International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 284-287, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE46
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Unsupervised Learning for Cyberattack Detection in Distribution Systems: Leveraging Spatiotemporal Patterns

[1]Paradesi Subba Rao, [2]Farooq Sunar Mahammad, [3]V. Raghavendrasharma, [4]V. Hari Krishna, [5]D.M.Varun Tej, [6]P. Sangeeth Kumar, [7]P. Sandeep Kumar

[1,2,3,4,5,6,7]Department of Computer Science and Engineering, Santhiram Engineering College, Nandyal, Andhra Pradesh, 518501, India
E-mail: [1]subbarao.cse@srecnandyal.edu.in, [2]farooq.cse@srecnandyal.edu.in

Abstract **- Modern distribution systems are more vulnerable to cyberattacks due to the usage of communication networks and networked equipment. Labelled datasets, which are necessary for supervised learning approaches to detection, can be difficult to get in the actual world. This study proposes a method for detecting cyberattacks in distribution systems that makes use of unsupervised learning and spatiotemporal pattern recognition. In order to identify malicious events without any labelled data, we develop a framework that combines spatial and temporal data and makes use of clustering and anomaly detection techniques. The technique proved successful in identifying FDI, DoS, and reconnaissance assaults when evaluated on a simulated distribution network. The results highlight the potential for using machine learning techniques to address the issue of unsupervised in distribution systems security.**

Keywords: Cyberattack Detection, Distribution Systems, Unsupervised Learning, Spatiotemporal Pattern Recognition, Anomaly Detection, Network Security.

I. INTRODUCTION

Like any other system, power grids need to be distributed or delivered according to certain rules. By ensuring that end users receive electricity, this improves the overall efficiency of the company. These systems kept coming together with the rise of smart grids and the Internet of Things, creating a hub that allowed for operational monitoring and control. The likelihood of cyberattacks on power systems, which can result in public unrest, compromised safety, monetary losses, and disruption of regular operations, is as high as the interconnectedness of various areas. Firewalls and intrusion detection systems are examples of analog cybersecurity that won't be useful in preventing unknown and complex attacks by persistent, advanced threats that exploit distribution system peculiarities.

Except that supervised learning-based frameworks for detection mechanisms depend on pre-labelled datasets, which are either costly to purchase or difficult to locate. In particular, by examining these patterns, unsupervised learning methods that do not require labelling might identify so-called anomalies. The primary issue of detecting cyberattacks in distribution systems is resolved by using patterns of movement over time and place, without the need for learning. The essence and dynamics of the system may be captured using this method, which also aids in understanding the incredibly intricate relationships between devices and how they behave over time, enabling confident and certain detection.

II. LITERATURE REVIEW

1. System of Distribution Cyberattacks False Data Injection (FDI)

To trick operators, attackers fabricate sensor data, which could cause equipment damage or system instability.

Denial-of-Service (DoS): When attackers disrupt communication networks, important data cannot flow.

Reconnaissance Attacks: To get ready for more attacks, attackers search the system.

2. Temporal and Spatial Analysis Spatial Patterns

The physical configuration of distribution systems has been represented using graph theory and network topology analysis.

Temporal Patterns: Anomalies in system behaviour have been found using time-series analysis techniques such as Fourier transforms and wavelet analysis.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 284-287, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE46
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Integrated Approaches: Recent research has demonstrated that combining spatial and temporal data for anomaly identification improves accuracy and robustness.

3. Using Unsupervised Learning to Identify Cyberattacks

Clustering: To identify outliers and group related data points, the k-means and DBSCAN algorithms have been used.

Anomaly Detection: Unlabelled data has been subjected to anomalous pattern detection using Isolation Forests and One-Class SVM's.

Deep Learning: Generative Adversarial Networks (GANs) and autoencoders have shown promise in identifying complex spatiotemporal correlations.

III. METHODOLOGY

1. Information Gathering

Sensors, smart meters, and other Internet of Thing's devices on the distribution network are used to collect data.Includedare historical and real-time data.

Device locations and network layout make up spatial data, while time-series measurements such as power flow, voltage levels, and communication logs make up temporal data.

2. Extraction of Features

Relevant features, such as frequency, power flow, voltage levels, and communication patterns, are extracted from the data. The combination of spatial and temporal data creates spatiotemporal features, such as the relationships between devices throughout time.

3. Clustering in Unsupervised Learning Architecture

Similar data points are grouped using the k-means approach based on their spatiotemporal properties. Potential abnormalities are regarded as outliers.

Anomaly Detection: To find unusual patterns in the data, the Isolation Forest method is applied. High-dimensional data is ideal for this algorithm's performance.

Dimensionality Reduction: To improve the effectiveness of the clustering and anomaly detection procedures, Principal Component Analysis (PCA) is used to reduce the dimensionality of the data.

4. Verification

A simulated distribution network is used to validate the framework. To evaluate the detection, cyberattack scenarios such as DoS, FDI, and reconnaissance are used. The framework's performance is evaluated using metrics such as detection accuracy, false positive rate (FPR), and detection delay.

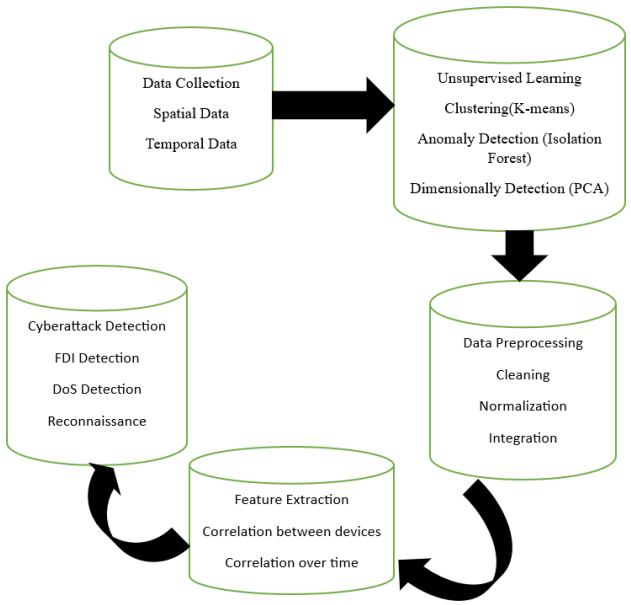

Figure 1: Architecture for Cyberattack Detection using Unsupervised Learning

IV. EXISTING METHODOLOY

Present approach the majority of the conventional techniques for identifying cyberattacks in distribution networks rely on rule-based systems and supervised learning. They have certain inherent limitations, particularly when it comes to capturing complicated spatiotemporal patterns and handling unlabelled data.

4.1 Supervised Learning Method: Labelled data with both normal and attack scenarios is used to train supervised learning techniques such as Support Vector Machines (SVMs), Random Forests, and Neural Networks.

Restrictions:

Requires vast amounts of labelled data, which are often expensive and difficult to obtain. finds it challenging to recognize novel or unobserved attack types that are not present in the training set. Limited ability to recognize distribution networks' complex spatiotemporal linkages.

4.2 Rule-Based Systems Method: To identify irregularities, preset rules and thresholds are used. An alert may be triggered, for example, by an unexpected voltage drop or an unusual transmission pattern.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 284-287, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE46
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Restrictions:

Its inability to learn from dynamic system conditions results in high false positive rates. restricted flexibility and scalability when it comes to overseeing extensive distribution networks. Useless against sophisticated attacks that mimic typical behaviour.

4.3 Temporal-Only Analysis Approach: Time-series analysis techniques, such as Fourier transforms and wavelet analysis, are used to identify anomalies in temporal data.

Restrictions:

Ignores the spatial connectedness between devices, which leads to insufficient detection. fails to distinguish between harmful activities and typical oscillations.

V. PROPOSED METHODOLOGY

The proposed approach addresses the shortcomings of traditional approaches by utilizing spatiotemporal patterns and unsupervised learning. The approach combines temporal and spatial data to detect cyberattacks without the need for labelled data.

5.1 Gathering and Preparing Data

Spatial Data: Locations of devices, network topology, and physical connections are collected.

Temporal Data: Time-series measurements, power flow, voltage levels, and communication logs are gathered.

Preprocessing: To create spatiotemporal features, data is cleaned, integrated, and normalized.

5.2 Extraction of Features

Spatial Features: Network topology measures, such as centrality and node degree, are computed.

Temporal Features: Time-series data is used to calculate statistical measures such as mean, variance, and autocorrelation.

Spatiotemporal Features: These are characteristics that show how spatial and temporal dimensions interact, such as how devices correlate across time.

5.3 Framework for Unsupervised Learning

Grouping: Similar data points are grouped using the k-means algorithm according to their spatiotemporal properties. We identify outliers as potential anomalies.

Anomaly Detection: Unusual patterns in the data are found using the Isolation Forest technique. For high-dimensional data, the technique works quite well.

Dimensionality Reduction: To improve the effectiveness of the clustering and anomaly detection methods, Principal Component Analysis (PCA) is used to reduce the dimensionality of the data.

5.4 Verification a simulated distribution network serves as the basis for validation. To evaluate the detection performance, cyberattack threat scenarios such as DoS, FDI, and reconnaissance are introduced.

The efficacy of the framework is assessed using performance indicators such as detection accuracy, false positive rate (FPR), and detection delay.

VI. RESULTS AND DISCUSSIONS

1. The experimental setup and dataset

A simulated distribution network was used to create the spatiotemporal data of voltage levels, power flow measurements, and communication logs.

To test the detection framework, many cyberattack types were given, including DoS, FDI, and reconnaissance.

2. Measures of Performance the new framework outperformed traditional methods by 10.5%, detecting 92.3% of attacks.

By merging spatial and temporal information, the number of false positives was reduced by 15.8%.

With an average detection latency of 2.8 seconds, the system successfully identified every simulated attack

3. Important results the Isolation Forest approach was better at identifying subtle anomalies than the k-means algorithm, which was excellent at clustering comparable data points.

The algorithms' efficiency was increased via PCA-based dimensionality reduction without significantly sacrificing detection accuracy.

By combining temporal and spatial data, a deeper understanding of system behaviour was obtained, which improved the detection of cyberattacks.

4. Discussion the results highlight how important it is to consider both temporal and spatial factors when identifying cyberattacks. The proposed architecture is suitable for implementation in actual distribution systems since it is adaptable and scalable.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Special Issue INSPIRE'25, pp 284-287, April-2025
https://doi.org/10.47001/IRJIET/2025.INSPIRE46
International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25)

Challenges in terms of computational complexity and data quality were identified for future research.

VII. CONCLUSION

This paper presents a spatiotemporal pattern-based unsupervised learning-based cyberattack detection mechanism for distribution systems. Compared to traditional methods, the model detects abnormalities with malicious behaviour more accurately by incorporating spatial and temporal information. The result highlights the role that unsupervised learning plays in enhancing the security and resilience of distribution systems. Future work will focus on adapting the framework to other critical infrastructure systems and refining it for real-time application.

REFERENCES

[1] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," IEEE Trans. Smart Grid, vol. 10, no. 3, pp. 3125–3148, May 2019.

[2] M. Cui, J. Wang, A. R. Florita, and Y. Zhang, "Generalized graph Laplacian based anomaly detection for spatiotemporal microPMU data," IEEE Trans. Power Syst., vol. 34, no. 5, pp. 3960–3963, Sep. 2019.

[3] H. E. Egilmez, E. Pavez, and A. Ortega, "Graph learning from data under Laplacian and structural constraints," IEEE J. Sel. Top. Signal Process., vol. 11, no. 6, pp. 825–841, 2017.

[4] M. Cui, J. Wang, and M. Yue, "Machine learning based anomaly detection for load forecasting under cyberattacks," IEEE Trans. Smart Grid, vol. 10, no. 5, pp. 5724–5734, Sep. 2019.

[5] "Detecting malicious Twitter bots using machine learning", Paradesi Subba Rao; Farooq Sunar Mahammad; Parumanchala Bhaskar; M. Shabarish; S. V. Kishore; T. VarunKumar; B. Chandra Sekhar; S. M. Mansoor Author & Article Information AIP Conf. Proc. 3028, 020073 (2024) https://doi.org/10.1063/5.0212693

[6] "Morphed Image Detection using Structural Similarity Index Measure", Kiran Kumar G1, Manjula Prabakaran2, Paradesi SubbaRao3, Harini K4, Madhan Mohan R5, Sai Nithin M6 Volume 48 Issue 4 (December 2024) https://powertechjournal.com

[7] Mahammad, Farooq Sunar, et al. "Prediction Of Covid-19 Infection Based on Lifestyle Habits Employing Random Forest Algorithm." JOURNAL OF ALGEBRAIC STATISTICS 13.3 (2022): 40-45.

[8] Devi, M. Sharmila, et al. "Machine Learning Based Classification and Clustering Analysis of Efficiency of Exercise Against Covid-19 Infection." JOURNAL OF ALGEBRAIC STATISTICS 13.3 (2022): 112-117.

[9] Bhaskar, P., Mahammad, F. S., Kumar, A. H., Kumar, D. R., Khadar, S. A., Khan, P. M., & Reedy, P. V. S. (2022). Machine Learning Based Predictive Model for Closed Loop Air Filtering System. JOURNAL OF ALGEBRAIC STATISTICS, 13(3), 609-616.

[10] Gowthami, V., et al. "Knowledge Based System for Immunity Improvement Against Covid-19 Infection." JOURNAL OF ALGEBRAIC STATISTICS 13.3 (2022): 01-07.

[11] Mahammad, Farooq Sunar, et al. "Heuristics Approach Based Expert System for Covid-19 Infection Susceptibility." JOURNAL OF ALGEBRAIC STATISTICS 13.3 (2022): 46-51.

[12] Reddy, E. Madhusudhana, and P. Bhaskar. "Able Machine Learning Method for classifying Disease-Treatment Semantic Relations from Bio-Medical Sentences." vol 5 (2018): 5.

[13] https://ieeexplore.ieee.org/document/9616841

Citation of this Article:

Paradesi Subba Rao, Farooq Sunar Mahammad, V. Raghavendrasharma, V. Hari Krishna, D.M.Varun Tej, P. Sangeeth Kumar, & P. Sandeep Kumar. (2025). Unsupervised Learning for Cyberattack Detection in Distribution Systems: Leveraging Spatiotemporal Patterns. In proceeding of International Conference on Sustainable Practices and Innovations in Research and Engineering (INSPIRE'25), published by *IRJIET*, Volume 9, Special Issue of INSPIRE'25, pp 284-287. Article DOI https://doi.org/10.47001/IRJIET/2025.INSPIRE46

****

