrix recreates the convolutionary decoder which detect and diagnostic the remaining irregularities by using inter-sensor similitudes and the tempo information maps. Extensive observation - based on virtual dataset and an existing plant data set experiment shows that MSCRED can use additional simple methods.

For the issue of detecting and diagnosing abnormalities, they proposed it in that article, and created a groundbreaking approach, MSCRED, which integrates model reasoning. Multi-scale (resolution) device signature matrices are used to characterize the state of the entire system at various time segments, and a deep encoder-decoder framework is used to produce reconstructed signature matrices. The system is able to model both inter-sensor associations and temporal dependencies in multivariate time series. After the residual signature matrices have been extracted, they are further used to identify and diagnose any abnormalities. In a comprehensive series of observational tests, which compared the output of MSCRED on synthetic data and a power plant dataset, it was noticed that MSCRED outperforms the industry norms by a large margin [3].

3.2. Unsupervised Anomaly Detection Using LSTM-Based Auto encoders

A method for categorizing and detailing anomalies in data sets has been identified as an anomaly. Correct identification of anomalies today is crucial because pure data volumes prevent the hand-marking of outliers. Auto detection system operations include theft detection, physician supervision, error detection and incident detection. In this topic the main issue is that there are no anomalies. Therefore, conventional methods of machine learning cannot be used for model training because time series labels are impossible.

Many other classical anomaly detection systems exist, for example

- Anomaly Detection Based on Clustering
- Isolation Forests
- Support Vector Machine
- The application of Gaussian distribution for anomaly detection

Moreover, all these methods describe the outlier merely because of its magnitude, but not because of the values of previous stages. In most other respects, when using such methods, temporal data is not taken into account. Consequently, classical systems had little success. Irregularities in time series data can be mentioned among the observed algorithms.

- Outlier detection based on signal decomposition (classical decomposition, STL) [4]
- Space vector model, Holt Winters, ARIMA Exponentially smoothing
- Deep learning: auto encoders based on feed forward, recurrent and LSTM neural network layers [5]
- Dimensionality reduction: RPCA, SOM, discords, piecewise linear

The easiest way to achieve this is to break down during normal time series: seasonal components and designs can first be excluded from the signal, along with an outside traditional tracking system. For example, this approach works well on hotel price data where constant fluctuations occur per year. Prices are now that year-on-year due to inflation.

So, we should look at the data point after being removed from the seasonal and the pattern vectors if it is far from empty. However, where time series are not accurate (e.g., foreign trade or sound), machine learning methods can only be employed. The self-coding discovery of anomalies is one of the better machine learning approaches. An automobile encoder is an artificial nerve network used to encrypt data effectively and without monitoring. An encoder's objective is to learn representation by training a network to disregard the signal 'noise' to reduce dimensionality for a wide range of data.

A rebuild and a reduction side are developed where the auto encoder plans to produce the same image as the original input and term of the reduced encoding. A single encoder and a decoder layer also have automated encoders, while profound encoders and decoders are profit-making. The encoder and decoder are used as two units. The features behind a stage are identified by an encoder. These features are typically smaller. The decoder reconstructs the original data from these.

Feed forward's neural network can be used to build auto encoder. We will therefore construct an LSTM-centric auto-encoder to accept temporal details. In comparison to a neural feed forward network, we use information to refer to LSTM one at a time. Each RNN unit is an extension of the RNN to preserve awareness of its importance in the neural network in a time sequence.

It is best to pick a neural network's design and post-processing variables based on data to be fed into the device. The most relevant things to remember are:

- Neural network's range of layers
- Layer LSTM cell count
- Window scale for neural network feeding
- Window scale smoothing
- Threshold, where we regard as anomaly the height of residues

Other approaches may be helpful, but only the auto encoder technique is universal and efficient enough for all sorts of time series [6].

3.3. Unsupervised Anomaly Detection with LSTM Neural Networks

Examine the detection of anomalies inside an unattended framework and integrate long-term network neural memory (LSTM). These sequences are transferred in particular via our LSTM frame and achieve defined period in the specified sequences of variable data lengths. Then, you will notice our anomaly detector's decision-making feature based on OC SVM and the single-class help-vector-definition (SVDD) algorithms. Because our first solution to collaborative preparation and optimization of LSTM and OC-SVM algorithms is to use extremely effective gradients and quadratic programming. In order to incorporate gradient training methods, we modify the original objective criteria of OC-SVM and SVDD algorithms if the current objective criteria converge with the original criteria. Our unattended formulation is often applied to semi-controlled and professionally supervised processes. It helps us achieve algorithms for anomaly detection that can process and deliver high efficiency data sequences, particularly in time series data. Our approach is generic enough that our LSTM and GRU-based architecture can be directly substituted with the Gated Recurrent Units approach (GRU). In our research, our conventional algorithms show substantial increases in performance.

Anomaly detection is researched and LSTM algorithms are presented in a non-supervised setting. Especially for the processing of variable-long data sequences, we implemented a general LSTM framework. Following the acquirement of defined sequences via our LSTM-based architecture, we add a ranking feature of our OC-SVM [6] and SVDD [7] algorithms for anomaly detectors. The parameters of both LSTM architectures and the final scoring function for the OC-SVM (or SVDD) formulation are optimized as a first time in literature. We have also conducted regression and Quadratic Programming-based training sessions with various algorithmic values to refine the parameters for our algorithms together, so that our derivatives for these algorithms can be applied to the half-checked and totally regulated frameworks. We change the OC-SVM and SVDD formulations in order to implement the gradient-based training mechanism and then include the convergence effects of the revised formulations with the original formulations. Therefore, we get highly efficient anomaly detection algorithms, specifically for time series data that can process data sequences of varying lengths. We also have GRU-based anomaly detection algorithms in our simulations owing to the generic structure of our method. We demonstrate major performance improvements obtained with the traditional methods through our algorithms [7], [8] and [9] through a broad variety of actual and virtual data sets via comprehensive experiments.

3.4. Time-Series Anomaly Detection Service at Microsoft

Large businesses must monitor their software and facilities via various indicators in real time (e.g., website views and revenues). With Microsoft time series, we provide an anomaly detector service that allows customers to monitor time series on a permanent basis and alert against events in time. We present in this text the pipeline and algorithm of our Anormal Detection Service for the unique, efficient and general purposes. The pipeline comprises three main components, namely intake of data, analysis tools and online computing. This method aims to resolve the problem of identification of an anomaly in time series by constructing a new algorithm based on Spectral Resin (SR) and CNN (CNN). Our work was the first effort to identify abnormalities for time series by taking the SR model from the field of visual saliency detection. We also merge SR with CNN for innovative enhancement of the performance of SR models. Our approach provides superior experimental results, unlike the existing baselines on both the public and Microsoft data output platforms.

The aim of identification of anomalies is to identify unusual patterns or uncommon artifacts. Data mining has become and is a critical analytical area for the business application as one of the most popular sectors. The real time identity anomaly will save a company's resources by eliminating downtime, mitigating brand damage and maintaining the company's image unchanged. Criminal justice experts conclude that inaccurate, positive claims are primarily liable for regulation that, along with financial companies who offer their own AMSI services to manage the condition, commodities and the wellbeing of their industry, pose the greater burden of this problem. When anomalies are detected, administrators are advised that they should take action to cope with injuries as soon as possible. Yahoo's release of EGADS [10] is an excellent example, which aims to track and boost alerts for the millions of time series for Yahoo's numerous properties in multiple applications. We build a detection service in Microsoft which tracks millions of measures from Bing, Workplace and Azure, and helps engineers function faster on a web. In this post, we emphasize the pipelines and the algorithm of our time series anomaly detection method.

The system contains three core elements: data ingest, platform testing and online estimates. We will install the whole pipeline first before describing these elements. By ingesting time series on a device, users will report monitoring incidents. It facilitates the use of time series from different data sources (including azure storage, databases, and online streaming data). The ingestion manager shall vary with the granularities indicated at each stage, e.g., minute, hour or day. Series points are stored in a time series database on the streaming pipeline via Kafka. Online state input time series anomaly test anomaly detector processor. Consumers concurrently eat a number of time series in a typical situation of demand measurements. For example, for different markets and channels, the Bing squad used a time series. If an event occurs, alert systems combine time series anomalies, and provide email and payment services to customers. Cumulative abnormalities mean the average condition of an injury and allow users to reduce diagnostic problems.

Anomaly identification in time series is important for maintaining the consistency of online services. In real applications an inexpensive, robust and reliable anomaly detection process is useful. We also released an Anomaly Detecting Service at Microsoft in this article. More than 200 teams, among them Bing, Office and Azure, have used the service in Microsoft. Anomalies in the production are detected from a maximum of 4 million time series per minute. Moreover, for the first time in time series anomaly detection we apply the Spectral Residual (SR) model and innovatively merge the SR model with the CNN model to deliver excellent performance. In future, we expect to combine together in order to provide our customers with a better anomaly detection service. Besides internal service, as part of our Cognitive Service, our time series anomaly detection system will soon be accessible to external customers through Microsoft Azure [11].

3.5. Deep Learning Approach for Unsupervised Outlier Detection in Time Series

In standard anomaly detection techniques, current points and seasonal fluctuations normally found in streaming data cannot be observed on the basis of distances and density, thereby allowing temporal anomalies to be identified in the existing IoT cycle. We use a new approach to time series data to solve this problem, called the Deep Learning Detection approach, which is applicable for non-streaming scenarios (DeepAnT). DeepAnT is able to detect a vast variety of anomalies, including dots, background and time series discord. In comparison to methods of anomaly identification under which anomalies are detected, DeepAnT uses uncontrolled data to gather and understand the dissemination of information used for deterring natural behavior. DeepAnT has two components: the time series predictor and abnormality detector. The time series Modulator predicts when the defined horizon is next marked by the profound neural networks (CNN). This module takes a window with a time series and measures the next time line (used as a background). The forecast value is then transferred to the detector module, which marks the time stamp continuously or irregularly. Even without removing exceptions from the data collection, DeepAnT can be trained. In general, for the deep learning methods of a model, several data are required. DeepAnT can acquire very small data sets, but the close exchange of CNN parameters guarantees a strong generalization power. As DeepAnT does not recognize the anomalies, it does not depend on unusual model generation marks. This technique can also be used specifically in real-life situations, where a large amount of data from heterogeneous sensors in natural and anomalous areas may be hard to distinguish. In addition to 10 anomaly detection bench marks, we have conducted a detailed evaluation of 15 algorithms, consisting of 433 actual and synthetic time series. Experiments show that DeepAnT is comparable with other anomaly detection techniques in most situations.

The proposed DeepAnT comprises two modules. The first module is the Time Series Index. The second module is liable for the normal or abnormal marking of data points in a given amount of occasions. Time-scale is the second module. Deep learning was used primarily for a vast range of applications because of its potential to automatically discover complex features without domain knowledge. This artificial neural network learning ability enables a powerful anomaly detection nominee in time series. Therefore DeepAnT utilizes raw figures, which includes CNN. It is often strong to change, in contrast to other neural networks and mathematical models. Literature [12][13] is considered to function well with the ability of LSTM to derive long-term trends from time-series. However, we have seen that CNN can be an excellent alternative to standardized and multi-varying time series results because of its parameter performance. CNN and LSTM are generally used in literature classification issues of time series [14] and [15], but we use CNN for time series regression (and LSTM for comparison).

In cases where a significant number of data is accessible without the risk of naming this method will practically be implemented. The data modeling process may therefore be hampered by low data quality. In the other hand, if the amount of pollution is above 5%, the device will attempt to model the instances, then, as they are assumed to be natural at the time of delimitation. The network design is chosen and the resulting hyper parameters are another constraint. The modern architecture quest techniques [16] are the usage of human technological expertise for this function to be circumvented. One of the most serious constraints is perhaps the adverse examples [17] which restrict the use of this technique in protection scenarios (and the majority of previous data-driven methods). In knowing and defending against these adversary examples important measures have been made. But no generic strategy to overcome this problem has yet been created [18].

4. CONCLUSION

At the conclusion of this paper, a few different hypotheses on time series anomaly detection are offered. The anomaly is just making a difference proportional to context. If a distinction is drawn between natural and abnormal behavior, it is pointless to regard one as another. The definition and meaning of anomaly vary, depending on where one is in the application. Because of this, say, the height of a 568 cm isn't deemed out of the ordinary when it is applied to a person, but for a CEO earning vastly more than the rest of the workers, it is when a comparison is drawn over a group. Therefore, specialists on the Internet programming deal with the reasons would influence the truthfulness of the algorithm's results in determining whether or not the algorithm detects accurately. To inform users of just the questionable or odd data in order to grab their attention Here we have an example of an extra or different condition, then viewed in relation to the data, we may call it an anomaly detection. Since standard multivariate data processing and multidimensional expansion, the results of the MTS seem to be very different. The factors have several interrelationships; because of this, the results would likely have a multivariable structure. As there are many possible reasons for detecting the MTS variation, the MTS anomaly test is conducted by doing a systematic study of each component. Specifically, time series detection is still in MTS anomaly remains immature, especially in MTS anomaly detection of deviance. In addition, the existing anomaly detection algorithm is not streamlined, so more reduction in complexity is needed in order to cope with dynamic time series analysis is still needs to be investigated.

ACKNOWLEDGE

We thank Mr. Usman Ahmed Raza, our faculty member, who encouraged us and supported us to achieve this goal.

REFERENCES

[1] Bowen. Z. and Shahriar. S, "Finding needle in a million metrics: anomaly detection in a large-scale computational advertising platform", Proceeding of the 2nd International Workshop on Ad Targeting at Scale, San Francisco, CA, USA, pp.1062-1065, Feb, 2016

[2] Ozkan, Huseyin.O, Fatih. O, Suleyman. S, "Online anomaly detection under markov statistics with controllable Type-I error ", IEEE Transactions on Signal Processing, Vol.64, No.6, pp.1435-1445, March, 2016.

[3] Chuxu Zhang, ∗Dongjin Song, Yuncong Chen, Xinyang Feng, Cristian Lumezanu, Wei Cheng, Jingchao Ni, Bo Zong, Haifeng Chen, Nitesh V. Chawla. A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data, The Thirty-Third AAAI Conference on Artificial Intelligence, 2019.

[4]] Robert B. Cleveland, William S. Cleveland, Jean E. McRae, Irma Terpenning. STL: A seasonal-trend decomposition procedure based on loss. Journal of Official Statistics, Vol 6., No. 1, 1990, pp. 3-73. http://www.nniiem.ru/file/news/2016/stl-statistical-model.pdf

[5] A Tutorial on Deep Learning Part 2: Autoencoders, Convolutional Neural Networks and Recurrent Neural Networks http://robotics.stanford.edu/~quocle/tutorial2.pdf

[6] Oleksandr I. Provotar, Yaroslav M. Linder, Maksym M. Veres. Unsupervised Anomaly Detection in Time Series Using LSTM-Based Autoencoders IEEE 2019

[7]] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution,” Neural Comput., vol. 13, no. 7, pp. 1443–1471, 2001.

[8]] D. M. J. Tax and R. P. W. Duin, “Support vector data description,” Mach. Learn., vol. 54, no. 1, pp. 45–66, Jan. 2004. doi: 10.1023/ B:MACH.0000008084.60811.49.

[9] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, Long Short Term Memory Networks for Anomaly Detection in Time Series. Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain, 2015, p. 89

[10] Nikolay Laptev, Saeed Amizadeh, and Ian Flint. 2015. Generic and Scalable Framework for Automated Time-series Anomaly Detection. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, NY, USA, 1939–1947

[11] Hansheng Ren, Bixiong Xu, Yujing Wang, Chao Yi, Congrui Huang, Xiaoyu Kou∗ Tony Xing, Mao Yang, Jie Tong, Qi Zhang. Time-Series Anomaly Detection Service at Microsoft. arXiv:1906.03821v1 [cs.LG] 10 Jun 2019

[12] F. A. Gers, D. Eck, and J. Schmidhuber, ‘‘Applying LSTM to time series predictable through time-window approaches,’’ in Neural Nets WIRN Vietri-01. London, U.K.: Springer, 2002, pp. 193–200

[13] S. Hochreiter and J. Schmidhuber, ‘‘Long short-term memory,’’ Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997

[14] Z. C. Lipton, D. C. Kale, C. Elkan, and R. Wetzell. (2015). ‘‘Learning to diagnose with LSTM recurrent neural networks.’’ [Online]. Available: https://arxiv.org/abs/1511.03677

[15] Y. Zheng, Q. Liu, E. Chen, Y. Ge, and J. L. Zhao, ‘‘Time series classification using multi-channels deep convolutional neural networks,’’ in Proc. Int. Conf. Web-Age Inf. Manage. Cham, Switzerland: Springer, 2014, pp. 298–310.

[16] B. Zoph and Q. V. Le, ‘‘Neural architecture search with reinforcement learning,’’ CoRR, vol. abs/1611.01578, pp. 1–16, Nov. 2016.

[17] A. Kurakin, I. J. Goodfellow, and S. Bengio, ‘‘Adversarial machine learning at scale,’’ CoRR, vol. abs/1611.01236, pp. 1–17, Nov. 2016.

[18] Mohsin Munir, Shoaib Ahmed Siddiqui, Andreas Dengel, Sheraz Ahmed. DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. IEEE January 7, 2019.

[19]AzlizaYacob ,ZirawaniBaharum , NurSukinah Aziz, Noor SuhanaSulaiman and Wan Mohd Amir Fazamin Wan Hamzah, “A Review of Internet of Things (IoT): Implementations and Challenges”, International Journal of Advanced Trends in Computer Science and Engineering, 9(1.3), 2020, 373 - 376