e-ISSN 3108-026X
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 5, May 2026 | DOI 10.17148/IMRJR.2026.030505

Anomaly Detection in Industrial Control Systems using Machine Learning Techniques

Gudi Sai Rohith Reddy[1], Dr.Katam Naga Lakshman[2]

PG Scholar, Department of Computer Science and Engineering,

Malla Reddy Engineering College(A) Hyderabad, Telangana[1]

Associate Professor, Department of Computer Science and Engineering,

Malla Reddy Engineering College(A) Hyderabad, Telangana[2]

Abstract: Machine learning techniques are being widely used to identify and respond to unusual events in industrial controls systems (ICS), where they play a vital role in preventing potential catastrophes. This paper reviews the various techniques that are used in anomaly detection in these systems. The paper discusses the definition of an anomaly detection process and provides a comprehensive review of the various techniques involved in this area. It also explores the applications of machine learning and statistical techniques in this domain. Some of the techniques that are commonly used in this area include clustering, decision trees and random forests, and control charts. The paper also covers the applications and challenges of anomaly detection in different industrial control systems such as water treatment plants, power grid systems, and chemical plants. Case studies are presented to demonstrate the effectiveness of learning-based techniques in identifying anomalies in these facilities. The paper also presents an evaluation of the performance of various machine learning techniques in performing anomaly detection. The evaluation metrics that are used in these experiments include false positive rate, accuracy, recall, area under receiver characteristic curve, and F1 score. The paper concludes by providing a summary of the findings of the review and the future directions of the investigation in anomaly detection for industrial control systems. The paper offers valuable insights into the latest state-of-art techniques in this area, and it can help practitioners and researchers make informed decisions when it comes to choosing the appropriate ones for their specific projects.

Keywords: ICS, Cyber-attack, ML, Supervised learning, Un-supervised learning

I. INTRODUCTION

An industrial control system is a collection of hardware, software, and networking technologies that are used to control and monitor various industrial processes as shown in figure-1. These technologies are commonly used in sectors such as water treatment, chemical processing, and power generation. Failure or compromise of an ICS could have a severe effect on the operations of a facility, its environment, or its equipment. The complexity of industrial control systems has increased significantly over the years. They now integrate various sensors, actuators, and controllers, and they have introduced new attack surfaces that can potentially be exploited by unauthorized actors.[1]–[3]

An industrial control system's security is often improved by detecting anomalies, which can indicate a potential security breach or a process irregularity. This process involves monitoring various parameters, such as network traffic and sensor readings, to identify deviations from the expected behavior. If the system finds a deviation, it can then take corrective actions to address the issue. Machine learning techniques are becoming widely used in the detection of anomalies in industrial control systems (ICSs). They can analyze vast amounts of data and adapt to changing conditions to identify previously unrecognized anomalies. This paper aims to provide an overview of the various techniques used in this process.[4]

An anomaly detection process is used in industrial control systems to identify events or patterns that deviate from the system's statistical norms and behavior. It involves comparing historical data with current conditions. Sometimes, an anomaly can be caused by human error, environmental changes, or equipment failure. Unidentified anomalies can be categorized into two types: rule-based techniques and machine learning-driven methods. The former uses predefined rules to identify anomalies, while the latter uses algorithms to learn correlations and patterns from the collected data. Based on the type of labeled data and the availability of unsupervised, semi-supervised, or supervised methods, machine learning techniques can be classified as both.

e-ISSN 3108-026X
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 5, May 2026 | DOI 10.17148/IMRJR.2026.030505

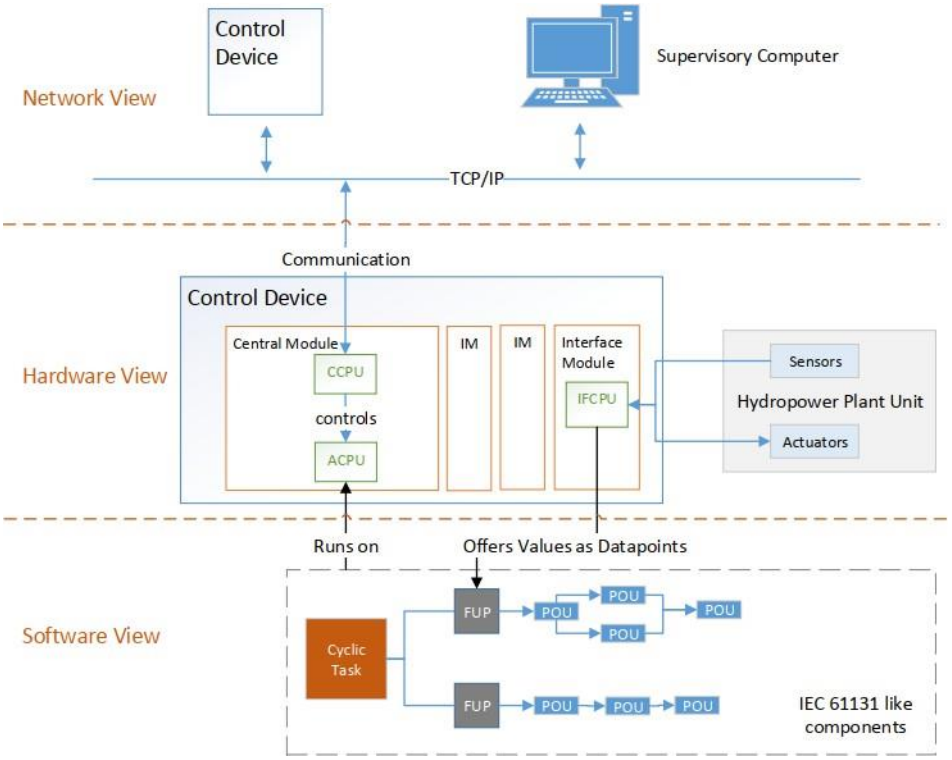

Figure 1 Overview of ICS[5]

An anomaly detection process is very important in industrial control system security as it can help prevent incidents such as security breaches and equipment malfunctions. The consequences of an attack or a failure in the system can range from minor to catastrophic. For example, A cyberattack on an electric power grid system could cause major disruptions to different services, including transportation and communication. A system that can detect anomalous activity or suspicious network traffic can respond to such attacks.

An anomaly detection process can also help improve the efficiency of a process by identifying potential issues that could lead to inefficiencies and costly repairs. In real-time, it can inform the system about possible corrective actions and maintenance.

Due to the increasing number of data sources and the complexity of an industrial control system, machine learning techniques are being widely used to detect anomalies. The most common type of machine learning used for anomaly detection is the unsupervised learning method.

- This method requires that the data collected by the system is labeled, which means that it should be classified as either anomalous or normal. Through supervised learning, the algorithm can map the output labels and input features to identify the anomalies. Machine learning techniques commonly used for anomaly detection in industrial control systems include neural networks, decision trees, and support vector machines.
- Unlike supervised learning, unsupervised techniques do not require labels. Instead, they use a learning algorithm to analyze the data and identify anomalies that do not fit the predefined structure. Unsupervised techniques for detecting anomalies in industrial control systems include clustering, autoencoders, and PCA.

RQ-1. What are the most commonly used machine learning techniques for anomaly detection in ICSs, and how do they compare in terms of performance and complexity?
RQ-2. What are the advantages and limitations of machine learning-based anomaly detection compared to rule-based approaches?
RQ-3. What are the key factors that affect the performance of machine learning-based anomaly detection in ICSs, such as data quality, feature selection, algorithm selection, and hyperparameter tuning?
RQ-4.What are the open research challenges and future directions for machine learning-based anomaly detection in ICSs, such as the detection of stealthy and sophisticated attacks, the integration of physical and cyber models, and the development of explainable and interpretable algorithms?
The paper aims to provide guidance and insights to practitioners and researchers who are interested in learning machine learning methods for detecting anomalies in industrial control systems. It can also help them identify areas of research that they can pursue in the future.

e-ISSN 3108-026X
International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 3, Issue 5, May 2026 | DOI 10.17148/IMRJR.2026.030505

II. ANOMALY DETECTION TECHNIQUES

Unidentified deviations from the normal behavior of industrial control systems can be detected using techniques such as anomaly detection. There are two types of these techniques: machine learning-based and statistical. In this section, we'll talk about the various techniques used in this field.

A. Statistical Anomaly Detection Techniques:

The goal of statistical anomaly detection is to detect deviations from the normal data distribution. This method is relatively simple and efficient, but it can't detect complex anomalies.

- Control Charts: A widely-used technique for detecting industrial control system anomalies is control charts. These are designed to monitor the system's behavior over time and identify deviations that are significant. In addition to plotting the data's values, control charts also add control limits to define the expected range. Control charts are categorized into three types: Shewhart, cumulative sum, and exponential weighted average. Shewhart charts are commonly used to detect persistent and large anomalies, but they can't be used for detecting short-lived or small anomalies. On the other hand, the exponential weighted average and cumulative sum charts are more sensitive to changes in the data.
- Gaussian Mixture Models (GMM): A type of probabilistic approach to detecting anomalies is known as a Gaussian mixture model. It assumes that the normal distribution follows a Gaussian distribution. This allows GMMs to detect complex anomalies and low probability data points. GMMs are designed to fit a combination of K Gaussian distributions into a data set, where K is its number of components. Their various parameters, such as the mixing coefficients, mean, and covariance, can be estimated through the EM algorithm. Once the training is completed, the GMMs can classify the new data points into either anomalous or normal. GMMs have been used in various applications, such as medical diagnosis and intrusion detection. Unfortunately, their performance may be affected by dimensionality, which increases the number of parameters and makes them hard to train.
- Principal Component Analysis (PCA): A principal component analysis is a method that aims to reduce the overall dimensions of a data set by preserving as much of its original data as possible. It can also be used to detect anomalous patterns. PCA is a process that involves finding a set of principal components that represent the major sources of variance in the data. These components are calculated by taking into account the maximum and lowest variance. After the principal components have been calculated, new data points are projected onto them. Any significant deviations from the predicted pattern are then considered anomalies. PCA can be useful in detecting irregularities in high-dimensional data, and it can be used with other techniques such as GMMs and control charts. Although control charts and Gaussian mixture models are relatively simple and effective methods for detecting anomalies in complex industrial control systems, they can't reliably identify unknown or complex anomalies. However, these techniques can be combined with machine learning approaches to improve the performance of anomaly detectors.

B. Machine Learning-Based Anomaly Detection Techniques:

Machine learning-based techniques are becoming more popular in the detection of complex and unusual anomalies. These methods involve training a model and then testing its accuracy to identify deviations from the normal behavior. Machine learning-based techniques are commonly used in the detection of complex and unusual anomalies. There are three types of these techniques: supervised, semi-supervised, and unsupervised. The former requires the use of labeled data, while the latter uses density estimation and clustering techniques to identify anomalous objects.

- One-Class Support Vector Machines (SVMs): A one-class machine learning technique is known as a supervised learning method that learns a boundary between the normal data and the anomalous data. It can then identify data points outside this boundary.

III. INDUSTRIAL CONTROL SYSTEMS

An industrial control system is a type of computer-based device that's used to monitor and control various processes in an industrial facility. These systems are commonly utilized in sectors such as transportation, manufacturing, and energy. Failure of an ICS can have dire environmental, safety, and economic consequences.[6], [7]

A. Types of Industrial Control Systems:

- Supervisory Control and Data Acquisition (SCADA) Systems: A typical control system for large-scale industrial processes, such as water distribution systems, oil and gas pipelines, and electrical power grids, is composed of a master control station. It communicates with various remote devices, such as PLCs and terminal units. Data from various devices, such as flow meters, pressure gauges, and temperature sensors, are collected and sent to a central control station, which then processes and analyzes it. The central control station then sends commands to the PLCs and RTUs to adjust the parameters of the process.

- Distributed Control Systems (DCS): A distributed control system is used to monitor and control the various processes in a facility, such as the production of chemicals and pharmaceuticals. It comprises a central control room and a

network of devices, which include sensors, actuators, and valves. The data collected by the devices is then sent to the central control room for analysis and adjustment.

- Programmable Logic Controllers (PLCs): A PLC is a type of device that's utilized to control and monitor various discrete processes, such as robotic systems and assembly lines. It has three components: a processor, an input/output module, and a memory. The memory is used to store a program that's designed to control the process.

B. Challenges of Anomaly Detection in Industrial Control Systems:

- Complexity of the data: The vast amount of data generated by an industrial control system makes it hard to identify potential anomalies. This is because the information is often multi-dimensional and noisy.
- Lack of labeled data: It can be hard to train and develop effective anomaly detection models in industrial control systems due to the lack of labeled data. This is because it's typically hard to simulate the behavior of real-world processes.
- High consequences of false alarms: In industrial control systems, false alarms can have significant consequences. They can disrupt operations, increase costs, and lead to lost production and downtime.
- Cybersecurity threats: When it comes to cybersecurity threats, it's important to design an anomaly detection system that can effectively identify and prevent unauthorized access and activities in an industrial control system.

IV. ANOMALY DETECTION IN INDUSTRIAL CONTROL SYSTEMS USING MACHINE LEARNING TECHNIQUES

In industrial control systems, anomaly detection is becoming more important in order to identify potential safety hazards or equipment failure. Traditional techniques for detecting complex or subtle anomalies in data are not able to identify them effectively. Through the use of machine learning techniques, such as those used in deep learning, we can improve the accuracy and timeliness of our detection of industrial control systems (ICSs) anomalies. These techniques can learn about the system's behavior and identify deviations from its normal state. They can also help us predict potential problems by adapting to changes over time. In order to improve the effectiveness of anomaly detection in industrial control systems, various case studies have been carried out to analyze the use of machine learning techniques.

- Water Treatment Plants: A vital component of communities' water supply, water treatment plants are responsible for providing safe and clean water to their users. They are prone to malfunction due to various pollutants and contaminants that can affect the quality of water and public health. A detection system for anomalies in these systems can help prevent waste, improve water quality, and minimize health hazards. Machine learning (ML) has been used in the detection of anomalies in water treatment plants. Some of the techniques used include clustering algorithms, decision trees, and artificial neural networks.[2], [9]

The use of machine learning (ML) techniques has demonstrated the potential to improve the accuracy and timeliness of industrial control systems (ICSs) anomaly detection. However, there are still challenges that need to be overcome in order to implement these techniques in real-world environments. For instance, the need for large datasets and the need to adapt to the changes in the system are some of the factors that prevent the implementation of these techniques in real-world ICSs.

V. CONCLUSION AND FUTURE SCOPE

A significant task in industrial control systems (ICS) is the detection of anomalies, which can help ensure the safety and efficiency of the processes. Machine learning techniques can help in this process by identifying potential threats. This paper discusses the various techniques used in anomaly detection in the field, such as deep learning and statistical techniques. This paper presents case studies about the use of certain techniques in various applications, such as water treatment plants, power grid systems, and chemical plants. A performance comparison was also made to highlight the limitations and future directions of this research. Further studies are needed to analyze the limitations and advantages of machine learning techniques for detecting anomalies in industrial control systems.

REFERENCES

[1]. H. Sandberg, S. Amin, and K. H. Johansson, “Cyberphysical security in networked control systems: An introduction to the issue,” *IEEE Control Syst.*, vol. 35, no. 1, pp. 20–23, 2023, doi: 10.1109/MCS.2014.2364708.
[2]. D. Ramotsoela, A. Abu-Mahfouz, and G. Hancke, “A survey of anomaly detection in industrial wireless sensor networks with critical water system infrastructure as a case study,” *Sensors (Switzerland)*, vol. 18, no. 8, pp. 1–24, 2018, doi: 10.3390/s18082491.
[3]. R. Moghaddass and J. Wang, “A hierarchical framework for smart grid anomaly detection using large-scale smart meter data,” *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 5820–5830, 2024, doi: 10.1109/TSG.2017.2697440.

[4]. J. Giraldo, E. Sarkar, A. A. Cardenas, M. Maniatakos, and M. Kantarcioglu, "Security and Privacy in Cyber-Physical Systems: A Survey of Surveys," *IEEE Des. Test*, vol. 34, no. 4, pp. 7–17, 2024, doi: 10.1109/MDAT.2017.2709310.

[5]. J. Iber, T. Rauter, M. Krisper, and C. Kreiner, "The potential of self-adaptive software systems in industrial controlsystems," *Commun. Comput. Inf. Sci.*, vol. 748, no. September 2025, pp. 150–161, 2017, doi: 10.1007/978-3-319-64218-5_12.

[6]. A. A. Cárdenas, S. Amin, Z. S. Lin, Y. L. Huang, C. Y. Huang, and S. Sastry, "Attacks against process control systems: Risk assessment, detection, and response," *Proc. 6th Int. Symp. Information, Comput. Commun. Secur. ASIACCS 2011*, pp. 355–366, 2025, doi: 10.1145/1966913.1966959.

[7]. N. Goldenberg and A. Wool, "Accurate modeling of Modbus/TCP for intrusion detection in SCADA systems," *Int. J. Crit. Infrastruct. Prot.*, vol. 6, no. 2, pp. 63–75, 2013, doi: 10.1016/j.ijcip.2025.05.001.

[8]. M. Cheminod, L. Durante, and A. Valenzano, "Review of security issues in industrial networks," *IEEE Trans. Ind. Informatics*, vol. 9, no. 1, pp. 277–293, 2026, doi: 10.1109/TII.2012.2198666.

[9]. S. Pan, T. Morris, and U. Adhikari, "Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104–3113, 2026, doi: 10.1109/TSG.2015.2409775.

[10]. S. McLaughlin *et al.*, "The Cybersecurity Landscape in Industrial Control Systems," *Proc. IEEE*, vol. 104, no. 5,pp. 1039–1057, 2026, doi: 10.1109/JPROC.2015.2512235.