KPI Anomaly Detection

Anomaly detection is the foundation of intelligent operation and maintenance (O&M), and detection objects are evaluated by key performance indicators (KPIs). For almost all computer O&M systems, KPIs are usually the machine-level operating data. Moreover, these high-frequency KPIs show a non-Gaussian distribution and are hard to model, i.e., they are intricate KPI profiles. However, existing anomaly detection techniques are incapable of adapting to intricate KPI profiles. In order to enhance the performance under intricate KPI profiles, a seasonal adaptive KPI anomaly detection algorithm *ASAD* (Adaptive Seasonality Anomaly Detection) was presented.

1. Introduction

Computer operation and maintenance is always a vital component in guaranteeing the high availability of the application systems. Operation and maintenance must evolve from manual detection to intelligent detection with the explosive increase in the volume of application data. According to Gartner's report, more than 40% of global enterprises have replaced their outdated O&M systems with intelligent solutions as of 2020. In these intelligent systems, anomaly detection is critical to detect important performance indicators (KPIs) such as CPU utilization, memory utilization and so on. To ensure a stable and reliable O&M system, a rising number of researchers are investigating KPI anomaly detection methods [1][2].

Traditional statistics, supervised learning and unsupervised learning algorithms are the three types of KPI anomaly detection techniques. First, seasonal length is required as an input parameter by traditional statistical approaches such as Argus [3] and TSD [4], but it is frequently given manually. It may cause seasonality to be disrupted in intricate KPI profiles, leading to erroneous anomaly detection. Secondly, supervised learning algorithms such as Opperence [5] and EGADS [6] relied on classical statistical techniques, and they also did not recognize seasonal length under intricate KPI profiles. Finally, among unsupervised learning methods, Zhao, N. [7] developed a periodic adjustable approach called Period. This entry considers time series data to be related to daily human activities, and it directly assumed that the basic seasonal length of time-series data is 1 day. However, KPI time series data containing intricate KPI profiles are very common, the seasonal length in these non-Gaussian distributed data is difficult to estimate [8]. In general, there are three key challenges to overcome. To begin, precise seasonal characteristics are hard to extract from the intricate KPI time series data. Second, due to the long sub-sequence length, the clustering process will take too much time. Third, noise and anomalies in the KPI time series data could also result in bad sub-sequence clustering results. Facing the above problems, existing KPI anomaly detection algorithms cannot obtain good performance under intricate KPI profiles.

2. Background

- KPI: Key Performance Indicator (KPI) consists of many background system metrics including CPU utilization, memory utilization, network throughput, system response time and so on. Above types of KPI time series data can cover the main information from hardware to software, and reflect the status of the entire system from the bottom up. In brief, it is the focus of the operation and maintenance system.

- Intricate KPI Profiles: In KPI time series data, time is the independent variable and KPI value is the dependent variable. The shape of the KPI time series data graph is known as the KPI profile. For most operation and maintenance systems, as time passes, the KPI profile will take on new forms, i.e., the graph of KPI time series data usually contains many KPI profiles. In researchers' work, the situation where many types of KPI profiles exist in KPI time series data graph is referred to as intricate KPI profiles.

- KPI Anomaly: KPI anomalies are data that do not meet expectations in KPI time series data [5][9]. Anomalies in KPIs are usually a sign that something is wrong with the system. For example, the system's CPU utilization remains excessively

high, indicating that the number of computing tasks executed by the system exceeds the typical level, posing a crash risk. Early detection of KPI deviations can aid in the diagnosis and analysis of issues.

- Seasonality of Time Series Data: When time series data vary with seasonal influences, they are said to have seasonality [10]. For example, if time series data frequently exhibit fixed characteristics in a certain time interval, this can indicate that the data are seasonal. The seasonal length is the time between repetitions, and it occurs at an observed or predicted period.

3. KPI Anomaly Detection Algorithm

Existing KPI anomaly detection algorithms shown in **Table 1** are divided into three categories, including traditional statistical, supervised learning and unsupervised learning algorithms [11]. For traditional statistical algorithms, Yaacob, A.H. et al. [12] studied the problem of network attack detection based on ARIMA in 2010. In 2012, Yan, H. et al. [3] developed the end-to-end service quality evaluation problem based on Holt–Winter. In 2013, Chen, Y. et al. [4] studied the view of web search response time based on TSD. The disadvantage of them is that they all need to input seasonal fitting parameters and cannot adapt to intricate KPI profiles.

Table 1. KPI Anomaly Detection Algorithms.

Name	Time	Type
Yaacob, A.H. et al. [12]	2010	traditional statistical
Yan, H. et al. [3]	2012	traditional statistical
Chen, Y. et al. [4]	2013	traditional statistical
Liu, D. et al. [5]	2015	supervised learning
Laptev, N. et al. [6]	2015	supervised learning
Zhou et al. [13]	2019	ensemble learning
Himeur et al. [14]	2020	deep neural network
Himeur et al. [15]	2021	deep neural network
Deng et al. [16]	2021	graph deviation network
Chen et al. [17]	2021	transformer-based architecture
Zhou et al. [18]	2021	federated learning
Xu, H. et al. [19]	2018	unsupervised VAE
Himeur et al. [20]	2021	unsupervised temporal autoencoder
Li et al. [21]	2021	unsupervised learning
Li et al. [22]	2021	fast unsupervised learning
Carmona et al. [23]	2021	unsupervised learning

For supervised learning algorithms, in 2015 Liu, D. et al. [5] proposed Opperence based on traditional statistical algorithms to solve the problems of service quality monitoring and performance anomaly detection. In the same year, Laptev N et al. [6] presented system anomaly monitoring based on traditional KPI anomaly detection methods. In 2019, Zhou et al. [13] designed an ensemble learning scheme based on extreme learning machine (ELM) algorithm and majority voting method to detect abnormal electricity consumption. In 2020, Himeur et al. [24] firstly discussed the anomaly detection in building energy consumption. It comprehensively introduced a method to classify existing algorithms based on different factors, such as the machine learning algorithm, feature extraction approach, detection level, computing platform, application scenario and privacy preservation. Then they introduced a new solution [14] to detect energy consumption anomalies. Besides micro-moment features extraction, they developed a deep neural network architecture for efficient abnormality detection and classification. In 2021, they also used the autoencoder and micro-moment analysis to detect abnormal energy usage [15]. To provide an explainable model, Deng et al. [16] propose a novel Graph Deviation Network (GDN) approach. It can learn a graph of relationships between sensors, and detects deviations from these patterns. Similarly, Chen et al. [17] presented a new framework for multivariate time series anomaly detection (GTA) that involves automatically learning a graph structure, graph convolution and modeling temporal dependency using a Transformer-

based architecture. Recently, Zhou et al. [18] put forward an anomaly detection framework. Firstly, this captures more detailed data regarding the time series' shape and morphology characteristics. Then, it utilizes interval representation to realize data visualization and mine the internal relationships. However, these supervised methods are unable to adapt to intricate KPI profiles due to the inherent lack of labeled anomalies in historical data.

For unsupervised learning algorithms, in 2018 Xu, H. et al. [19] studied application monitoring problems based on VAE. In 2021, Thill et al. [25] designed a novel unsupervised temporal autoencoder architecture based on convolutional neural networks (TCN-AE). It can utilize the information from different time scales in the anomaly detection process. Then, Himeur et al. [20] developed two different schemes to detect abnormalities in energy consumption. These are an unsupervised abnormality detection based on one-class support vector machine (UAD-OCSVM) and a supervised abnormality detection based on micro-moments (SAD-M2). In the same year, Li et al. [21] proposed a clustering-based approach to detect anomalies concerning the amplitude and the shape of multivariate time series. They generate a set of multivariate subsequences by setting the sliding window. To improve the detection efficiency, Li et al. [22] proposed FluxEV, a fast and effective unsupervised anomaly detection framework. It can extract appropriate features to indicate the degree of abnormality, and make the features of anomalies as extreme as possible. Recently, Carmona et al. [23] presented a framework Neural Contextual Anomaly Detection (NCAD) that scales seamlessly from the unsupervised to supervised setting. It is a window-based approach which can facilitate learning the boundary between normal and anomalous classes by injecting generic synthetic anomalies into the available data. Moreover, it adopted the moments method to speed up the parameter estimation in the automatic thresholding. Although they achieved good performance, the defect is also unsuitable for intricate KPI profiles due to lack of significant seasonality in original data. In order to solve the above problem, Zhao, N. et al. [7] devised a periodic adaptable algorithm Period, to enhance the accuracy of KPI anomaly detection. The authors of this work assumed that the intricate KPI profiles had a 1-day seasonal length and split the KPI data. However, this strategy is not universal, because not all intricate KPI profiles have a 1-day seasonal length.

In fact, KPIs may show distinct patterns in different time intervals, which are referred to as KPI profiles, such as weekly, quarterly or other imperfect or complex periodicity. To deal with the situation described above, a new algorithm to recognize intricate KPI profiles with uncertain seasonal lengths must be developed. Therefore, this entry proposes an adaptive seasonality anomaly detection algorithm under intricate KPI profiles. The notations list of researchers' research is shown in **Table 2**.

Table 2. Notations List.

Notation	Description
S	seasonal component of KPI time series data
S_i	ith sample seasonal component from S
sum_i	sum of sample points for ith sample sequence
PO_{all}	set of powers derived from the periodogram
po_i	ith power from PO_{all}
$period_i$	ith element in the candidate period set
$period_{S'}$	period of scaled seasonal component S'
U	discrete cosine feature matrix
D	matrix after discrete cosine transform
Z	standard matrix for discrete cosine transform
Q	principal information of matrix D

References

1. He, S.; Li, Z.; Wang, J.; Xiong, N.N. Intelligent Detection for Key Performance Indicators in Industrial-Based Cyber-Physical Systems. IEEE Trans. Ind. Inform. 2021, 17, 5799–5809.

2. Wu, D.; Zhou, D.; Chen, M.; Zhu, J.; Yan, F.; Zheng, S.; Guo, E. Output-Relevant Common Trend Analysis for KPI-Related Nonstationary Process Monitoring With Applications to Thermal Power Plants. IEEE Trans. Ind. Inform. 2021, 17, 6664–6675.

3. He, Y.; Flavel, A.; Ge, Z.; Gerber, A.; Massey, D.; Papadopoulos, C.; Shah, H.; Yates, J. Argus: End-to-end service anomaly detection and localization from an isp's point of view. In Proceedings of the 2012 Proceedings IEEE INFOCOM, Orlando, FL, USA, 25–30 March 2012; pp. 2756–2760.

4. Chen, Y.; Mahajan, R.; Sridharan, B.; Zhang, Z.L. A provider-side view of web search response time. ACM Sigcomm Comput. Commun. Rev. 2013, 43, 243–254.

5. Liu, D.; Zhao, Y.; Xu, H.; Sun, Y.; Pei, D.; Luo, J.; Jing, X.; Feng, M. Opprentice: Towards practical and automatic anomaly detection through machine learning. In Proceedings of the 2015 Internet Measurement Conference, Tokyo, Japan, 28–30 October 2015; pp. 211–224.

6. Laptev, N.; Amizadeh, S.; Flint, I. 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, Sydney, Australia, 10–13 August 2015; pp. 1939–1947.

7. Zhao, N.; Zhu, J.; Wang, Y.; Ma, M.; Zhang, W.; Liu, D.; Zhang, M.; Pei, D. Automatic and Generic Periodicity Adaptation for KPI Anomaly Detection. IEEE Trans. Netw. Serv. Manag. 2019, 16, 1170–1183.

8. Chen, W.; Xu, H.; Li, Z.; Pei, D.; Chen, J.; Qiao, H.; Feng, Y.; Wang, Z. Unsupervised anomaly detection for intricate kpis via adversarial training of vae. In Proceedings of the IEEE INFOCOM 2019-IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; pp. 1891–1899.

9. Ch, V.; Banerjee, A.; Kumar, V. Anomaly detection: A survey. ACM Comput. Surv. (CSUR) 2009, 41, 15.

10. Hyndman, R.J.; Athanasopoulos, G. Forecasting: Principles and Practice; OTexts: Melbourne, Australia, 2018.

11. Wang, Y.; Wang, Z.; Xie, Z.; Zhao, N.; Pei, D. Practical and White-Box Anomaly Detection through Unsupervised and Active Learning. In Proceedings of the 29th International Conference on Computer Communications and Networks (ICCCN), Honolulu, HI, USA, 3–6 August 2020.

12. Yaacob, A.H.; Tan, I.K.T.; Chien, S.F.; Tan, H.K. Arima based network anomaly detection. In Proceedings of the 2010 Second International Conference on Communication Software and Networks, Singapore, 26–28 February 2010; pp. 205–209.

13. Fang, Z.; Cheng, Q.; Mou, L.; Qin, H.; Zhou, H.; Cao, J. Abnormal electricity consumption detection based on ensemble learning. In Proceedings of the 2019 9th International Conference on Information Science and Technology (ICIST), Hulunbuir, China, 2–5 August 2019; pp. 175–182.

14. Himeur, Y.; Alsalemi, A.; Bensaali, F.; Amira, A. A novel approach for detecting anomalous energy consumption based on micro-moments and deep neural networks. Cogn. Comput. 2020, 12, 1381–1401.

15. Himeur, Y.; Elsalemi, A.; Bensaali, F.; Amira, A. Detection of appliance-level abnormal energy consumption in buildings using autoencoders and micro-moments. In Proceedings of the Fifth International Conference on Big Data and Internet of Things (BDIoT), Rabat, Morocco, 17–18 March 2021; pp. 1–13.

16. Deng, A.; Hooi, B. Graph neural network-based anomaly detection in multivariate time series. Proc. Aaai Conf. Artif. Intell. 2021, 35, 5.

17. Chen, Z.; Chen, D.; Zhang, X.; Yuan, Z.; Cheng, X. Learning graph structures with transformer for multivariate time series anomaly detection in iot. IEEE Internet Things J. 2022, 12, 9179–9189.

18. Zhou, Y.; Ren, H.; Li, Z.; Pedrycz, W. An anomaly detection framework for time series data: An interval-based approach. Knowl.-Based Syst. 2021, 228, 107153.

19. Xu, H.; Chen, W.; Zhao, N.; Li, Z.; Bu, J.; Li, Z.; Liu, Y.; Zhao, Y.; Pei, D.; Feng, Y.; et al. Unsupervised anomaly detection via variational auto-encoder for seasonal kpis in web applications. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 187–196.

20. Himeur, Y.; Alsalemi, A.; Bensaali, F.; Amira, A. Smart power consumption abnormality detection in buildings using micromoments and improved K-nearest neighbors. Int. J. Intell. Syst. 2021, 36, 2865–2894.

21. Li, J.; Izakian, H.; Pedrycz, W.; Jamal, I. Clustering-based anomaly detection in multivariate time series data. Appl. Soft Comput. 2021, 100, 106919.

22. Li, J.; Di, S.; Shen, Y.; Chen, L. 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, Virtual, 8–12 March 2021; pp. 824–832.

23. Carmona, C.U.; Aubet, F.X.; Flunkert, V.; Gasthaus, J. Neural contextual anomaly detection for time series. arXiv 2021, arXiv:2107.07702.

24. Himeur, Y.; Ghanem, K.; Alsalemi, A.; Bensaali, F.; Amira, A. Anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. Appl. Energy 2020, 289, 116601.

25. Thill, M.; Konen, W.; Wang, H.; Bäck, T. Temporal convolutional autoencoder for unsupervised anomaly detection in time series. Appl. Soft Comput. 2021, 112, 107751.