International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(REVIEW ARTICLE)

Check for updates

The usage of AI-Based Anomaly Detection in SCADA Systems

Khalid Adnan Ali *

Department of Computer Science and Engineering, American University of Sharjah, Sharjah P.O. Box 26666, United Arab Emirates.

International Journal of Science and Research Archive, 2025, 17(01), 770-773

Publication history: Received on 05 September 2025; revised on 16 October 2025; accepted on 18 October 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.17.1.2839

Abstract

Supervisory Control and Data Acquisition (SCADA) Systems are critical for industrial infrastructures, especially to ensure the ability to seamlessly monitor and control assets remotely. However, the significant threats of cybersecurity breaches and operational faults stand in the way of complete reliance on SCADA systems. This review paper investigates the recent advancements in artificial intelligence (AI) and its ability to serve as an effective tool to detect anomalies which traditional rule-based systems often miss. AI-Driven anomaly detection is examined from two perspectives: operational anomalies-including process deviations, abnormal readings and equipment malfunctions-and cybersecurity anomalies such as intrusions and network attacks. Various modern AI methodologies are evaluated, alongside potential challenges like real-time processing demands and explainability.

Keywords: SCADA; Artificial Intelligence; Anomaly Detection; Cybersecurity

1. Introduction

Supervisory Control and Data Acquisition (SCADA) systems are essential for remote, centralized control and monitoring across key sectors like energy, water, and manufacturing. While these systems improve efficiency, they are exposed to significant threats, particularly operational anomalies (such as sensor and equipment failures) and cybersecurity issues (including unauthorized access and DoS attacks). Traditional rule-based anomaly detection struggles to keep up with the scale and complexity of modern SCADA environments, especially after the rise of IoT and high-profile attacks like BlackEnergy on Ukraine's grid.

Artificial Intelligence (AI) has emerged as a powerful tool for detecting both operational and cyber anomalies. AI algorithms, including machine learning (ML), deep learning (DL), and hybrid methods, offer greater accuracy by learning from historical and real-time data. ML models use supervised methods like Support Vector Machines (SVM) for labelled data and unsupervised methods like k-Nearest-Neighbors (kNN) for clustering anomalies. DL models, such as autoencoders, excel at identifying subtle or previously unseen anomalies by recognizing patterns in complex datasets.

This review evaluates how AI-based approaches can improve SCADA anomaly detection across both operational and cybersecurity domains, aiming to identify optimal techniques for different anomaly types.

1.1. Background

Understanding AI's role in SCADA anomaly detection requires a look at system architecture: SCADA integrates field devices (RTUs, PLCs), communication networks (using protocols like Modbus or IEC 60870-104), and central control centres. Operational anomalies typically arise in the field layer, such as sensor faults or control logic failures, while

* Corresponding author: Khalid Ali

cybersecurity anomalies often affect the communication layer but can impact the entire system. Notable incidents like Stuxnet and BlackEnergy highlight the potential for cyber threats to cause major physical damage.

With SCADA systems now more interconnected due to technological advances, new vulnerabilities have emerged, making advanced detection essential. AI techniques—spanning ML, DL, and hybrid approaches—are proving vital in identifying subtle and novel threats, helping secure SCADA's evolving landscape.

2. Operational Anomaly Detection

Operational anomalies in SCADA systems, such as equipment malfunctions and sensor faults, can significantly impact system reliability and may lead to major failures if not detected and rectified in a timely manner. These operational anomalies cover a wide range of possibilities, ranging from minimal impacts such as temporary data inconsistency to severe consequences such as area-wide blackouts or threats to human life. In general, the detection of these anomalies relies on the operator to notice an abnormal alarm in between hundreds of other alarms, hence the importance of effective alarm management in SCADA systems. However, placing the responsibility of detection solely on operators susceptible to human error puts the infrastructure at risk which could be avoided through the utilization of AI models. By training AI models on historical data they are able to learn patterns and identify any deviations in real-time data, which could suggest faults or failures.

AI-based anomaly detection methods enhance system reliability by providing early warning of potential faults and enabling predictive maintenance. By automatically identifying abnormal patterns before they escalate into major failures, these methods reduce downtime, prevent costly breakdowns, and help operators prioritize maintenance activities. This proactive approach improves the efficiency of maintenance planning and reduces the reliance on manual monitoring, which is prone to human error, ultimately safeguarding the continuity and safety of critical infrastructure.

Recent case studies demonstrate the effectiveness of AI in operational anomaly detection within SCADA systems. For example, AI-driven predictive maintenance has been successfully applied in water treatment plants to anticipate pump failures, and in power grids to detect transformer faults before outages occur. These implementations have shown significant reductions in unplanned downtime, improved maintenance scheduling, and enhanced overall system resilience. The literature highlights the transition from rule-based alarm management to data-driven AI solutions as a key factor in achieving these improvements.

3. Cybersecurity Anomaly Detection

Cybersecurity anomalies in SCADA systems, such as unauthorized access attempts, data breaches, or malicious command injections, pose significant threats to the integrity and availability of critical infrastructure. Unlike operational anomalies that stem from physical or equipment-related faults, cybersecurity anomalies originate from intentional or accidental digital actions that can compromise control systems and lead to severe operational disruptions. These incidents may range from minor unauthorized logins or abnormal traffic patterns to large-scale cyberattacks capable of halting essential services or causing widespread blackouts.

Traditionally, cybersecurity monitoring in SCADA systems relies on static rule-based intrusion detection and human oversight. However, with the increasing complexity and interconnectivity of modern control networks, relying solely on predefined rules and manual analysis is no longer sufficient. Human operators may overlook subtle indicators of compromise, especially when embedded within massive volumes of network traffic and event logs [1].

AI-based cybersecurity anomaly detection addresses these challenges by leveraging historical network and system behavior data to identify deviations that could indicate potential cyber threats. Machine learning algorithms can learn the normal operational patterns of communication between devices and detect anomalies such as abnormal data flows, unexpected command sequences, or unauthorized system access in real time.

By implementing AI-driven cybersecurity anomaly detection, organizations can achieve proactive threat identification and rapid response, minimizing the risk of system compromise. These models enhance situational awareness, reduce false alarms, and support continuous monitoring without overburdening human operators. Recent advancements demonstrate the success of AI in detecting zero-day attacks, identifying malware-infected nodes, and uncovering insider threats before they escalate [3]. This shift from reactive, rule-based defenses to adaptive, data-driven models represents a transformative step toward securing critical infrastructure against evolving cyber threats.

4. AI Techniques Used

Prominent AI/ML techniques in SCADA anomaly detection include Support Vector Machines (SVM), Random Forest, Neural Networks (such as deep learning models and LSTM), and Autoencoders. SVM and Random Forest are favored for their robustness with tabular and structured data, while neural networks and autoencoders excel in capturing complex patterns in time series and high-dimensional datasets. Another option is combining multiple algorithms to create hybrid or ensemble approaches to improve detection accuracy and robustness. Examples include stacking classifiers, voting ensembles, and blending supervised with unsupervised methods. These strategies leverage the complementary strengths of different models, helping to reduce false positives and adapt to diverse anomaly types in SCADA environments. In general, each algorithm has its strengths and weaknesses, making the choice of algorithm differ from case to case. SVM and Random Forest are interpretable and efficient but may struggle with complex temporal dependencies. Whereas neural networks and LSTM handle sequential data and high dimensionality well, though they require large datasets and are less transparent. As for autoencoders, they are effective for unsupervised anomaly detection but can be sensitive to hyperparameter tuning and data quality.

Transfer learning and reinforcement learning are also gaining traction. Transfer learning allows models to leverage knowledge from related tasks, improving performance with limited SCADA data. Reinforcement learning optimizes detection strategies through feedback-driven learning, enabling adaptive responses to evolving threats and dynamic system conditions.

Through advanced AI techniques, we are able to address SCADA data's high dimensionality, imbalance, and time series nature through feature selection, data augmentation, and temporal modeling (e.g., LSTM). Ensemble methods and specialized loss functions help mitigate class imbalance, while deep learning architectures can process complex, sequential sensor data effectively.

5. Challenges and Limitations

AI-driven SCADA anomaly detection faces challenges such as data scarcity and label imbalance. Attack events are rare compared to normal operations, resulting in limited labelled data for training. This imbalance can cause models to favor normal patterns, making it difficult to accurately detect anomalies. Additionally, acquiring high-quality, representative datasets from critical infrastructure environments is often constrained by privacy and operational concerns.

Real-time processing is crucial in SCADA environments to ensure timely detection and response to threats. AI models must analyze high volumes of continuous, streaming data with minimal delay. However, complex algorithms can introduce significant latency, potentially delaying alerts and response actions. Therefore, achieving both high detection accuracy and low latency is critical but challenging as it requires efficient model design and optimization for real-time deployment.

Due to the sensitive nature of SCADA systems, explainability is essential for operators to understand and trust AI-driven alerts. Many advanced models, such as deep neural networks, act as "black boxes," making it difficult to interpret their decisions. Lack of transparency can hinder operator confidence and slow incident response, emphasizing the need for interpretable AI methods that provide clear, actionable explanations for detected anomalies.

AI systems although may be used in SCADA cybersecurity to detect and prevent attacks, they can also introduce new attack vectors, such as adversarial attacks that manipulate model inputs to evade detection or trigger false alarms. Attackers may also target the training data or model parameters, compromising the integrity of the detection system. Ensuring the robustness and security of AI models is therefore a critical consideration in their deployment.

6. Conclusion

AI significantly improves anomaly detection in SCADA systems by leveraging advanced machine learning techniques to identify unusual patterns and threats more accurately. Its ability to process complex, high-dimensional, and sequential sensor data enables faster and more reliable detection, supporting proactive security and operational monitoring. Key AI/ML methods—including SVM, Random Forest, neural networks, and autoencoders—offer complementary strengths in detecting anomalies. Ensemble and hybrid approaches further enhance accuracy, while transfer and reinforcement learning support adaptability. These techniques help mitigate class imbalance and effectively manage time-series data for robust anomaly detection. Challenges include limited labelled data, class imbalance, latency in real-time detection, and model explainability. Privacy and operational constraints restrict dataset quality. Future research should focus on

improving model transparency, efficiency, and resilience against adversarial attacks to ensure reliable deployment in critical SCADA environments. AI adoption in critical infrastructure enhances resilience and security by enabling early detection and rapid response to threats. Improved anomaly detection supports operational continuity and safeguards essential services, contributing to the overall robustness of industrial systems.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Anwar, M., Lundberg, L. & Borg, A. Improving anomaly detection in SCADA network communication with attribute extension. *Energy Inform* **5**, 69 (2022). https://doi.org/10.1186/s42162-022-00252-1

[2] Almalawi, A., Hassan, S., Fahad, A., Iqbal, A., & Khan, A. I. (2025). Hybrid Cybersecurity for Asymmetric Threats: Intrusion Detection and SCADA System Protection Innovations. *Symmetry*, *17*(4), 616. https://doi.org/10.3390/sym17040616