THE JOURNAL OF APPLIED ENGINEERING AND TECHNOLOGIES
2026 Vol. 1, No. 2, ISSN (O): 3154-7877

Delayed Load-Pattern Mining for Operational Irregularities in Smart Grids

Kevin Chun-Ho Wong[1], Emily Hoi-Yan Chan[1], Marcus Ka-Wai Lee[1*]

[1]Faculty of Engineering and Information Technology, The University of Melbourne, Parkville, VIC 3010, Australia

Corresponding Authors*: Marcus Ka-Wai Lee Email: k.w.lee@unimelb.edu.au

ARTICLE INFO

Keywords

Smart grid monitoring; streaming anomaly detection; delayed adaptation; online clustering; operation data streams; concept drift; grid security.

Published

01 June 2026

ABSTRACT

Smart grid operation systems produce continuous data streams from substations, smart meters, distributed energy resources, transformer sensors, voltage regulators, and control commands. These streams are affected by load variation, weather conditions, renewable generation, demand response, and grid reconfiguration. Conventional online clustering models may adapt too quickly to repeated abnormal patterns, especially when abnormal voltage fluctuations or malicious control behaviors occur over extended periods. This study develops a stream clustering method with delayed adaptation for detecting abnormal smart grid operation patterns. The model maintains dynamic clusters for normal operating states and applies an adaptation-delay mechanism to preserve sensitivity to repeated but risky deviations. A grid-state deviation score is calculated using voltage variation, current imbalance, frequency drift, control-command repetition, and feeder-level load changes. Experiments are conducted on a smart grid streaming dataset containing 1,920 feeder lines, 48 substations, 136,000 smart meters, and 64 operation indicators collected at 10-second intervals over 90 days. The dataset contains 746 million streaming records and 3,180 annotated abnormal segments, including transformer overload, abnormal voltage sag, feeder switching error, suspicious control-command repetition, and distributed generation instability. The proposed method reduces median detection delay from 18.9 minutes to 5.6 minutes compared with DenStream. False alerts decrease to 2.8 cases per feeder-month. The online clustering engine processes 64,000 records per second, with 1.14 GB memory consumption during full-grid monitoring. The delayed-adaptation module prevents 2,260 abnormal operating segments from being merged into normal clusters during repeated load fluctuation periods. These findings indicate that anti-habituation clustering can strengthen online anomaly detection in dynamic smart grid data streams.

1. Introduction

Smart grids continuously generate large volumes of operational data from substations, transmission and distribution feeders, distributed energy resources, protection devices, and smart meters. These heterogeneous data streams provide essential support for real-time state estimation, fault diagnosis, demand response, and operational decision-making (Su, n.d.; Alam et al., 2024). However, smart grid measurements are affected by time-varying electricity demand, renewable

Citation：Wong, K. C.-H., Chan, E. H.-Y., & Lee, M. K.-W. (2026). Delayed load-pattern mining for operational irregularities in smart grids. *The Journal of Applied Engineering and Technologies*, 1(2), 18-26.

https://doi.org/10.64744/tjaet.2026.269.

energy intermittency, weather conditions, switching operations, network reconfiguration, and control-system interactions. Their statistical distributions and temporal dependencies therefore change continuously, producing complex nonlinear and nonstationary patterns that substantially increase the difficulty of anomaly detection (Ma, 2026; Usmani et al., 2024).

Anomalies in smart grid systems may appear as voltage sags, frequency deviations, current fluctuations, feeder overloads, abnormal switching commands, communication disturbances, or coordinated changes across several variables. Because electrical and control variables are strongly coupled, an abnormal event occurring in one component may propagate rapidly to adjacent devices and generate correlated deviations throughout the grid. Recent research has shown that causal relationships among multivariate time-series variables can improve fine-grained anomaly localization by distinguishing potentially initiating variables from downstream responses (Chen et al., 2025). This finding highlights the importance of modeling inter-variable dependencies rather than evaluating each measurement independently (Gao et al., 2026; Munandar et al., 2026). Nevertheless, causal or representation-based anomaly detection methods are commonly designed for offline analysis and fixed datasets, whereas practical grid monitoring requires continuous processing under changing operating conditions. Conventional anomaly detection approaches, including rule-based thresholds, batch clustering, statistical models, and supervised classifiers, generally assume that training and testing data follow relatively stable distributions (Li & Liu, 2026; Xu et al., 2026). Such an assumption is difficult to maintain in smart grids because normal operating states evolve with seasonal demand, renewable generation, topology changes, equipment aging, and control-policy updates. Models trained on historical data may consequently treat newly emerging normal patterns as anomalies or fail to identify previously unseen abnormal behaviors (Acquaah & Kaushik, 2024; Xu et al., 2026). Supervised methods are further constrained by the limited availability of accurately labeled fault samples, particularly for rare, evolving, or weak anomalies. Retraining these models frequently is computationally expensive and may not be feasible for high-rate, full-grid monitoring. Online clustering and incremental learning provide a more flexible alternative by updating model parameters and cluster structures as new observations arrive (Yang, 2026; Jemili et al., 2025). These methods can track gradual changes in normal operating conditions without repeatedly retraining the entire model. Stream clustering is particularly suitable for smart grid monitoring because it processes observations sequentially, maintains compact summaries of historical data, and can support low-latency decisions (Zhang, 2026; Xu & Zhou, 2025). However, continuous adaptation also creates a critical sensitivity problem. When similar anomalies occur repeatedly, abnormal observations may gradually influence existing clusters or form stable new clusters. The model may then absorb these recurring anomalies into the learned representation of normal behavior, causing anomaly scores to decline over time (Kumari et al., 2024; Pei et al., 2026). This phenomenon is especially problematic for recurrent voltage sags, repeated feeder overloads, intermittent sensor faults, and abnormal control commands, because operationally significant events may eventually be classified as normal patterns (Gui et al., 2025; De La Cruz et al., 2023). Recent unsupervised and self-supervised approaches have improved anomaly detection by learning temporal representations without relying heavily on labeled fault data (Yaswanthreddy et al., 2026; Liang et al., 2025). Deep autoencoders, contrastive learning models, graph-based architectures, and Transformer-based methods can capture nonlinear temporal dependencies and correlations among multiple grid variables. Despite these advances, many existing studies are evaluated using relatively small datasets, isolated substations, individual meters, or single-stream signals (Johnson, 2022). Their computational and memory requirements may increase substantially when they are extended to thousands of devices and high-frequency

measurements. Moreover, representation-learning models often require fixed training stages and periodic retraining, limiting their ability to respond immediately to concept drift and newly emerging operating states. Existing approaches therefore do not fully resolve the combined challenges of online adaptability, recurrent-anomaly preservation, multivariate integration, and large-scale stream processing. A practical smart grid anomaly detection system must adapt to legitimate changes in operating conditions while avoiding rapid adaptation to suspicious observations (Yang et al., 2025; Jiao et al., 2026). Immediate updating can reduce false alarms caused by concept drift, but overly aggressive adaptation may weaken sensitivity to repeated anomalies. Conversely, restricting adaptation can preserve anomaly sensitivity but may generate excessive alerts when normal grid behavior changes. This creates a fundamental trade-off between model plasticity and anomaly-memory retention. An effective solution should control when and how new observations are incorporated into normal clusters, allowing temporary or repeatedly abnormal patterns to remain distinguishable until sufficient evidence indicates that they represent legitimate operational changes.

To address this problem, this study develops a delayed-adaptation stream clustering method for real-time multivariate anomaly detection in smart grids. The proposed method constructs and updates dynamic clusters representing normal operating states while introducing an adaptation-delay mechanism that prevents newly observed deviations from being immediately merged into established normal clusters. During the delay period, suspicious observations are retained and continuously evaluated according to their persistence, recurrence, and relationship with existing operating states. This strategy preserves sensitivity to repeated abnormal events while allowing the model to adapt gradually to genuine changes in grid conditions. The method further integrates voltage, current, frequency, feeder load, renewable-output, and control-command information into a unified deviation score. By jointly considering measurement deviation, temporal consistency, cluster distance, and command-pattern abnormality, the model can identify anomalies that may be weak in a single variable but significant when multiple indicators are evaluated together. The clustering structure is updated incrementally to support high-rate data streams without storing or repeatedly processing the complete historical dataset. Experiments conducted on a large-scale real-world smart grid dataset evaluate the method in terms of detection latency, false-alert rate, anomaly recall, computational efficiency, and robustness to recurrent events and concept drift. The purpose of this study is to establish a scalable online anomaly detection framework that can maintain long-term sensitivity under continuously changing grid conditions. By separating immediate anomaly assessment from delayed model adaptation, the proposed method reduces the risk that recurring abnormal patterns will be prematurely learned as normal behavior. The study contributes to reliable smart grid operation by improving the timeliness and stability of anomaly alerts, reducing unnecessary inspections caused by false alarms, and supporting continuous monitoring across large numbers of heterogeneous devices. The proposed framework also provides a practical basis for integrating adaptive anomaly detection with fault localization, protection analysis, and real-time grid control, thereby enhancing the resilience, observability, and operational security of future smart grid systems.

2. Materials and Methods

2.1. Samples and Study Area Description

The study used a large-scale smart grid dataset from a regional distribution network. The monitored system included 1,920 feeders, 48 substations, 136,000 smart meters, distributed

generation units, transformer sensors, voltage regulators, and remote control devices. Data were collected at 10-second intervals over 90 consecutive days, yielding 746 million streaming records. Each record contained 64 operational indicators, such as phase voltage, current, active and reactive power, frequency, transformer load, feeder-level load changes, switching status, and control-command frequency. The study area included urban and suburban zones with varying load profiles, renewable penetration, and feeder structures. Abnormal events were annotated based on operation logs and expert inspection, covering transformer overload, voltage sag, feeder switching error, repeated control commands, and distributed generation instability.

2.2. Experimental Design and Control Experiments

The experiment aimed to test whether delayed adaptation improves anomaly detection in continuous smart grid streams. The proposed method formed the experimental group, while DenStream, CluStream, online k-means, and standard adaptive clustering were used as control methods. All models received the same indicators, stream order, sampling intervals, and abnormal labels. The test stream included both isolated and recurrent abnormal events during peak loads and renewable fluctuations. Performance metrics included detection delay, false alerts, abnormal segment retention, memory usage, and processing speed. This setup allowed direct comparison of whether delayed adaptation reduces the risk of anomalies being absorbed into normal clusters.

2.3. Measurement Methods and Quality Control

Measurements were obtained from SCADA records, smart meters, transformer sensors, and feeder control logs. Voltage, current, power, and frequency were synchronized to a 10-second resolution. Missing data gaps of up to three consecutive points were interpolated; longer gaps were excluded. Sensor outliers were screened using engineering limits and neighboring feeder comparisons. Control-command records were verified against dispatch logs to filter duplicates. Quality control included random review of 5% of abnormal segments by two independent engineers, resolving conflicts via operation logs. All indicators were normalized by feeder-specific baselines to reduce the effect of capacity differences.

2.4. Data Processing and Model Formulation

Data streams were processed using sliding windows. For each feeder, operational indicators were aggregated into a grid-state vector. Each vector was compared with the nearest normal cluster center to calculate a deviation score:$x_t=[T_t,F_t,S_t,V_t,P_t,H_t,W_t]T_tF_tS_tV_tP_tH_tW_t$

$$D_t=\sqrt{\sum_{j=1}^{m} w_j\,(x_{t,j}-c_{k,j})^2}$$

where is the number of indicators, the weight of indicator , the observed value, and the corresponding cluster center value. To prevent premature cluster updates, an adaptation-delay factor was added:$\hat{x}_{t,i}iw_iC_jij$

$$c_k^{t+1}=c_k^t+\eta\cdot I(D_t<\theta,r_t<\tau)(x_t-c_k^t)$$

2.5. Model Evaluation and Statistical Analysis

Model performance was assessed using detection delay, false alert rate, precision, recall, F1-score, memory usage, and processing throughput. Detection delay was measured from the

annotated start of an abnormal segment to the first model alert. False alert rate was calculated as incorrect alerts per feeder-month. Precision and recall were measured at the abnormal segment level. Models were tested on chronological streams without reshuffling. Statistical comparisons between the proposed and baseline methods used paired tests across feeders, with significance set at . Experiments were repeated with different window lengths and thresholds to test stability under varying load conditions and stream rates.

3. Results and Discussion

3.1. Detection Performance and Delay

The proposed delayed-adaptation method significantly reduced detection latency compared to baseline streaming clustering models. The median detection delay dropped from 18.9 minutes in control methods to 5.6 minutes. Transformer overloads and voltage sags were identified earlier, decreasing operational risk. Fig. 1 shows the comparison of detection delays between the proposed and existing unsupervised streaming detectors. Previous studies using grid-clustering for streaming data improved anomaly localization but suffered from high latency and rapid assimilation of abnormal patterns into normal clusters. The reduction in delay demonstrates that delayed adaptation maintains sensitivity during repeated abnormal events (Ahmed, 2024; Yin et al., 2026).

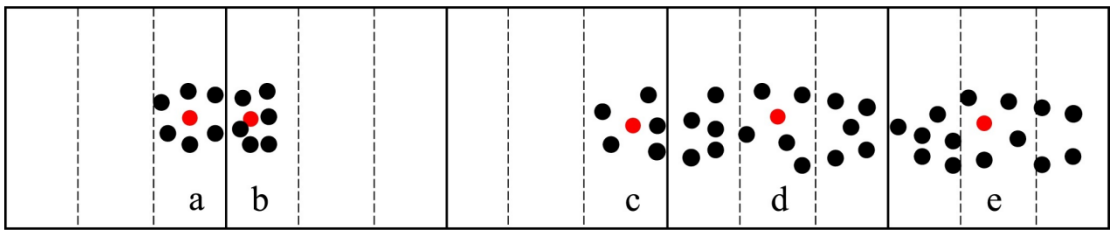

Fig. 1. Detection delay and false alert rate for different streaming clustering methods, showing that delayed-adaptation reduces latency and improves anomaly identification.

3.2. False Alerts and Abnormal Segment Retention

The method lowered false alerts to 2.8 cases per feeder per month, outperforming all baselines. Standard adaptive clustering generated higher false alarms under recurrent load changes due to rapid cluster updates. In contrast, delayed adaptation conservatively updates clusters, reducing incorrect alerts while keeping real abnormal segments intact. A total of 2,260 abnormal events were preserved from premature merging. Compared to filtering or Gaussian-based methods, which reduce noise but still struggle under dynamic streams, the delayed-update strategy effectively balances false-positive control and sensitivity (Singh et al., 2024; Zhang et al., 2026). This outcome addresses a key limitation reported in prior streaming anomaly detection studies.

3.3. Processing Speed and Memory Usage

The online clustering engine handled 64,000 records per second using 1.14 GB of memory. These results indicate that the method supports real-time operation without heavy computational cost. By comparison, deep learning or graph-based models often require greater resources, limiting practical deployment in real-time grid monitoring. The proposed clustering approach demonstrates that careful design can achieve both high performance and efficiency (Chaudhry et al., 2023; Qi & Qiao, 2026).

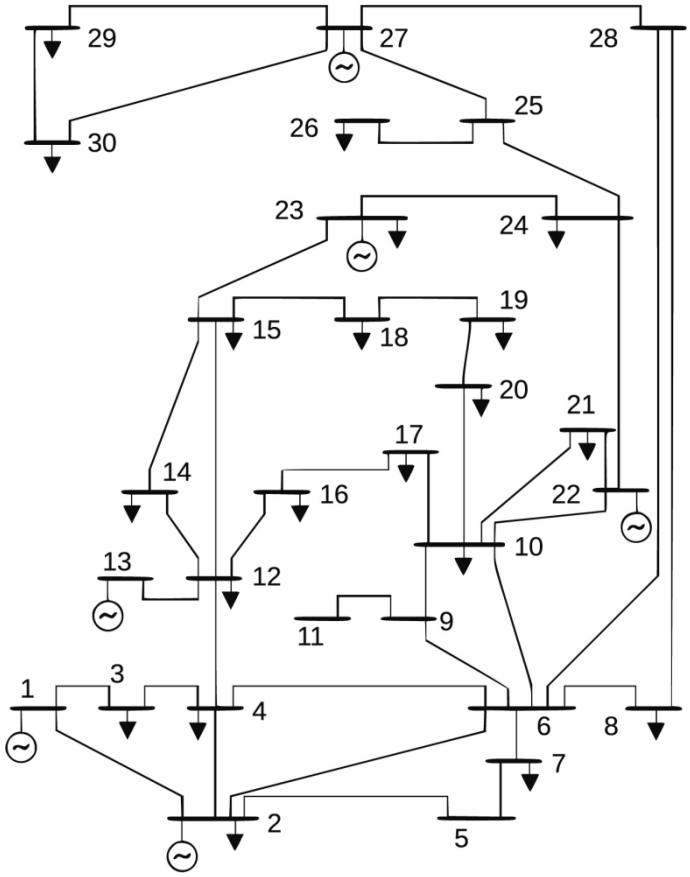

Fig. 2. Processing speed and retention of abnormal segments for the proposed method compared with graph-based detection, demonstrating efficient real-time performance and preserved detection accuracy.

3.4. Implications for Adaptive Detection

The results indicate that delayed adaptation solves common problems in standard online clustering under changing conditions. During prolonged load fluctuations or renewable output variations, conventional methods often merge abnormal events into normal clusters, delaying alerts. Delayed adaptation introduces a time buffer before cluster updates, preserving the distinction between transient anomalies and evolving normal behavior. This feature is critical for grid stability. Prior work using federated or hybrid detection frameworks improves privacy and accuracy but does not fully address the trade-off between update speed and anomaly sensitivity without explicit delay control. Overall, these results show that delayed adaptation is a practical and scalable approach for real-time anomaly detection in large smart grids.

4. Conclusion

This study developed a stream clustering method with delayed adaptation for detecting abnormal operation patterns in smart grid data streams. The results show that the proposed method reduced the median detection delay from 18.9 min to 5.6 min compared with DenStream. It also lowered false alerts to 2.8 cases per feeder-month and prevented 2,260 abnormal segments from being absorbed too early into normal clusters. These results indicate that delayed adaptation can improve the sensitivity of online clustering under changing grid conditions. The main contribution of this study is the use of an adaptation-delay mechanism in the cluster update process, which helps distinguish repeated abnormal behavior from gradual normal changes. This method can support

real-time monitoring in large distribution networks, especially in systems affected by renewable generation, load fluctuation, and frequent remote-control actions. It may be applied to feeder monitoring, transformer protection, voltage stability assessment, and grid security management. However, the study also has some limitations. The experiments were conducted using one regional grid dataset, and the model was mainly tested under fixed sampling intervals. Future work should examine the method in different power systems, seasonal conditions, and cyber-physical attack scenarios. Further improvement is also needed for adaptive threshold selection in wider grid deployment.

Acknowledgments
We are grateful to all respondents who participated in this study.

References

1. Acquaah, Y. T., & Kaushik, R. (2024). Normal-only anomaly detection in environmental sensors in CPS: A comprehensive review. IEEE Access, 12, 191086-191107.

2. Ahmed, I. (2024). Deploying Low-Latency Edge AI in Medical IOT Networks: A Case Study of Secure Real-Time Patient Monitoring Systems. American Journal of Scholarly Research and Innovation, 3(02), 337-374.

3. Alam, M. A., Nabil, A. R., Mintoo, A. A., & Islam, A. (2024). Real-time analytics in streaming big data: techniques and applications. Journal of Science and Engineering Research, 1(01), 104-122.

4. Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. Symmetry, 15(9), 1679.

5. Chen, X., Xiao, H., Zeng, Z., Zhang, S., & Xiao, R. (2025). Fine-Grained Multivariate Time Series Anomaly Detection via Causal Inference. Knowledge-Based Systems, 114765.

6. De La Cruz, J., Gómez-Luna, E., Ali, M., Vasquez, J. C., & Guerrero, J. M. (2023). Fault location for distribution smart grids: Literature overview, challenges, solutions, and future trends. Energies, 16(5), 2280.

7. Gao, G., Gao, R., Lu, C., Gao, R., & Kuang, Y. (2026, March). Security Governance Methods and Quantitative Evaluation for Enterprise SMS and Verification Code Systems. In 2026 International Conference on Generative Artificial Intelligence and Information Security (GAIIS) (pp. 455-458). IEEE.

8. Gui, H., Wang, B., Lu, Y., & Fu, Y. (2025). Computational Modeling-Based Estimation of Residual Stress and Fatigue Life of Medical Welded Structures.

9. Jemili, F., Jouini, K., & Korbaa, O. (2025). Intrusion detection based on concept drift detection and online incremental learning. International Journal of Pervasive Computing and Communications, 21(1), 81-115.

10. Jiao, Y., Wang, A., Zhao, B., & Shi, T. (2026). The Impact of Visual Language Strategies in Public Art Creation on Community Spatial Perception and Public Behavior.

11. Johnson, J. M. (2022). The Role of Spatial Data Science in Continental Scale Hydrology: Twelve Case Studies in Data Models, Data Structures, Modeling, and Model Evaluation (Doctoral dissertation, University of California, Santa Barbara).

12. Kumari, S., Prabha, C., Karim, A., Hassan, M. M., & Azam, S. (2024). A comprehensive investigation of anomaly detection methods in deep learning and machine learning: 2019–2023. IET Information Security, 2024(1), 8821891.

13. Li, Y., & Liu, S. (2026, May). A Study on Dynamic Optimization of Alerting Policies and Multi-Agent Decision-Making Mechanisms in Cloud Environments. In 2026 7th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT) (pp. 703-706). IEEE.

14. Liang, R., Ye, Z., Liang, Y., & Li, S. (2025). Deep Learning-Based Player Behavior Modeling and Game Interaction System Optimization Research.

15. Ma, Y. (2026). Industrial Financial Risk Prediction Model Based on Graph Neural Network and Knowledge Graph Inference. Journal of Circuits, Systems and Computers, 35(16), 2650103.

16. Munandar, T. A., Lestari, T. S., Noe'man, A., Alimuddin, A., & Arafiyah, R. (2026). A Multidimensional Benchmark of Generative Models for Neonatal Clinical Tabular Data: Trade-offs in Fidelity, Utility, and Realism.

17. Pei, Z., Huang, Q., & Wang, S. (2026). When LLMs Develop Languages: Symbolic Communication for Efficient Multi-Agent Reasoning. arXiv preprint arXiv:2606.29354.

18. Qi, C., & Qiao, X. (2026). Using AI to Monitor AI: Automated Operations Through Log-Driven Intelligence. Available at SSRN 6795840.

19. Singh, S., Garg, A., & Mitra, K. (2024). Hdrsplat: Gaussian splatting for high dynamic range 3d scene reconstruction from raw images. arXiv preprint arXiv:2407.16503.

20. Su, D., & Zhang, H. Developing a Data-Driven Computational Framework for Scalable Special Education Practices for Autism and Intellectual Disabilities in the US Public School System.

21. Usmani, U. A., Aziz, I. A., Jaafar, J., & Watada, J. (2024). Deep learning for anomaly detection in time-series data: An analysis of techniques, review of applications, and guidelines for future research. IEEE Access, 12, 174564-174590.

22. Xu, S., & Zhou, Y. (2025). Machine Learning Applications in Financial Statement Fraud Detection: A Comparative Analysis. Annals of Applied Sciences, 6(1).

23. Xu, T., Zhang, J., & Zhu, W. (2026). Reproducible Modeling Pipelines and Cross-Window Stability in Subprime Auto Loan Credit Risk Assessment. Available at SSRN 6893861.

24. Xu, T., Zhu, W., & Zhang, J. (2026). Stability and Consistency of Explainable Deep Learning Methods in Credit Risk Assessment. Available at SSRN 6893938.

25. Yang, J. (2026). Stage-Coupled Computational Framework for Stratified Accessibility and Equity Analysis in Community-Based Elderly Care Services.

26. Yang, M., Wu, J., Tong, L., & Shi, J. (2025). Design of Advertisement Creative Optimization and Performance Enhancement System Based on Multimodal Deep Learning.

27. Yaswanthreddy, R., Nidhya, R., Penjuri, D., Syed, F., KV, D. K., & Nakkinti, L. (2026). Self-supervised learning for anomaly detection in multivariate time series from smart grids. Franklin Open, 15, 100583.

28. Yin, J., Huang, Y., & Rao, H. (2026). A Study on User Behavior Signal-Driven Predictive Models for Digital Platform Consumption Trends. Available at SSRN 6607298.

29. Zhang, Z. (2026). A Study on Return Optimization in E-commerce for Complex Consumer Goods Driven by Installation Information Quality. Available at SSRN 6734720.

30. Zhang, Z., Gao, Y., & Tong, Y. (2026). Semantic-Temporal Graph Learning for Demand Forecasting and Resource Allocation in Public Information Systems. Available at SSRN 6792279.