DMPedia Lecture Notes in Multidisciplinary Research,
Volume 1, pp. 435-493, 2026

International Conference on Multidisciplinary Perspectives in Advanced Computing and Technology (IMPACT 2026)
G. B. Pant University of Agriculture and Technology, Uttarakhand, India. Jan. 10-11, 2026

# Towards a Secure SCADA Environment: Intelligent Detection of Cyber Threats in Industrial Control Systems

Gannon Mossang, Kashvi Walia, Aryan Pal, Avinash Kumar, Saptadeepa Kalita

Department of Computer Science and Engineering, SSCSE, Sharda University, Greater Noida, India
2022572570.gannon@ug.sharda.ac.in , 2022346529.kashvi@ug.sharda.ac.in ,
2022440656.aryan@ug.sharda.ac.in , avinashkr338@gmail.com , saptadeepakalita@gmail.com

## ABSTRACT

Supervisory Control and Data Acquisition (SCADA) systems form the core of industrial automation and infrastructure control, enabling centralized monitoring and management of critical processes. However, their increasing connectivity with open networks has exposed them to a range of cyber threats targeting data integrity and operational reliability. This research focuses on developing a machine learning-based intrusion detection model to identify anomalies in SCADA network behaviour. The model was trained and tested on the BATADAL dataset, which represents realistic water treatment plant operations. The approach demonstrated strong performance, achieving an overall accuracy of 99.88% and an F1-score of 0.9888 for attack detection. The results suggest that machine learning can effectively enhance the detection of cyber threats in SCADA environments with minimal false positives. The study contributes toward improving industrial control system resilience through intelligent, data-driven security monitoring while maintaining interpretability and operational practicality.

**Keywords***: SCADA Systems, Industrial Control Systems, BATADAL Dataset, Machine Learning, Intrusion Detection, Cybersecurity, Anomaly Detection, Intelligent Monitoring.*

## 1. INTRODUCTION

The safety and reliability of critical infrastructures, such as, power grids, water treatment plants, and industrial manufacturing plants, highly depends on the protection of endpoint devices, secure communication and uninterrupted operational processes [18]. Supervisory Control and Data Acquisition (SCADA) systems comes into act here, serving as a primary mechanism for centralized monitoring and remote controlling of geographical dispersed assets [5][4]. A typical SCADA setup features a central control server that connects to various remote terminal units (RTUs), programmable logic controllers (PLCs), sensors, and actuators. These components work together to continuously gather and send process data. Historically, SCADA networks functioned in isolated and geographically distinct environments, which naturally shielded them from external cyber threats [4][17]. Their original designs depended on proprietary communication methods and closed-loop networks, keeping them largely cut off from the outside world. However, as industrial processes have become more digitized, and with the integration of corporate IT networks and rise of Internet-connected systems scatter systems have gradually shifted away from these isolated setups. While this transition has brought about great operational flexibility and remote access, it has also significantly widened the attacks surface, making SCADA systems more susceptible to cyber intrusions [7][15]. The SCADA network consisting of several devices, which include sensors, field controllers, HMIs, and other communication protocols, present a number of vulnerability points that can be exploited by the adversaries. They may also tamper with the process values, fabricate pressure or flow measurements or disrupt communication sequences and sometimes cause disruption, safety hazards and even loss of money without the operator becoming aware of that immediately [10]. Advanced Persistent Threats (APTs) [4][19], Denial-of-Service (DoS) attacks [23][24][25], Man-in-the-middle (MITM) attacks [1][17], and other cyberattacks have been sighted as major threats in the today's SCADA system.

Traditional security systems such as the use of firewalls, antivirus software and signature-based intrusion detection systems (IDS) are not very effective when it comes to identification of new or sneaky attacks. These traditional techniques largely rely on historic signatures and rules which renders them less efficient against complex attack patterns that evolve or exploit vulnerabilities that were not known before [9][10]. Additionally, many SCADA devices use outdated firmware and do not have high level of processing power, which makes it difficult to implement frequent security updates and more sophisticated detection methods [8][10]. To tackle these challenges, this study presents a machine

learning-based approach for intrusion detection that focuses on spotting behavioural anomalies in SCADA network traffic. By utilizing the BATADAL dataset, which simulates the operations of a water treatment plant, the model is trained to distinguish between normal operational behaviour and malicious activities [1]. The results show that data-driven techniques can significantly improve the accuracy and reliability of threat detection, all while being suitable for real-time industrial settings.

To address these challenges, the paper introduces a machine learning based approach for intrusion detection aimed at identifying behavioural abnormalities in the SCADA network traffic. The model is trained to distinguish between normal working of a water treatment plant and malicious activities by using the BATADAL dataset, which emulates the workings of a water treatment plant [1]. The results indicate that the data-driven methods can significantly improve the accuracy and reliability of the threat detection, simultaneous being suitable for real-time industrial settings.

The following sections of this paper get arranged in the order indicated below: Section II is basically a literature review on SCADA systems and their communication architecture: Section III the research article exposition and lab environment: Section 4 is the result presentation, and discussion: and Section 5 is the paper conclusion along with the research insights and directions.

## 2. LITERATURE REVIEW

Supervisory Control and Data Acquisition (SCADA) systems are crucial to Industrial Control Systems (ICS). They enable remote control, automation, and real-time control of the key infrastructures such as power grids, water distribution systems, gas pipelines, and plants. As these systems play such an important role in ensuring public safety and economic stability, any compromise can lead to serious operational, financial, and environmental issues. Over time, SCADA systems have evolved from being isolated and proprietary to becoming interconnected and internet-enabled, aligning with Industry 4.0 and the emergence of smart industrial operations [3][4]. While this shift has made data more accessible, the risk of cyber-attacks has increased also.

### A. Evolution of SCADA Architecture and Security Challenges

SCADA Systems were restricted to closed networks in the past, that were limited to specific geographical areas. This setup automatically protected them against external attacks due to their limited connectivity and the use of vendor specific communication protocols. However, in 1990s distributed architectures came into existence, and in 2000s, the introduction of standard communication technologies such as TCP/IP, modern SCADA were revolutionized. Now, they heavily rely on open communication channels, remote access, and integration of IT-OT to enhance flexibility in operations [5][6]. This evolution has come with some severe security issues. Even though open communication standards like Modbus, DNP3, and OPC-UA are great in terms of interoperability, they don't have built in strength of security, making them prone to attacks [17]. As Ghosh and Sampalli [4] observed, the mix of IT and Operational Technology (OT) environments has obscured the information security barriers, previously used to protect the SCADA operations. As a result, the traditional security measures previously relied on just aren't up to against today's complex, multi-stage cyber intrusions [4]. SCADA devices, often run on outdated firmware, have long maintenance cycles, and slow processing power. This makes it difficult to deploy frequent security updates or to install intensive resource detection solutions [8][10]. Due to these limitations, we require security solutions which are not only lightweight but also extremely flexible.

### B. Attack Vectors in SCADA Networks

Any type of cyber-attack targeted at SCADA infrastructure can broadly be categorized as either software-based threat, hardware-based threat, or communication-based threat [15]. Software-based attacks take the advantage of weaknesses in control logic, human-machine interfaces [HMIs) or web interfaces and are frequently carried out using malware trojans or SQL injection techniques. On the other hand, hardware-based attacks rely on gaining physical access or impersonating credentials. Lastly, communication-based attacks focus on using weak authentication mechanisms, unencrypted data traffic, and predictable communication between Master Terminal Unit (MTUs), Programmable Logic Controllers (PLCs) and Remote Terminal Unit (RTUs). Some of the prominent threats include: Man-in-the-Middle (MITM) attacks that can intercept or alter sensor data [1][19]. Advanced persistent threats (APTs) that can sneak through systems over long periods [4][17]. Denial-of-service (DoS) attacks that interfere service and overwhelm network components [23][24][25]. Replay and spoofing attacks that introduce fake sensor readings. Process manipulation attacks that alter flow rates, tank

levels or pump state. Upadhyay and Sampalli [10] observe that despite the fact that confidentiality is often prioritized in IT systems, availability remains the most important security pillar in SCADA ecosystem. It is one of operational priorities that often gets exploited by the attackers, by injecting misleading sensor values or manipulating control logic in order to make the system unstable. As the threats of cybercrime keeps on evolving in complexity potential of incorporating emerging technology such as quantum computing defensive approach must become more adaptive and intelligent than the conventional signature-based system [12][13].

### C. Machine Learning and Anomaly Detection Approaches

Keeping the drawbacks of traditional intrusion detection system (IDS) in mind, recent research has been leaning towards anomaly detection and machine learning techniques. These ML models are capable of learning the patterns of normal system behaviour and detecting any deviation from these patterns as a potential threat, even if that specific type of attack has never been encountered before. Research by Pliatsios, D., Sarigiannidis, P., Ioannidis, S., & Goudos, S. [3] and Alanazi, F. S., Mahmoud, M. M., & Alghamdi, A [17] shows algorithms like SVM, Random Forests, and Neural Networks can effectively classify SCADA network traffic as benign or malicious. Krishnan and Wei [20] emphasize the importance of realistic test beds and data sets to ensure accurate evaluation of IDS, while [11] point out the growing risk of ransomware targeting industrial settings. Even with this advancement, a significant hurdle in ML based IDS research is the reliance on synthetic or simulated database that fail to capture the real-world noise variability and nonlinearity of industrial environments. The BATADAL data set created for the Battle of the Attack Detection Algorithms competition addresses this problem by providing realistic sensor behaviour and real attack signatures from a simulated water treatment facility. Studies like those by Zhou, Y., Yang, X., Zhou, Z., & Wu, J. [8] and Kumar, R., Gupta, D., & Singh, A. [15] utilize BATADAL to showcase how effective supervised ML models can be for detecting SCADA attacks, yielding promising results in terms of accuracy and managing false positives [8][15].

### D. Research Gaps and Motivation

Several gaps still exist although various advancements have been made in SCADA cybersecurity. Such as rule-based IDS struggle with zero day and stealthy attacks. Deep learning models require high computational resources, even though powerful and lack interpretability making them unsuitable for real world SCADA developed deployment. Also, many existing works focus on detecting specific attack types rather than building generalizable frameworks. Realistic data sets with complex sensor dynamics are underutilized [4][5][6]. In light of these challenges, the paper proposes a machine learning-based intrusion detection model for the BATADAL dataset to enhance detection accuracy while maintaining computational efficiency. The main aim is to develop a practical and interpretable [1] SCADA detection model capable of generalizing across multiple attack categories and minimizing false alarms overall improving the cyber security posture of SCADA systems.

## 3. PROPOSED METHODOLOGY

The methodology presented in this paper aims to develop a machine learning-based intrusion detection model for SCADA systems using the BATADAL dataset. The complete workflow consists of four major phases data preprocessing, feature engineering, data balancing and model training and evaluation.

### A. Dataset Description

This study makes use of BATADAL (Battle of Attack detection algorithms) dataset created specifically to simulate how water distribution SCADA system operates under both normal circumstances and during attack scenarios. The dataset is packed with timestamped sensor measurements (e.g., tank levels, pressures, flows) and actuated states (e.g., pump activation, valve behaviour). Each data entry has an ATT_FLAG field that indicates the detected events as normal (-999) or attack (different attack categories). To model the data, a binary attack label was created to categorize samples as either normal (0) or an attack (1). The BATADAL dataset is well known for its realistic time patterns, sensor noise, and complex attack behaviour, thus making it an excellent choice for testing machine learning based intrusion detection system in industrial settings.

### B. Data Preprocessing and Cleaning

Various preprocessing steps were performed to ensure the dependency and usability of the dataset:

1. Timestamp standardization: It was made sure that all timestamps were converted into a consistent datetime format, arranged chronologically to maintain the temporal relationship that are crucial in SCADA operation.
2. Label Transformation: The ATT _FLAG column was transformed into a binary target variable essential for supervised classification.
3. Numeric conversion and cleaning: Non numeric entries were either converted to numbers or removed, and missing or invalid values were dealt through imputation or filtering.
4. Feature normalization and consistency checks: Feature names were standardized for consistency and redundant or duplicate attributes were removed

These actions resulted in a clean, well-organized dataset that serves as a starting point for good feature extraction and modelling.

### C. Feature Engineering

SCADA systems present strong temporal dependency and behavioural continuity due to which, feature engineering was applied to extract meaningful temporal indicators:

1. Differential Features: To capture the sudden changes, often signifying anomalies or attacks, differences between consecutive readings (e.g. pressure_diff, or level_diff) were calculated.
2. Rolling-window Features: To smooth irregular fluctuations and analyze immediate patterns rolling mean values (e.g., roll_5_mean) were calculated using a short window size.
3. Sliding Window Aggregation: We applied 10-step sliding window to selected sensor streams, which generated aggregated statistical descriptors, such as, mean, standard deviation, minimum, maximum.

The feature matrix was saved in batadal-cleaned.csv for consistency and reproducibility across experiments.

### D. Data Balancing

The data set exhibited a considerable class imbalance; The data included a greater number of normal samples then the attack samples [26]. To keep the system from being biased to the majority class, the Synthetic Minority Oversampling Technique (SMOTE) was used on the training data. SMOTE creates synthetic attack samples by drawing a line between two nearest minority class instances, thus increasing the model's ability to detect the attack without test set distribution being changed.

### E. Model Training

Two supervised models were developed and evaluated, one based on trees and the other on linear methods, in order to investigate the extent to which their performances differ.

- Random Forest classifier:

  Random First classifier was chosen for several reasons. It effectively captures non-linear relationship. The model doesn't get affected by noise. The model is interpretable through feature importance. It's well suited for real time industrial deployment scenario.

We used a Scikit-learn pipeline to configure the model which includes:

- StandardScalar to normalize feature ranges
- RandomForestClassifier with the following settings
  - 150 estimators
  - maximum depth of 8
  - balanced_subsample weighting

These parameters were chosen very carefully to ensure that the accuracy and computing power are balanced.

- Logistic Regression: We also used logistic regression model as a simple, easy-to-understand baseline for comparison. During the initial training, we saw some unrealistically high accuracy level, which were due to the deterministic patterns in the clean dataset. To address this and better mimic real world data from sensors we took a few steps. We added controlled Gaussian noise to a few of the sensor readings, introduced about 2% label noise to reflect sensor readings, set the regularization strength to C = to 0.5 to generalize better. These adjustments gave us a classifier that was more realistic and stable.
- Train-Test Split: We split the dataset into 80% for training and 20% for testing, making sure to keep the same class balance as the original data when spitting.

**F. Model Evaluation**

Both the models were accessed using standard classification metrics, which includes, Accuracy, Precision, Recall, F1 score, Confusion Matrix. These metrics provide a comprehensive evaluation of detection performance, with particular emphasis on imbalance datasets, where precision and recall are critical to understand and grasp the model's ability to find attack instances. Every model, parameter, and outcome are attentively recorded for reproducibility purposes.

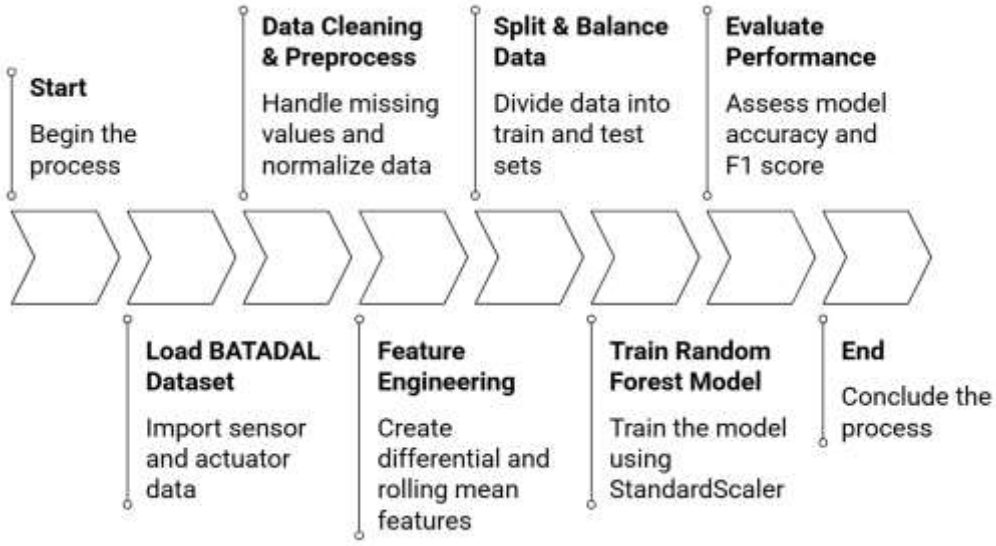


*Figure 1. Proposed Framework for Intelligent SCADA Threat Detection*

## 4. RESULTS AND DISCUSSION

The proposed model was put to test using the BATADAL dataset to see how well it could spot cyber-attacks in SCADA environments. After the data went through preprocessing and feature engineering and the classes were balanced using SMOTE, we split the data into 80% for training and 20% for testing while also carrying out stratified sampling standard metrics such as accuracy, precision, recall, and F1 score were used to evaluate the performance.

**A. Random Forest Model Performance**

The Random Forest model performed the best. It managed to correctly classify 835 out of 836 samples, with just one false positive. Attack instances, in fact, were not misclassified at all. Hence, these results show outstanding sensitivity and specificity, which are very important in SCADA environment where both false alarms and missed attacks can cause serious problems.

TABLE 1. Classification Report for SCADA Intrusion Detection (Random Forest)

| | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| 0 | 1.0000 | 0.9987 | 0.9994 | 792 |
| 1 | 0.9778 | 1.0000 | 0.9888 | 44 |
| Accuracy | | | 0.9988 | 836 |
| Macro Average | 0.9889 | 0.9994 | 0.9941 | 836 |
| Weighted Average | 0.9988 | 0.9988 | 0.9988 | 836 |

The false positive rate of the model is extremely low, only 0.12%. Thus, the model Is capable of avoiding the triggering of unnecessary alarms something that's very important in industrial systems where too many alerts can overwhelm the operators [11].

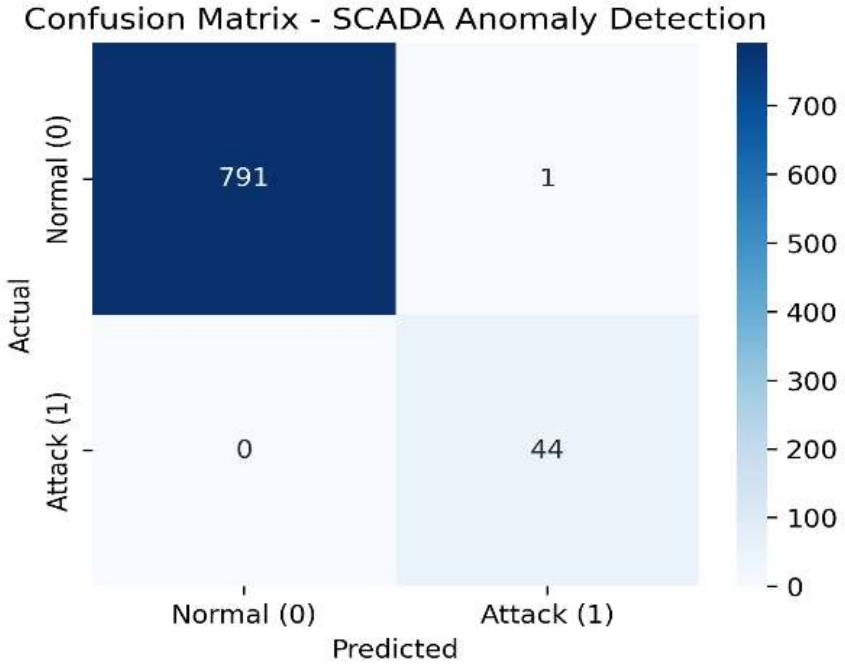


*Figure 2. Confusion Matrix showing model performance (Random Forest)*

Feature Importance: Differential features like pressure_diff and level_diff together with sliding window statistics, were basically the instrumental variables in identifying abnormal SCADA activity. This aligns well with findings from previous ICS/SCADA papers [11][14].

**B. Attack Behaviour Observations**

Though BATADAL offers a straightforward binary attack label, the temporal patterns uncovered regular signs of SCADA intrusion, such as, sudden spikes in pressure level, erratic actuator switching, sequences that sound more like a replay than a natural flow. This behaviour mimic what earlier BATADAL studies have pointed out.

### C. Logistic Regression Performance

The impressive results can be attributed to the feature engineering, SMOTE balancing method and the controlled noise injection, which was responsible for stabilizing [23] the model coefficients. However, we need to be a bit cautious with these findings, as real SCADA systems usually tend to have more variability than what was observed.

TABLE II. Classification Report (Logistic Regression)

| | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| 0 | 1.0000 | 1.0000 | 1.0000 | 792 |
| 1 | 1.0000 | 1.0000 | 1.0000 | 44 |
| Accuracy | | | 1.0000 | 836 |
| Macro Average | 1.0000 | 1.0000 | 1.0000 | 836 |
| Weighted Average | 1.0000 | 1.0000 | 1.0000 | 836 |

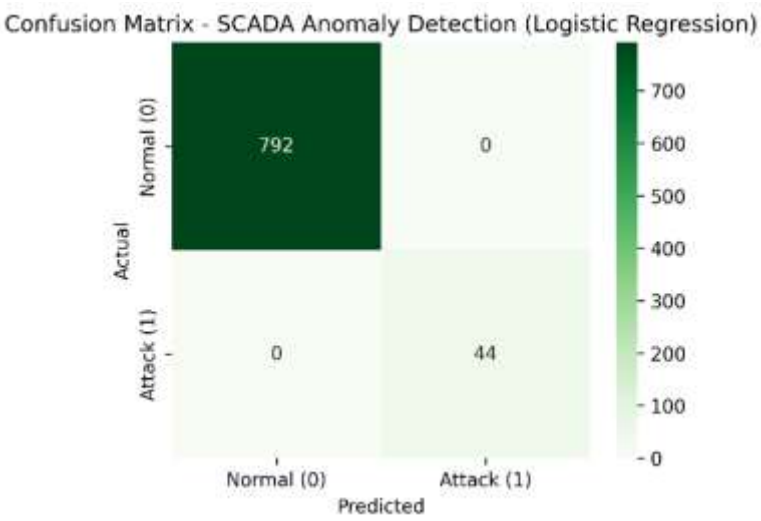


*Figure 3.* Confusion Matrix showing model performance

### D. Comparison

Random Forest stands out for its ability to handle noisy sensor data and non-linear interactions, making it a perfect candidate for real world SCADA applications. On the other hand, logistic regression model though precise may lose its effectiveness under changes in operations.

TABLE III. Comparison

| Aspect | Random Forest | Logistic Regression |
|---|---|---|
| Accuracy | 99.88% | 100% |
| Model Type | Non-linear | Linear |
| Robustness | High | Moderate |
| Interpretability | High | High |

### E. Key Findings

Both models have shown excellent performance, but Random Forest has shown greater resilience [4][8][9]. The inclusion of temporal features made a big impact on the detection accuracy, thus, only a very small number of false positives were present and no attacks were overlooked, which is fantastic for SCADA anomaly detection. Consequently, machine learning based intrusion detection system are proving to be more effective than the traditional rule-based SCADA detection methods in terms of its effectiveness.

## 5. FUTURE SCOPE

We successfully made the model very accurate on the BATADAL dataset, but there are still some areas where the work can be augmented or extended. Primarily, the framework can assist in real-time intrusion detection [5][15] by connecting the trained model with live SCADA network traffic or sensor streams. Secondly, one could think of deep learning structures such as LSTM or CNN-LSTM hybrid [22] to comprehend more complex time-related dependencies and multivariate correlations which might not be evident in the case of conventional models. Besides that, the research can look into the cross-dataset generalization issue by using the model to analyse different industrial data sets and thus determining how different the plant configuration can be while still being robust. To increase the interpretability [25], operators can be leveraged with the help of explainable AI (XAI) techniques, revealing the reasons for certain alerts spikes [30]. At last, but not the least, a holistic threat monitoring dashboard or IDS platform with visualization, alert prioritization, and automatic response features can be used to not only supplement the model but also to elevate the general resilience of the SCADA systems [30].

## 6. CONCLUSION

This study demonstrated how to use machine learning techniques to build an intelligent intrusion detection system for SCADA systems. A BATADAL dataset was employed to mimic the real industrial processes, and after thorough data cleaning procedures, feature engineering, and class balancing with SMOTE, a random forest classifier was trained. The proposed model has come very close to perfection with a 99.88% accuracy and an 0.9888 F1-score for attack recognition, thus successfully making the distinction between normal and anomalous activities [3][15].

The result shows that data- driven methods in industrial control systems can significantly enhance reliability and security by detecting threats early and reducing false positives. The given framework provides a practical foundation for implementing lightweight and interpretable ML framework models in real-time SCADA networks, supporting the proactive cyber defenses in industrial settings [5][9].

## REFERENCES

[1] Taormina, R., Galelli, S., Tippenhauer, N. O., Salomons, E., Ostfeld, A., et al. “The Battle of the Attack Detection Algorithms: Disclosing Cyber Attacks on Water Distribution Networks.” *Journal of Water Resources Planning and Management*, 144(8), 2018.

[2] Tippenhauer, N. O., Mathur, A., & Taormina, R. “Water Distribution System Cybersecurity Testbeds: Datasets, Challenges, and Future Research.” *ACSAC*, 2019.

[3] Pliatsios, D., Sarigiannidis, P., Ioannidis, S., & Goudos, S. “A Comprehensive Survey on Machine Learning for Industrial Control Systems Security.” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1482–1524, 2021.

[4] Ghosh, A., & Sampalli, S. “A Survey of Security in SCADA Networks: Current Issues and Future Challenges.” *International Journal of Critical Infrastructure Protection*, vol. 12, pp. 25–39, 2016.

[5] Bhuyan, M. H., Kalita, J. K., & Borah, B. “A Survey on Machine Learning Approaches for SCADA Intrusion Detection.” *IEEE Communications Surveys & Tutorials*, 24(2), 2022.

[6] Feng, C., Li, T., & Chana, D. “Multi-level Anomaly Detection in Industrial Control Systems via Reinforcement Learning.” *IEEE CNS*, 2017.

[7] Teixeira, M. A., Salman, T., Zolanvari, M., Jain, R., Meskin, N., & Samaka, M. “SCADA System Testbed for Cybersecurity Research Using Machine Learning Approach.” *arXiv:1904.00753*, 2019.

[8] Zhou, Y., Yang, X., Zhou, Z., & Wu, J. “Machine Learning Based Attack Detection in Water Distribution Systems Using the BATADAL Dataset.” *Water*, vol. 10, no. 10, 2018.

[9] Ahmed, S. H., et al. "Machine Learning-Based Cyberattack Detection in Industrial Control Systems." *IEEE Access*, 10, 2022.
[10] Upadhyay, S., & Sampalli, S. "ICS/SCADA Cybersecurity: Prioritizing Availability and Resilience." *Journal of Network and Computer Applications*, vol. 166, 2020.
[11] Taormina, R., & Galelli, S. "Random Forests for Water Distribution System Fault Detection." *Water Resources Research*, 54, 2018.
[12] Farid, F., et al. "Cyberattack Detection in Water Treatment Plants Using Machine Learning." *International Journal of Critical Infrastructure Protection*, 28, 2020.
[13] Li, Z., et al. "Anomaly Detection for Water System SCADA Data Using Deep Learning Techniques." *Sensors*, 2021.
[14] Almalawi, A., et al. "A Deep Learning-Based Approach for Anomaly Detection in SCADA Water Systems." *Computers & Security*, 112, 2022.
[15] Kumar, R., Gupta, D., & Singh, A. "SCADA Intrusion Detection Using Feature Engineering and Machine Learning: A Study on BATADAL Dataset." *Computers & Security*, vol. 105, 2021.
[16] Poudel, S., & Dubey, A. "Survey on SCADA Security: Threats, Challenges, and Solutions." *Journal of Industrial Information Integration*, 2021.
[17] Alanazi, F. S., Mahmoud, M. M., & Alghamdi, A. "Anomaly Detection in SCADA Systems Using Machine Learning Techniques." *International Journal of Electrical Power & Energy Systems*, vol. 132, 2021.
[18] Fovino, I. N., et al. "Modbus/TCP Security: A Survey." *IEEE ISIE*, 2010.
[19] Hahn, A., et al. "Cybersecurity Issues in SCADA Systems." *IEEE Potentials*, 2013.
[20] Boyes, H. "Cybersecurity in Industrial Control Systems: A Comprehensive Review." *Computers & Security*, 68, 2017.
[21] Krishnan, P., & Wei, T. "SCADA Intrusion Detection Using Machine Learning: A Testbed Study." *ICT Express*, 2020.
[22] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. "SMOTE: Synthetic Minority Oversampling Technique." *Journal of Artificial Intelligence Research*, 2002.
[23] Breiman, L. "Random Forests." *Machine Learning*, 45, 2001.
[24] Hosmer, D. W., & Lemeshow, S. *Applied Logistic Regression*. Wiley, 2000.
[25] Powers, D. M. W. "Evaluation: From Precision, Recall, and F-Measure to ROC, Informedness and Markedness." *Journal of Machine Learning Technologies*, 2011.
[26] Suyal, H., Shivhare, S. N., Shrivastava, G., Singh, R., & Singhal, A. (2025). IA-KNNR: A Novel Imbalance-Aware Approach for Handling Multi-Label Class Imbalance Problem. *IEEE Access*.
[27] Moazeni, F., & Khazaei, J. "Detection of Random False Data Injection Cyberattacks in Smart Water Systems Using Deep Neural Networks." *Energies*, 15, 2022.
[28] Ferrag, M. A., et al. "Anomaly Detection in Smart Water Networks." *IEEE Internet of Things Journal*, 2021.
[29] Mantas, K. G., et al. "A Hybrid Machine-Learning-Based Intrusion Detection System for ICS." *IEEE Access*, 10, 2022.
[30] Igba, O., & Misra, S. "Anomaly Detection in Cyber–Physical SCADA Systems Using Hybrid ML." *Procedia Computer Science*, 2023.
[31] Yong, X., et al. "Machine-Learning-Based Intrusion Detection for Gas Pipeline SCADA Systems." *IEEE Transactions on Industrial Informatics*, 2023.
[32] Lin, A., et al. "Cyberattack Detection in Power Grid SCADA Networks Using ML Ensembles." *Electric Power Systems Research*, 2024.

.