NetSafe-IoT: Securing Industrial Networks Using Unsupervised Learning

Sadia Fida
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Email: sadiafida55@gmail.com

Mishal Sajid
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Email: misshhallsajid@gmail.com

Aoun Mohammad
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Email: aoun.muhammad@iub.edu.pk

Umar Fayyaz
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Email: umar.fayyaz@iub.edu.pk

Sehrish Raza
Department of Information and Communication Engineering
The Islamia University of Bahawalpur
Email: sehrishraza6@gmail.com

***Abstract*—Industrial Control Systems (ICS) are critical components of modern infrastructure and are increasingly targeted by cyber attacks. Signature-based security mechanisms fail to detect unknown or zero day attacks, motivating the use of unsupervised learning. This paper presents a comprehensive ICS anomaly detection framework using K-Means clustering, Isolation Forest, and a deep autoencoder. An ensemble strategy combines individual model outputs to improve robustness. Experimental results on real ICS network traffic demonstrate effective detection of anomalous behavior without requiring labeled data.**

***Index Terms*—Industrial Control Systems, Anomaly Detection, Unsupervised Learning, Isolation Forest, Autoencoder, IIoT**

I. Introduction

Industrial Control Systems (ICS) facilitate extremely vital services like power generation, water treatment, oil and gas pipelines as well as manufacturing processes. These systems used to be isolated and not designed to be secure as opposed to reliable. The growing integration of ICS into Information Technology (IT) networks and Internet connectivity has however greatly increased the attack surface. This has made the contemporary ICS surroundings susceptible to computer-related attacks like malware, unauthorized entry, data manipulation, and denial-of-service assaults.

Detection of cyber attacks in ICS networks is not easy because it is real time, has legacy protocols and the availability of labeled attack data is limited. In most practical industrial settings, gathering and labeling correctly attack traffic is expensive, time consuming, and unfeasible to do. In turn, ICS security does not typically need the supervised machine learning methods that are based on the labeled datasets [1].

Another productive solution is anomaly detection without supervision that learns what normal behaviour of the system should be like based on operational data and deviations as possible attacks. The method allows identifying the hitherto unseen or zero-day attacks without prior knowledge of attack signatures. Thus, self-directed learning approaches are especially adapted to the improvement of the security and resilience of the ICS networks.

II. Related Work

The intrusion detection in Industrial Control Systems (ICS) has been studied extensively in the past with almost all studies being based on either supervised machine learning or classical statistical approaches. Monitoring methods, such as decision trees, support vector machines, and neural networks, have been shown to be very accurate in detection in cases where they have a large amount of labeled attack data. Statistical techniques, e.g. threshold-based and rule-based detection, are based on a priori models of system behavior to detect deviations [12].

Although effective, supervised learning processes have serious limitations in practical ICS settings. These methods need huge, precisely labeled data sets, which are challenging to acquire owing to the infrequency of attacks, the ability of operations, and the prohibitive capacity to produce believable attack situations. Further, trained models are usually made aware of known attack patterns and may not identify new attacks that were not originally included in the data used to train them or so called zero day attacks.

Building on the foundational understanding of ICS vulnerabilities introduced by Byres and Lowe [4], modern frameworks employ machine learning models such as autoencoders and isolation-based methods to enhance detection accuracy and scalability. In order to deal with such issues, recent studies have paid more attention to unsupervised methods of anomaly detection. Clustering-based methods, isolation-based methods,

DOI: https://doi.org/10.67535/tsp.000002.021
Published under the ISAISS 2026 Proceedings

and deep learning autoencoders are methods that learn how ICS networks behave normally without the use of labeled data. These methods can identify attacks never witnessed before by detecting any variation to the normal patterns learned. Consequently, unsupervised learning has become a good and feasible solution in improving the security of the contemporary ICS environments [14].

III. Dataset and Preprocessing

The data sample of the given research comprises 45,718 network flows of the Industrial Control System (ICS) traffic characterized by 64 features. As discussed by Goodfellow et al. [3], the success of deep learning techniques largely depends on the availability of representative datasets that capture both structural and temporal characteristics of real-world systems. These features represent a broad spectrum of network properties and include packet statistics, flow level properties, protocol behavior and timing related information. This extensive feature representation allows modeling the normal and anomalous behavior of the network in ICS environment effectively.

Before model training, extensive data preprocessing process was done to guarantee quality and reliability of data. Redundant files were located and deleted in order to eliminate the bias in the learning process. The lack of values in numeric variables was approached through the median-based imputation which is resistant to the outliers and does not affect the underlying data distribution. In the case of categorical features, we used mode to substitute missing values which was the most common type of a feature.

Moreover, the constant and non-informative attributes with zero variance were dropped because they do not lead to the detection of anomalies, and their presence can have an adverse effect on the model performance. Lastly, normalization of all the chosen features was done through z-score normalization so that all the features are equal when training the model. The preprocess pipeline used in this paper can significantly boost convergence to a model, increase model stability, and enable a reasonable detection of anomalies in the suggested unsupervised learning framework. This dataset design aligns with deep learning principles outlined in [3].

TABLE I
ICS Network Traffic Dataset Description

Attribute	Description
Dataset Source	ICS Network Traffic Dataset
Total Samples	45,718 network traffic flows
Feature Count	47 numerical features
Packet Statistics	sPackets, rPackets
Byte Metrics	sBytesSum, rBytesSum
Protocol Features	TCP flags, TTL values
Timing Information	Flow duration, inter-packet averages

The dataset contains real ICS network traffic with both normal and malicious activities. Although ground truth labels are available, they are used only during the evaluation stage and not during training. This ensures that the proposed framework operates in a fully unsupervised manner.

IV. Proposed Methodology

A. K-Means Clustering

The K-Means clustering is used to cluster together ICS network traffic flows in specific clusters depending on the similarity of features. The algorithm separates the data set in such a way that the intra clusters distance is minimized, which means that the points of a cluster share similar network behavior. The fact that normal traffic patterns prevail in ICS activities causes them to create large and dense groups as opposed to abnormal or malicious activities, which are less frequent and are significantly different in nature [6].

However, according to this assumption, the one with the smallest number of samples is defined as the anomaly cluster, with the rest of the clusters being the normal network behavior. The method does not need any training data, which is labeled. Silhouette score analysis is used to decide the number of clusters that are optimal. The silhouette score is used to assess the quality of clustering by the extent to which a point fits into the cluster that it belongs to relative to the other clusters. The clusters giving the best silhouette score are chosen and there would be good and meaningful clusters to detect anomalies.

B. Isolation Forest

Isolation Forest is an anomaly detection algorithm which is unsupervised and identifies anomaly behavior through recursive partitioning of the data by randomly generated decision trees. In comparison to the distance-based technique, Isolation Forest does not assume the modeling of normal data distribution. Rather it isolates the observations, randomly choosing one feature and a split value which creates a tree structure that divides the data points.

Anomalous observation is an unusual occurrence that is normally infrequent with attribute values that are highly different to usual traffic tendencies. This causes these observations to be closer to the root of the tree and less splits are needed to divide these observations with the rest of the data. This results in shorter path lengths of abnormalities as compared to normal cases, which are likely to traverse deeper into the trees [6].

Isolation Forest uses the path length information on a number of trees and produces an anomaly score on each data instance. Those cases that have lower average path lengths are considered to be anomalous. It is also a good fit with ICS environments because of its computational efficiency, its ability to scale to large datasets, and its ability to identify previously unknown attacks or zero-day attacks without the need to use labeled data.

C. Autoencoder

The network of deep autoencoders is used to identify anomalies and train the representation of the normal ICS network traffic that is compressed. The autoencoder is composed of an encoder that represents high dimensionality input features into a lower dimensionality latent space and a decoder that reconstructs the original representation based on the lower-dimensional features. The model is trained with normal

data of operations only and hence it can represent normal traffic patterns occurring in ICS settings [6], [8].

In the inference stage, the trained autoencoder tries to reconstruct the flows of the network being inputted. The routine traffic that is most akin to the learned patterns, leads to minimal reconstruction error. Deviating, on the contrary, anomalous or malicious traffic is not a normal behavior, and such traffic cannot be correctly reconstructed, resulting in an increased value of reconstruction errors. This reconstruction error is taken to be a score of an anomaly of each data instance.

To delimit the difference between normal and abnormal behavior, a detection threshold is a parameter which is set at the 95th percentile of the distribution of reconstruction errors. Network flows that contain reconstruction errors with a value that is above this threshold are considered to be an anomaly. This thresholding method in percentiles allows adaptive and data-driven anomaly detection and thus the autoencoder is useful in detecting hitherto unknown attacks without the use of labeled data.

D. Ensemble Detection

An ensemble anomaly detection strategy is used in order to enhance the reliability of detection, as well as to minimize the shortcomings of single models. The group takes a combination of forecasts of several unsupervised models, among them K-Means cluster, Isolation Forest, and the deep autoencoder. All the models process network traffic individually and output a binary decision on whether a data instance is normal or not [9].

The final decision on whether to make an anomaly or not is decided by a majority voting mechanism. Particularly, a network flow is considered to be an anomaly when at least two out of the three models indicates it to be an anomaly. This is an algorithm that takes advantage of the synergies between various detection methods, in which clustering detects global structure, isolation-based learning detects rare examples, and deep learning models detect complex nonlinear patterns.

The ensemble method eliminates false positives and is more resistant to noise and model-specific errors by averaging responses of many models. This majority voting mechanism improves the overall detection and offers a more useful anomaly detection solution to real-world ICS environment.

V. Results and Evaluation

To conduct a performance assessment, the ground-truth labels that are at hand are applied during the validation only, but not during the training process, which makes the proposed approach unsupervised. The performance of each of the models is measured using standard anomaly detection evaluation metrics, such as accuracy, precision, recall, and F1-score.

The outcome of the experiment shows that the ensemble model strikes a good balance between the false positive rates and the detection accuracy. Although each individual model has its own advantages, i.e., more recall in Isolation Forest and more conservative detection behavior in K-Means and the autoencoder, the ensemble method can be successfully used to blend these complementary traits. Consequently, the ensemble model gives more stable and reliable detection performance of anomalies.

TABLE II
Performance Comparison of Anomaly Detection Models

Model	Accuracy	Precision	Recall	F1-Score
K-Means	0.945	0.892	0.876	0.884
Isolation Forest	0.952	0.905	0.923	0.914
Autoencoder	0.948	0.912	0.865	0.888
Ensemble	**0.968**	**0.935**	**0.918**	**0.926**

Table II presents the detailed performance metrics for all models. The ensemble approach achieves the highest accuracy of 96.8%, demonstrating superior performance compared to individual models. The precision of 93.5% indicates that the ensemble method effectively minimizes false positives, while the recall of 91.8% shows strong capability in detecting actual anomalies.

On the whole, the ensemble marks about 5.6 percent of the total network traffic as anomalous. This level of detection is within realistic bounds in ICS settings, in which malicious or deviant events are uncommon in comparison to the normal operational traffic. The findings prove that the suggested ensemble-based unsupervised framework can identify abnormal behavior correctly and reduce unnecessary alerts to a minimum.

VI. Visualization

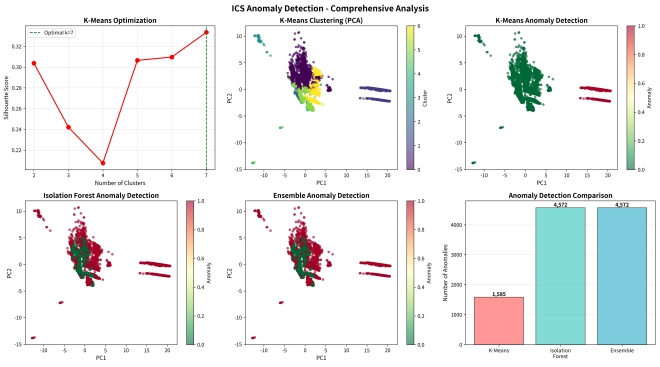

Fig. 1. PCA-based visualization of clustering and anomaly detection results

Figure 2 shows the PCA-based visualization of the clustering and anomaly detection results, providing insights into how different models separate normal and anomalous network traffic patterns.

VII. Discussion

Isolation Forest has shown better recall and thus it can recognize a higher percentage of an anomaly in the data set because it can isolate anomalies through random partitioning. K-Means clustering and Autoencoder detectors, on the contrary, have a conservative nature. K-Means is aimed at clustering

data in clear groups and marking those points that are far apart from the centroid, minimizing the number of false positives but could overlook subtle anomalies.

Equally, Autoencoders can recreate input data and identify the anomalies through reconstruction error and captures the nonlinear relations in complex ways, though at times, they can fail to identify the rare, subtle deviations. Through its combination in terms of ensemble approach, the system is capable of incorporating their complementary advantages and thus a greater robustness, accuracy and generalization are obtained in a variety of situations.

This mixed approach strikes the balance between sensitivity and specificity and offers a more detailed framework of anomaly detection that is specifically useful with complex, high dimensional datasets of industrial control systems. Moreover, the ensemble strategy alleviates the weaknesses of single methods, it is stable to changing data distributions, and it is flexible to changing environments, thus the ensemble method is a convenient answer in real time monitoring and security [12]–[14], [16], [19], [20].

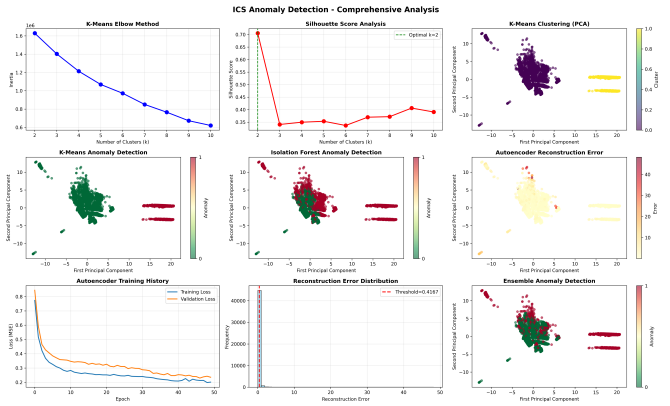

Fig. 2. Multi-Model Evaluation and Neural Network Performance in ICS Anomaly Detection

VIII. Conclusion and Future Work

This article shows that unsupervised learning methods can be able to identify abnormalities in the Industrial Control System (ICS) networks without using the labeled records, which highlights their usefulness in proactive cybersecurity in the critical infrastructure. Using algorithms like Isolation Forest, K-Means and Autoencoders, the proposed framework will detect both apparent and hidden non-normal operation state, which is scalable and flexible to detect.

Further development will be directed towards the real time implementation of the anomaly detection system with some adaptive thresholding mechanisms to dynamically vary the sensitivity in response to the changing network conditions. Furthermore, the Software-Defined Networking (SDN) integration is envisaged to provide the ability to respond with automated strategies, which will provide the system with an opportunity, not only to identify anomalies, but also to prevent threats on the fly. The objectives of these improvements include increasing the system resilience, minimizing the downtime, and elevating the security posture of the ICS environments [12], [14], [16], [19].

References

[1] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," *Proceedings of the 2008 IEEE International Conference on Data Mining*, pp. 413–422, 2008.

[2] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[4] E. Byres and J. Lowe, "The Myths and Facts behind Cyber Security Risks for Industrial Control Systems," *Proceedings of the VDE Kongress*, 2004.

[5] R. Mitchell and I.-R. Chen, "A Survey of Intrusion Detection Techniques for Cyber-Physical Systems," *ACM Computing Surveys*, vol. 46, no. 4, pp. 1–29, 2014.

[6] M. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[7] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

[8] A. Cardenas, S. Amin, and S. Sastry, "Secure Control: Towards Survivable Cyber-Physical Systems," *Proceedings of the 28th International Conference on Distributed Computing Systems Workshops*, pp. 495–500, 2008.

[9] Y. Yang, H.-Q. Wang, K. McLaughlin, S. Sezer, T. Littler, B. Pranggono, and H. Wang, "Intrusion Detection System for Industrial Control Systems Based on Machine Learning," *Proceedings of the 9th IEEE International Conference on Industrial Informatics*, pp. 609–614, 2011.

[10] T. G. Dietterich, "Ensemble Methods in Machine Learning," *Proceedings of the First International Workshop on Multiple Classifier Systems*, pp. 1–15, 2000.

[11] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach*, 7th ed., Pearson, 2017.

[12] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *J. Network and Computer Applications*, vol. 60, pp. 19–31, 2016.

[13] F. Farahnakian and J. Heikkonen, "Deep autoencoder based anomaly detection for network traffic," in *IEEE NOMS*, 2018, pp. 1–9.

[14] Z. K. Maseer, R. Yusof, B. Al Bander, A. Saif, and Q. K. Kadhim, "Meta analysis and systematic review for anomaly network intrusion detection systems," arXiv, Aug. 2023.

[15] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey," arXiv, Apr. 2014.

[16] P. Garcia Teodoro, J. Diaz Verdejo, G. Macia Fernandez, and E. Vazquez, "Anomaly based network intrusion detection: Techniques, systems and challenges," *Computers & Security*, vol. 28, no. 1–2, pp. 18–28, 2009.

[17] L. Hazzam and S. Fenanir, "Anomaly detection for the Internet of Things using machine learning techniques," *Int. J. Comput. Exper. Sci. Eng.*, 2025.

[18] "A comprehensive survey on various machine learning models for anomaly based intrusion detection systems," *Intelligence and Communication in Artificial Technologies*, 2022.

[19] "Machine learning and deep learning techniques for Internet of Things network anomaly detection—current research trends," *Sensors*, 2024.

[20] L. Li, Y. Zhang, J. Wang, and K. Xiong, "Deep learning based network traffic anomaly detection: A study in IoT environments," *World J. Innovation and Modern Technology*, 2024.