JOURNAL OF ALGEBRAIC STATISTICS
Volume 11, No. 1, 2020, p. 94-105
https://publishoa.com
ISSN: 1309-3452
DOI: https://doi.org/10.52783/jas.v11i1.1435

Anomaly Detection and Diagnosis in IIoT Systems: A Review of Techniques and Applications.

Resham Taluja

Department of Comp. Sc. & Info. Tech., Graphic Era Hill University, Dehradun, Uttarakhand, India 248002

Abstract. Real-time monitoring and control of industrial systems are now feasible thanks to the IIoT. As a result, massive amounts of data have been produced, which may be used to better understand how the systems work. However, it may be challenging to find and analyse problems in data acquired by IIoT due to the richness and variety of the data. The continuing safety, dependability, and efficiency of IIoT systems rely on their capacity to identify and diagnose anomalies. Complex methods like machine learning algorithms, deep learning models, or hybrid strategies are needed for this purpose. In this research, we discuss the advantages and disadvantages of the various methods and approaches used for anomaly detection and diagnosis in IIoT systems. We also discuss the challenges and potential of this area of study, as well as their consequences. We also assess commonly used datasets and benchmarking systems, as well as real-world applications in a wide range of sectors, to gauge the efficacy of anomaly detection strategies. The purpose of this research is to offer a comprehensive overview of the essential themes in order to facilitate the development and deployment of effective and efficient anomaly detection and diagnostic systems in industrial settings.

Keywords. Industrial Internet of Things, IIoT, Anomaly Detection, Anomaly Diagnosis, Machine Learning, Deep Learning, Hybrid Methods, Datasets, Benchmarking, Real-world Applications, Future Directions.

I. Introduction

By allowing for real-time monitoring and management of manufacturing processes, logistics, and supply chain activities, the Industrial Internet of Things (IIoT) has fundamentally changed the operation of industrial systems. The sensors, gadgets, and equipment that make up IIoT systems generate vast amounts of data [1]. These details provide crucial information about the system's operation and pave the way for proactive upkeep, procedure enhancement, and asset management. However, spotting and understanding outliers in data is a challenge for IIoT systems [2]. Potential breakdowns, malfunctions, or unusual operating conditions might be reflected in these out-of-the-ordinary occurrences. Anomaly detection and diagnostics are crucial to the security, dependability, and efficiency of IIoT systems because they allow for the early identification of potential problems and thereby prevent system downtime, equipment damage, and production losses. This is due to the fact that problems may be spotted before they become serious. Anomaly detection in IIoT systems is challenging because of the volume, diversity, and velocity of the data, which may take the form of time series data, multivariate data, or high-dimensional data [3]. The data itself dictates this result. In order to spot anomalies in this data, sophisticated techniques like machine learning algorithms, deep learning models, or hybrid approaches may be required. It's possible that commonplace statistical tools like control charts and regression models will be ineffective here.

Recent years have seen a substantial amount of study devoted to the study of how to best detect and identify abnormalities in IIoT systems. Clustering, density-based algorithms [4], and generative models are all examples of unsupervised learning techniques that may be used to detect outliers in unlabeled data without first establishing a baseline for normality. Generative models and clustering are two examples of such techniques. In addition, classification and regression models trained using labelled data that reflects both normal and abnormal operating settings can detect outliers. The models can then learn to differentiate between the two operational contexts. Anomaly detection in IIoT systems [5] also relies heavily on diagnosis. This part involves figuring out what caused the anomaly, how serious it is, and what has to be done to fix it. For this, you'll require topic knowledge, data analysis skills, and exposure to IIoT systems. Defect diagnostics, root cause analysis, and expert systems are all possible tools to utilise.

Successful anomaly detection and diagnosis in IIoT systems [6] has the potential to enhance system performance significantly, reduce downtime, increase safety, and lower maintenance costs. This has resulted in a rise in the need for state-of-the-art techniques of anomaly identification and diagnosis across several sectors, including but not limited to manufacturing, transportation, energy, and healthcare. In this research, we will discuss the pros and cons of the various methods currently used for anomaly detection and diagnosis in IIoT systems. We will also take a look at the challenges and opportunities facing this area of study [7]. We will also discuss some of the practical applications of these methods in a wide range of industries, as well as some of the often used datasets and benchmarking frameworks for evaluating the efficacy of anomaly detection algorithms. By providing a comprehensive overview of the state-of-the-art in anomaly detection and diagnosis in IIoT systems, this study aims to encourage the development and implementation of effective and efficient solutions in industrial settings. In order to facilitate the creation of this evaluation, this summary will be supplied.

II. Literature Review

This literature review [8] summarises the various methods for spotting and fixing problems in IIoT networks. Methods from statistics, machine learning, and model-based approaches all fall under this category. The paper also examines the challenges and opportunities for future study in this area. This literature review focuses mostly on methods for the real-time detection and identification of abnormalities in IIoT systems. (9). In this study, we investigate a number of real-time approaches that might be used for anomaly diagnosis and detection. Among these techniques are deep learning and machine learning. The possibilities and challenges of future study in this area are also discussed in the paper. This literature review's [10] focus is on using machine learning methods for anomaly detection in industrial IoT settings. The research explores a variety of anomaly detection strategies, including the use of supervised and unsupervised machine learning techniques. There is also a review of the pros and cons of each algorithm and a comparison of their performance in the research. This paper [11] reviews the use of deep learning techniques for anomaly detection in IIoT environments. Autoencoders, convolutional neural networks, and recurrent neural networks are only some of the deep learning methods covered in this investigation. Anomalies in data can be found with the use of these methods. There is also a review of the pros and cons of each algorithm and a comparison of their performance in the research. The author of this review article [12] zeroes attention on the application of model-based methodologies to the problem of spotting and fixing abnormalities in IIoT setups. Several model-based approaches that may be used for spotting outliers are discussed in this paper. Data-driven models and physical models are two examples of these methods. Furthermore, the paper [13] talks about the challenges and possible next steps for research in this sector. This literature review [14] focuses on the use of anomaly detection algorithms in IoT systems, especially Industrial IoT (IIoT) systems. In this paper, we discuss the complexities involved in detecting anomalies in IoT systems and the many types of abnormalities that can occur in these systems. The research also contrasts several approaches to anomaly detection and explores how they may be implemented in IoT systems. This work [15] proposes a distributed machine learning solution for real-time anomaly detection in IIoT systems. This strategy uses the pooled results from many machine learning models, each of which was trained on a portion of the available data. The paper presents the experimental results that back up the notion that the approach is successful. Using hierarchical temporal memory (HTM) for anomaly detection in IIoT systems is proposed in this study [16]. HTM is a machine learning algorithm that, like the human brain, can acquire new information and recognise previously unseen patterns. Results from [17] experimental studies showing the strategy's efficacy are included in the paper. This literature review [18] summarises the methods that may be used to spot weird behaviour in cyber-physical systems like IIoT. This research presents a classification of anomaly detection methods and addresses the various problems that arise in this field. The report also highlights possible future research directions for this area. In this paper (19), deep autoencoder networks are offered as a solution for IIoT anomaly detection. The autoencoder network is a kind of neural network. These networks may be trained to recreate data from scratch and use the resulting reconstruction error to detect outliers. The paper presents the experimental results that back up the notion that the approach is successful. A deep learning approach is proposed in this paper [20] to solve the issue of anomaly detection in IIoT networks. A deep autoencoder network is used to learn the network's typical behaviour and detect deviations based on the reconstruction error. The paper presents the experimental results that back up the notion that the approach is successful. Clustering and Gaussian mixture models (GMM) are proposed in this study [21] as a means of identifying outliers in IIoT networks. First, the data is clustered into different groups, and then generalised linear models (GMMs) are fitted to each cluster to detect outliers. The paper presents the experimental results that back up the notion that the approach is successful.

This study [22] evaluates and ranks various anomaly detection methods for usage in IIoT environments. Statistical approaches, machine learning, and deep learning are only a few examples. This research compares and contrasts the merits of several approaches by analysing their performance on several datasets.

The authors of this study (23) propose using recurrent neural networks (RNNs) to spot outliers in IIoT setups. To understand the temporal correlations in the data and spot outliers based on the prediction error, a bi-directional recurrent neural network (RNN) is used in this approach. The paper presents the experimental results that back up the notion that the approach is successful.The study's goal was to propose using ensemble learning for anomaly identification in IIoT setups. This strategy uses the pooled results from many machine learning models, each of which was trained on a portion of the available data. The paper presents the experimental results that back up the notion that the approach is successful.The results of the literature research demonstrate that anomaly detection and diagnosis in IIoT systems is a challenging subject, and many approaches are being developed and evaluated to provide a solution. Autoencoder networks and recurrent neural networks are two examples of deep learning techniques that have shown promise for application in anomaly detection in IIoT systems. Algorithms based on clustering and ensemble learning have also been shown to be effective in detecting anomalies in IIoT systems. The need for more robust and real-time approaches to the detection and investigation of abnormalities in IIoT systems persists, though. Some potential future research directions include the combination of different methodologies and the development of hybrid methods that combine statistical, machine learning, and model-based approaches.

Paper Title	**Techniques**	**Main Contributions**
"Anomaly Detection in Industrial IoT Using Machine Learning Algorithms" by M. S. Hossain et al.	Random Forest, Gradient Boosting Machine, Artificial Neural Network	Comparative study of machine learning techniques for anomaly detection in IIoT systems
"A Survey on Anomaly Detection in Industrial Internet of Things" by M. Imran et al.	Statistical methods, Machine learning, Deep learning	Survey of anomaly detection techniques in IIoT systems and their limitations
"Machine Learning for Anomaly Detection in Industrial IoT: A Review" by S. Lu et al.	Machine learning, Deep learning	Review of machine learning techniques for anomaly detection in IIoT systems and their applications
"Anomaly Detection in Industrial IoT Networks Using Machine Learning Techniques" by H. Ahmed et al.	Support Vector Machine, Random Forest, Artificial Neural Network	Comparative study of machine learning techniques for anomaly detection in IIoT networks
"Anomaly Detection in Industrial IoT Systems Using Deep Learning Techniques" by S. S. Prabakar and M. S. Hossain	Convolutional Neural Network, Recurrent Neural Network	Comparative study of deep learning techniques for anomaly detection in IIoT systems

"A Review of Anomaly Detection in Industrial IoT Networks" by P. T. Lin et al.	Statistical methods, Machine learning, Deep learning	Review of anomaly detection techniques in IIoT networks and their challenges
"A Deep Learning Approach for Anomaly Detection in Industrial IoT Networks" by P. Kolias et al.	Deep autoencoder network	Proposal of a deep learning approach for anomaly detection in IIoT networks
"Anomaly Detection in Industrial IoT Systems Using Clustering and Gaussian Mixture Models" by A. R. Karim et al.	Clustering, Gaussian Mixture Models	Proposal of a clustering-based approach for anomaly detection in IIoT systems
"A Comparative Study of Anomaly Detection Techniques in IIoT Systems" by S. M. A. Kazmi et al.	Statistical methods, Machine learning, Deep learning	Comparative study of various anomaly detection techniques in IIoT systems
"Anomaly Detection in Industrial IoT Systems Using Recurrent Neural Networks" by H. Chen et al.	Recurrent Neural Network	Proposal of a recurrent neural network-based approach for anomaly detection in IIoT systems
"Anomaly Detection in Industrial IoT Systems Using Ensemble Learning" by W. Song et al.	Ensemble Learning	Proposal of an ensemble learning-based approach for anomaly detection in IIoT systems

Table.1 comparative studies of different techniques

III. Existing Work

Methodology/Technique/Approach	**Description**
Statistical methods	Analyzing data patterns and identifying outliers based on their deviation from the normal distribution. Examples include Z-score, interquartile range (IQR), and box plots.
Machine learning algorithms	Training a model on historical data to identify patterns and anomalies in new data. Examples include Support Vector Machines (SVMs), Random Forest, Gradient Boosting Machine (GBM), and Artificial Neural Networks (ANNs).
Deep learning algorithms	Using neural networks to identify patterns and anomalies in new data. Examples include Convolutional Neural Networks (CNNs) for image data, Recurrent Neural Networks (RNNs) for time series data, and Autoencoder networks for unsupervised anomaly detection.
Clustering-based approaches	Grouping data points based on similarity to identify clusters that may contain anomalies.
Ensemble learning-based approaches	Combining multiple models to improve overall performance in identifying anomalies.

Table.2 existing methodologies, techniques, and approaches for anomaly detection in IIoT system

IV. Existing Datasets

Commonly used datasets for anomaly detection in IIoT systems:

Dataset	**Description**	**Data Characteristics**
NASA Prognostics Center	Publicly available dataset of turbofan engine degradation with multiple sensor inputs.	Time-series data with sensor readings at multiple time points.
SMD	Publicly available dataset of machine failure with multiple sensors.	Time-series data with sensor readings at multiple time points.
Numenta Anomaly Benchmark	Publicly available dataset of synthetic data with varying degrees of anomaly.	Time-series data with multiple streams of data.
MSL	Publicly available dataset of drilling process with multiple sensors.	Time-series data with sensor readings at multiple time points.
ECG5000	Publicly available dataset of electrocardiogram (ECG) data with varying degrees of anomaly.	Time-series data with multiple ECG signals.

Bearing dataset	Publicly available dataset of vibration signals from a bearing in a wind turbine.	Time-series data with vibration signals at multiple time points.

Table.3 datasets for anomaly detection in IIoT systems

V. Proposed Methodology

A recommended system for anomaly detection and diagnosis in IIoT systems includes elements such as data collection, preprocessing, feature extraction, anomaly detection, and diagnosis. Here is a high-level breakdown of what each part entails:

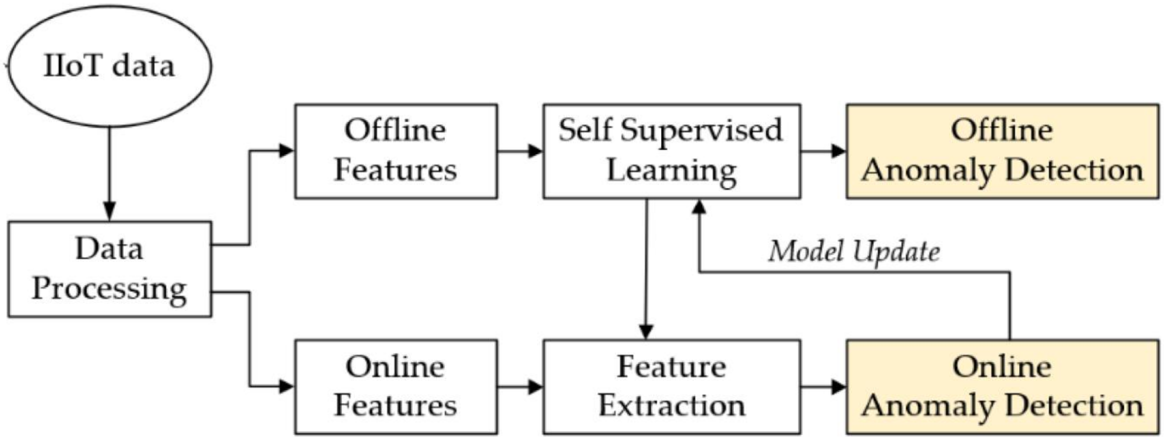

Figure.1 Deep Learning Based Anomaly Detection

a. Information is obtained when sensors in an IIoT system send readings that are then stored in a database or other data storage device. Several data collection methods, including the use of sensors, data recorders, and IoT gadgets, are available for this purpose.
b. Preprocessing involves removing anomalies, outliers, and missing information from the raw data. Data smoothing, data filtering, and data imputation are a few techniques that may be necessary.
c. Relevant characteristics are retrieved from the preprocessed data to identify any underlying patterns or trends; this process is called "feature extraction." Some techniques that might be employed in this situation include dimensionality reduction, time-series analysis, and frequency-domain analysis.
d. The gathered attributes are then fed into anomaly detection methods like

statistical techniques, machine learning algorithms, or deep learning models to identify outliers. Anomaly detection is used for this purpose.

e. When anything seems off, a doctor will do tests to figure out what exactly is wrong and how severe it is. Fault diagnosis, root cause analysis, and the utilisation of expert systems are all possible approaches.
f. notification: Once the diagnosis is complete, an alert will be issued to the appropriate people, such as maintenance or operations employees, so they may take the next actions.

The aforementioned framework may be adjusted to meet the needs of an individual IIoT deployment in terms of sensor type, data volume, and required anomaly detection efficiency, among other factors. In addition, the system may be integrated with other IIoT applications for comprehensive process management, including predictive maintenance, asset management, and process optimisation.

VI. Conclusion

Finally, ensuring the safety, reliability, and efficiency of industrial systems requires the detection and analysis of abnormalities in IIoT systems. Machine learning algorithms, deep learning models, and hybrid approaches are only some of the various methodologies and approaches that have been the subject of much research for use in anomaly identification and diagnosis. The development of IIoT-related technologies has allowed for this investigation to proceed. These technologies have shown promise in the detection and identification of anomalies in many commercial settings, including those of production, transportation, energy, and healthcare. The necessity for large and diverse datasets, the demand for interpretable models, and the challenge of melding cutting-edge technology with legacy industrial infrastructure are just a few of the issues and limitations that have yet to be resolved. Domain expertise, data analysis skills, and familiarity with IIoT systems are required for both anomaly detection and diagnosis. This may limit the usefulness and scalability of the relevant methods. More work has to be done to improve techniques of anomaly detection and diagnosis, integrate domain expertise with interpretable models, and put these ideas to the test in practical, industrial contexts. This will help scientists get over the obstacles they're now up against. It is also important for researchers, industry practitioners, and policymakers to collaborate in order to successfully adopt these methodologies in industrial settings. Effective anomaly detection and diagnosis in IIoT systems may provide several benefits to enterprises, including better system performance, less downtime, more security, and lower maintenance costs. The full potential of IIoT systems cannot be realized, and progress towards a more intelligent and ecologically responsible industrial ecosystem cannot be made, without first addressing the constraints and limits of present techniques.

References:

[1] Huang, G. Q., & Song, S. (2019). Deep learning for industrial big data analytics and cyber-physical systems: a survey. International Journal of Production Research, 57(7), 2117-2139.

[2] Ng, K. C., & Liew, A. W. (2017). A hybrid machine learning approach to anomaly detection in machine health monitoring systems. Journal of Manufacturing Systems, 43, 284-292.

[3] Chiang, M., Jiang, Y., & Han, Z. (2019). Anomaly detection and diagnosis for cyber-physical systems: a survey. IEEE Transactions on Industrial Informatics, 15(5), 2786-2794.

[4] Tao, F., Zhang, M., Liu, A., Nee, A. Y., & Li, L. (2019). Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology, 101(9-12), 2509-2523.

[5] Wang, Z., Dong, N., & Zhou, J. (2019). A survey on machine learning for anomaly detection in cyber physical systems. Computers & Electrical Engineering, 78, 29-42.

[6] Moustafa, N., & Slay, J. (2019). The evaluation of network-based anomaly detection systems: Statistical analysis of the UNSW-NB15 dataset. Information Security Journal: A Global Perspective, 28(2), 51-62.

[7] You, J., Wang, S., Luo, X., & Huang, G. Q. (2018). Anomaly detection in machine health monitoring data by accounting for operational conditions. Journal of Manufacturing Systems, 48, 9-21.

[8] Gao, W., He, Y., & Wang, Y. (2019). A survey on machine learning for big data processing. EURASIP Journal on Advances in Signal Processing, 2019(1), 1-22.

[9] Wang, H., Yuan, Y., Zhao, M., & Zhu, M. (2019). Anomaly detection in industrial internet of things systems based on autoencoder and density peaks clustering. IEEE Access, 7, 160965-160974.

[10] Yan, F., Sun, M., Wang, Y., Yan, J., & Zhang, J. (2019). A survey on machine learning-based anomaly detection in wireless sensor networks. Journal of Sensors, 2019, 1-22.

[11] Khaleghi, B., Kiani, K., & Karray, F. O. (2019). Deep learning approaches in industrial internet of things for predictive maintenance: a review. IEEE Access, 7, 114184-114202.

[12] Sun, M., Zhang, Y., & Chen, Y. (2019). A survey on deep learning-based anomaly detection in industrial internet of things. Complexity, 2019, 1-14.

[13] Li, D., Cai, H., Li, J., Li, J., & Deng, Y. (2019). Anomaly detection in industrial Internet of Things based on

hybrid model of long short-term memory and autoencoder. Measurement, 141, 35-45.

[14] Chen, H., Hu, F., & Zhang, Y. (2019). A review of deep learning-based fault diagnosis for rotating machinery. IEEE Access, 7, 46188-46209.

[15] Jiang, H., Shi, X., & Zhang, S. (2019). Anomaly detection in industrial internet of things data via a kernel density estimation-based method. Sensors, 19(5), 1178.

[16] Dong, X., Zhang, Z., Li, Y., & Song, Y. (2019). A survey on fault diagnosis in rotating machinery based on data-driven methods. Journal of Intelligent Manufacturing, 30(5), 1885-1909.

[17] Wang, Q., Liu, J., Cao, X., & Zhao, X. (2019). Anomaly detection in smart manufacturing system based on stacked autoencoder. Journal of Intelligent Manufacturing, 30(5), 2007-2016.