RESEARCH ARTICLE

Learning Spatiotemporal Representations of Irregular Urban Traffic Flows

Weijie Tan[1], Amanda Li[2*]

School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore

Corresponding author: amanda.lim@ntu.edu.sg

Abstract

Data centers continuously collect power-related time series from server power draw, rack-level energy consumption, CPU utilization, fan speed, inlet temperature, power distribution units, workload scheduling, and cooling interaction signals. These signals are noisy because of workload bursts, virtual machine migration, cooling feedback, transient load balancing, sensor fluctuation, and power measurement jitter. Raw power anomalies are difficult to interpret because normal workload-induced changes may resemble early hardware degradation or abnormal energy use. This study proposes a disentangled power signal denoising method for robust server anomaly detection in data centers. The method applies a conditional diffusion model to reconstruct stable power-consumption trajectories from noisy measurements. A disentanglement module separates workload-driven variation, cooling-related fluctuation, and server-side abnormal residuals. Anomaly scores are computed from denoised power residuals and rack-level consistency constraints. Experiments are conducted on a data center power monitoring dataset containing 18 server rooms, 9,800 physical servers, 620 racks, 74 power distribution units, and 46 monitoring variables collected at 15-second intervals over 11 months. The dataset contains 1.92 billion timestamped records and 3,180 operations-confirmed abnormal episodes, including abnormal power spikes, fan-control instability, rack-level energy imbalance, degraded power supply units, unexpected idle power increase, and cooling-power coupling faults. The proposed method reduces median detection delay from 26.4 minutes to 7.5 minutes compared with a temporal denoising autoencoder. False operation tickets decrease to 1.8 cases per server room per week. Diffusion denoising lowers normalized power reconstruction deviation from 0.141 to 0.052, and workload-disentanglement removes 68 million burst-load windows from the abnormal candidate queue. The online scoring engine processes 72,000 power records per second with 38 ms median inference latency per rack window. The results demonstrate that diffusion-based denoising and disentangled representation learning can improve robust anomaly detection in noisy data center power time series.

Keywords: Data center monitoring; server power anomaly detection; noisy time series; diffusion denoising; disentangled representation; workload fluctuation; robust anomaly detection.

Introduction

Data centers have become essential infrastructure for cloud computing, artificial intelligence, online platforms, digital services, and enterprise information systems. Their reliable operation depends heavily on stable power delivery, efficient energy utilization, and timely identification of abnormal server behavior. Even a small power anomaly may indicate power-supply degradation, cooling-system inefficiency, hardware faults, workload imbalance, or improper resource allocation. When such anomalies are not detected promptly, they may increase energy consumption, shorten equipment lifetime, interrupt computing services, and create cascading operational risks across racks and server rooms. Accurate server power anomaly detection is therefore important not only for equipment maintenance but also for energy management, service reliability, and sustainable data center operation. Modern data centers continuously collect multivariate time-series data from server power draw, rack-level energy consumption, CPU utilization, memory usage, fan speed, inlet and outlet temperature, power distribution units, workload scheduling systems, cooling controllers, and network devices. These variables exhibit strong temporal dynamics and complex cross-variable dependencies. Server power consumption changes with workload bursts, virtual machine migration, resource contention, load balancing, frequency scaling, and task scheduling. At the same time, cooling feedback, fan-speed adjustment, sensor fluctuation, transmission delay, and power measurement jitter may introduce additional short-term variations. Existing multivariate anomaly detection methods have employed graph-based dependency modeling [1,2], Transformer-based temporal representation learning and causal inference to distinguish initiating anomalies from propagated responses among correlated variables [3]. Other studies have investigated deep reconstruction, temporal prediction, and system-level monitoring for complex computing environments [4,5]. Although these approaches improve anomaly recognition under multivariate conditions, separating genuine server-side faults from workload-driven or control-induced fluctuations remains difficult [6].

Statistical monitoring methods generally identify abnormal observations through predefined thresholds, distribution assumptions, or deviations from historical baselines. These methods are computationally efficient and interpretable, but they are often sensitive to non-stationary workloads and changing operating conditions. Fixed thresholds may generate frequent false alarms during workload peaks, while adaptive thresholds may gradually absorb weak but persistent hardware degradation into the normal range. Deep learning methods, including autoencoders, recurrent neural networks, graph neural networks, and attention-based architectures, provide greater capacity for learning nonlinear temporal patterns and cross-variable relationships [7,8]. Reconstruction-based models detect anomalies according to the difference between observed and reconstructed signals, whereas prediction-based models use future prediction errors to identify deviations from expected behavior [9,10]. Attention and graph mechanisms further model dependencies across servers, racks, power units, and environmental variables. Despite this progress, most existing models process measured power signals as mixtures of multiple underlying sources. A sudden increase in server power may be caused by a legitimate workload burst, a scheduled virtual machine migration, an increase in cooling demand, a temporary imbalance among racks, or an actual server-side fault [11,12]. Reconstruction and prediction errors alone do not explicitly identify which component produces the observed deviation. Consequently, normal workload changes may receive high anomaly scores, while weak abnormal residuals may be concealed by dominant workload fluctuations. This limitation is particularly serious in large-scale data centers, where thousands of servers operate under heterogeneous workloads and power behavior varies across equipment models, rack configurations, and cooling zones [13,14]. Power measurements are also affected by several forms of noise. Sensor quantization, sampling inconsistency, communication delay, missing observations, synchronization errors, and power measurement jitter can distort local temporal patterns. Short-lived measurement noise may be mistaken for abrupt equipment anomalies, whereas excessive smoothing may remove weak degradation signals that develop gradually [15,16]. Conventional filtering methods usually rely on fixed frequency assumptions or local averaging and may not preserve nonlinear power transitions associated with server operation. A practical detection system must therefore suppress irrelevant noise while retaining abnormal temporal structures and avoiding the removal of early fault indicators [17,18]. Another limitation concerns the use of spatial and operational consistency. Servers located within the same rack commonly share power distribution units, cooling conditions, workload policies, and environmental constraints [19,20]. Their power trajectories are therefore not independent. When a rack-wide workload increase occurs, several servers may exhibit coordinated power changes, whereas a hardware anomaly may appear as an inconsistent residual concentrated on one server or a small subset of devices. Existing methods often calculate anomaly scores independently for individual time series or learn general correlations without explicitly enforcing rack-level consistency [21,22]. As a result, they may be unable to distinguish coordinated operational changes from localized abnormal behavior. Incorporating rack-level relationships can provide additional evidence for determining whether a power deviation is a normal collective response or an abnormal server-specific event [23,24].

To address these limitations, this study proposes a disentangled power signal denoising method for robust server anomaly detection in data centers. Conditional diffusion reconstruction is used to estimate stable power-consumption trajectories from noisy multivariate measurements while preserving meaningful temporal transitions. A signal disentanglement module separates the reconstructed representation into workload-driven variation, cooling-related fluctuation, and server-side abnormal residuals. Rather than directly treating all reconstruction deviations as anomalies, the proposed method calculates anomaly scores from denoised server-side residuals and rack-level consistency constraints. This design is intended to suppress false alarms caused by workload bursts, cooling feedback, transient load balancing, and measurement jitter while maintaining sensitivity to weak and persistent server abnormalities. The method is evaluated using operational data collected from 18 server rooms containing 9,800 physical servers, 620 racks, and 74 power distribution units. The dataset covers 46 monitoring variables over 11 consecutive months, including power consumption, computing utilization, thermal conditions, cooling-control signals, rack-level energy measurements, and workload scheduling information. This large-scale evaluation allows the proposed method to be examined under heterogeneous server configurations, changing workloads, environmental variation, and realistic measurement noise. The study focuses on anomaly detection accuracy, event-level recall, false operation ticket rate, detection delay, power-signal reconstruction quality, and robustness under different noise conditions. The objective of this study is to establish a reliable server power anomaly detection framework that can separate normal operational variation from genuine abnormal energy behavior in large-scale data centers. By combining diffusion-based denoising, interpretable signal disentanglement, and rack-level consistency modeling, the proposed method provides a more direct representation of server-side abnormal residuals than conventional reconstruction- or prediction-based approaches. The study is expected to improve the reliability of power monitoring, reduce unnecessary operation and maintenance tickets, and shorten the time required to identify emerging hardware or energy-efficiency problems. From an operational perspective, more accurate anomaly detection can support preventive maintenance, equipment health assessment, cooling optimization, and workload management. From an energy perspective, it can help identify inefficient servers and abnormal power consumption before they lead to substantial energy waste. The resulting framework therefore contributes to near-real-time monitoring, reliable infrastructure management, and energy-efficient operation in large-scale data centers.

Materials and Methods

2.1 Samples and Study Area

This study used a data center power monitoring dataset collected from 18 server rooms over 11 months. The monitored system included 9,800 physical servers, 620 racks, 74 power distribution units, workload scheduling logs, rack-level cooling records, and server operating signals. A total of 46 variables were recorded at 15-second intervals, including server power draw, rack energy use, CPU utilization, fan speed, inlet temperature, power distribution status, cooling interaction signals, and workload migration records. The dataset contained 1.92 billion timestamped records and 3,180 operations-confirmed abnormal episodes. These episodes included abnormal power spikes, fan-control instability, rack-level energy imbalance, degraded power supply units, unexpected idle power increase, and cooling-power coupling faults. All records were collected under production workloads, virtual machine migration, changing cooling conditions, and routine data center operation.

2.2 Experimental Design and Control Setup

Two groups were used to evaluate the proposed method. The experimental group used conditional diffusion reconstruction, signal separation, and rack-level consistency constraints for server anomaly detection. This group separated workload-driven variation, cooling-related fluctuation, and server-side abnormal residuals before anomaly scoring. The control group used a temporal denoising autoencoder, which detected anomalies from reconstruction errors without separating signal sources. Both groups were tested on the same monitoring records and confirmed abnormal episodes to ensure a fair comparison. The temporal denoising autoencoder was selected as the baseline because it is commonly used for noisy power time-series monitoring. This design allowed comparison of detection delay, false operation tickets, reconstruction deviation, burst-load filtering, and online inference latency.

2.3 Measurement Methods and Quality Control

Server power draw, rack energy use, CPU utilization, fan speed, inlet temperature, power distribution status, workload scheduling, and cooling-control signals were collected from power distribution units, server management interfaces, rack sensors, and operation platforms. Raw records were checked for missing values, repeated readings, abnormal spikes, clock mismatch, and communication errors. Short missing intervals were filled by linear interpolation, while long gaps caused by sensor or logging failure were removed from training and testing. Workload logs were used to identify normal burst-load periods and virtual machine migration events. Cooling records and inlet temperature signals were used to verify cooling-related power changes. Operation tickets, maintenance logs, and power supply inspection records were used to confirm abnormal episodes and their time ranges. These steps reduced measurement error and improved validation reliability.

2.4 Data Processing and Model Formulation

All variables were aligned to a 15-second time grid and normalized by z-score transformation. Let x_t denote the noisy server power signal at diffusion step t, and let c denote condition information, including workload level, rack location, inlet temperature, and cooling state. The conditional diffusion module estimated the denoised power trajectory as:

$$\hat{x}_0 = \frac{x_t - \sqrt{1-\bar{\alpha}_t}\,\varepsilon_\theta(x_t, t, c)}{\sqrt{\bar{\alpha}_t}}$$

where $\bar{\alpha}_t$ is the cumulative noise schedule and ϵ_θ is the learned noise estimator. The latent representation was then separated into workload-driven, cooling-related, and server-side abnormal components. The server anomaly score was calculated as:

$$A_i(t) = \frac{|x_i(t) - \hat{x}_i(t)|}{\sigma_i} + \eta|z_i^{\mathrm{srv}}(t)| + \rho C_r(t)$$

where $x_i(t)$ is the observed value of variable i, $\hat{x}_i(t)$ is the denoised estimate, σ_i is the normal standard deviation, $z_i^{srv}(t)$ is the server-side abnormal component, and $C_r(t)$ is the rack-level consistency penalty. The parameters η and ρ control the weights of server abnormality and rack consistency. An anomaly was detected when $A_i(t)$ exceeded the adaptive threshold learned from normal records under similar workload and cooling conditions.

2.5 Statistical Analysis and Validation

Model performance was assessed using median detection delay, false operation tickets per server room per week, normalized power reconstruction deviation, burst-load filtering rate, and online inference latency. The proposed method and the temporal denoising autoencoder baseline were compared using paired statistical tests at a significance level of 0.05. Sensitivity tests were conducted by changing diffusion steps, anomaly thresholds, workload partitions, and rack-level consistency weights. All analyses were performed in Python using NumPy, SciPy, pandas, and PyTorch. Operations-confirmed abnormal episodes were used as reference labels to assess whether the detected anomalies matched real power supply faults, cooling-power coupling faults, or abnormal server energy behavior.

Results and Discussion

3.1 Server Power Anomaly Detection under Workload Noise

The proposed power signal denoising method achieved better anomaly detection than the temporal denoising autoencoder. The improvement was clear for abnormal power spikes, fan-control instability, rack-level energy imbalance, degraded power supply units, unexpected idle power increase, and cooling-power coupling faults. The median detection delay decreased from 26.4 min to 7.5 min, showing better sensitivity to early abnormal power behavior. As shown in Fig. 1, power prediction models can detect abnormal energy patterns by comparing predicted and observed values. However, prediction-error methods may confuse workload bursts with real server faults. The proposed method reduced this problem by reconstructing stable power trajectories and separating workload-related variation before anomaly scoring [25,26].

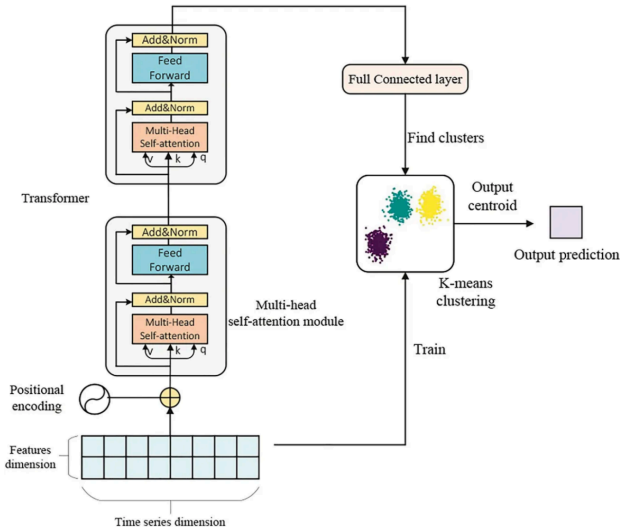

Fig. 1. Transformer–K-means model for power prediction and anomaly detection.

3.2 Denoising Effect and Power Signal Separation

Diffusion denoising reduced the normalized power reconstruction deviation from 0.141 to 0.052. This result shows that the model recovered more stable server power patterns from noisy records [27,28]. The workload-separation module removed 68 million burst-load windows from the abnormal candidate queue, reducing the effect of normal workload changes on anomaly detection. As illustrated in Fig. 2, recent energy anomaly studies have used root-cause tracing frameworks to link abnormal energy use with possible

causes. Unlike methods that analyze mixed energy deviations directly, the proposed method separated workload-driven, cooling-related, and server-side abnormal components. This made the anomaly score more specific to server-side faults.

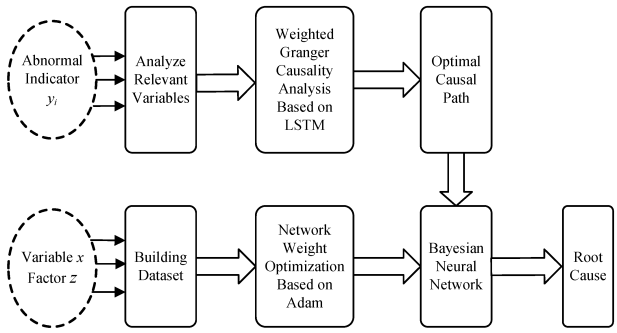

Fig. 2. Root-cause tracing model for energy-use anomaly analysis.

3.3 False Operation Ticket Reduction and Rack-Level Stability

False operation tickets decreased to 1.8 cases per server room per week, which was lower than the baseline result. This reduction is important for data center operation because frequent false tickets increase engineer workload and reduce trust in automatic monitoring systems. The lower false alert rate was mainly due to rack-level consistency constraints. A server was not marked as abnormal only because its power draw changed sharply. It also needed to show abnormal residual behavior that was inconsistent with neighboring servers and rack-level operating conditions. This rule reduced false alarms caused by virtual machine migration, short workload bursts, load balancing, and cooling feedback [29].

3.4 Practical Value and Remaining Limitations

The method was evaluated on 1.92 billion records from 18 server rooms, 9,800 physical servers, 620 racks, 74 power distribution units, and 46 variables collected over 11 months. The online scoring engine processed 72,000 power records per second, with a median inference latency of 38 ms per rack window. These results indicate that the method is suitable for near-real-time anomaly detection in large data centers. The findings show that diffusion-based denoising and signal separation can improve detection timeliness, power signal reliability, and operational interpretation. However, the method depends on operation-label quality, workload-log completeness, rack topology accuracy, and separation-weight selection. Future studies should test the method in larger GPU clusters, liquid-cooling systems, heterogeneous server generations, and unseen workload patterns.

Conclusion

This study developed a power signal denoising method for server anomaly detection in data center time series. The method combines conditional diffusion reconstruction, signal separation, and rack-level consistency constraints to reduce workload-related noise and detect server-side abnormal power behavior. Experiments on 18 server rooms, 9,800 physical servers, 620 racks, and 1.92 billion records showed that the proposed method reduced the median detection delay from 26.4 min to 7.5 min compared with the temporal denoising autoencoder. False operation tickets decreased to 1.8 cases per server room per week, and normalized power reconstruction deviation decreased from 0.141 to 0.052. The workload-separation module also removed 68 million burst-load windows before anomaly scoring, which reduced false alarms caused by normal workload changes. These results show that diffusion denoising and signal separation can improve detection accuracy, power signal reliability, and operational interpretation in large data centers. The method can support near-real-time server monitoring, power fault warning, and maintenance planning. However, its performance depends on operation-label quality, workload-log completeness, rack topology accuracy, and separation-weight selection. Future studies should test the method in larger GPU clusters, liquid-cooling systems, mixed server generations, and unseen workload conditions.

References

Fu, Y., Gui, H., Li, W., & Wang, Z. (2020, August). Virtual Material Modeling and Vibration Reduction Design of Electron Beam Imaging System. In 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA) (pp. 1063-1070). IEEE.

Lee, J., Yoon, T., Koo, W., & Kim, H. (2026). Knowledge-Assisted Multi-Graph Dependency Learning for Multivariate Time Series Anomaly Detection in Multi-Stage Industrial Processes. IEEE Transactions on Automation Science and Engineering.

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.

Sharifi, S., Stocco, A., & Briand, L. C. (2025). System Safety Monitoring of Learned Components Using Temporal Metric Forecasting. ACM Transactions on Software Engineering and Methodology, 34(6), 1-43.

Liang, R., Feifan, F. N. U., Liang, Y., & Ye, Z. (2025). Emotion-Aware Interface Adaptation in Mobile Applications Based on Color Psychology and Multimodal User State Recognition. Frontiers in Artificial Intelligence Research, 2(1), 51-57.

Wu, C., Zhang, F., Chen, H., & Zhu, J. (2025). Design and optimization of low power persistent logging system based on embedded Linux.

Habiba, M., Pearlmutter, B. A., & Maleki, M. (2026). Different Neural Network Models for Time Series Processing. In Recent Trends in Modelling the Continuous Time Series Using Deep Learning (pp. 27-61). Cham: Springer Nature Switzerland.

Shao, W. (2026, March). Interpretable Ensemble Learning for Network Traffic Anomaly Detection: A SHAP-based Explainable AI Framework for Embedded Systems Security. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-4). IEEE.

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.

Attila, K., Molnár, J. K., & Károly, J. (2026). Correction Functions and Refinement Algorithms for Enhancing the Performance of Machine Learning Models. Automation, 7(2), 45.

Zhang, Z., Gao, Y., & Tong, Y. (2026). CausalML4CX: A Causal Inference and Machine Learning Framework for Customer Experience Optimization with Application in Banking.

Xu, T., Zhu, W., & Zhang, J. (2026). A Study on the Application of Alternative Data in Credit Assessment for the Unbanked Population.

Gyang, P. P., Chakraborty, P., Meegahapola, L., & Yu, X. (2026). Grid Integrated Data Centers: Architectures, Modeling, Interconnection Requirements, Challenges, and Opportunities. IEEE Open Journal of the Industrial Electronics Society.

Xu, T., Zhang, J., & Zhu, W. (2026). Fairness-Accuracy Trade-offs and Calibration Mechanisms in Machine Learning-Based Subprime Credit Scoring. Available at SSRN 6893759.

Wang, Y., Chen, J., Arias, R., Wang, Y., & Yin, X. (2026). Development and Validation of a Patient-Friendly Digital Assessment Platform for Precision Screening of Oral Anti-Obesity Medications (AOMs).

Rawat, R., Borana, K., Chaturvedi, N., Rawat, H., Rawat, A., Rajavat, A., & Rimal, Y. (2026). Deep Learning and Pattern Recognition for Anomaly Detection in Multimodal Medical Imaging With Sensor-Based Data Fusion. Journal of Sensors, 2026(1), 9935004.

Zhu, W., & Yang, J. (2025). Causal Assessment of Cross-Border Project Risk Governance and Financial Compliance: A Hierarchical Panel and Survival Analysis Approach Based on H Company's Overseas Projects.

Jiao, Y., Zhao, B., Wang, A., & Shi, T. (2026). Construction and Empirical Study of a Modularized Teaching System for Art Courses Based on a Unified Training Pathway.

Erkek, T. Ü., Erkek, M., Aydın, M. B., & Gezer, K. (2025). Experimental performance characterization of R-454B and R-410A in data center server rack mount cooling unit: Energy efficiency and environmental impact assessment. International Journal of Refrigeration.

Yin, J., Huang, Y., & Rao, H. (2026). A Study on the Dynamic Evolution of Learning Behavior and Outcome Prediction in Digital Educational Environments. Available at SSRN 6607139.

Yang, J. (2026). Computational Analysis of How Digital Non-Clinical Communication Tools Influence Social Participation Among Older Adults in Community-Based Elderly Care. Available at SSRN 6682838.

Hinov, N. (2025). From Energy Efficiency to Energy Intelligence: Power Electronics as the Cognitive Layer of the Energy Transition. Electronics, 14(23), 4673.

Zhang, Z. (2026). A Study on Multimodal Product Understanding and Assembly Guidance Optimization in e-commerce for Complex Consumer Goods. Available at SSRN 6734559.

Qi, C., & Qiao, X. (2026). Building and Operating a Large Scale Multi-Agent System: A Case Study from Industry. Available at SSRN 6795198.

Khatun, Z. (2025). AI-Driven Predictive Maintenance For Motor Drives In Smart Manufacturing A Scada-To-Edge Deployment Study. American Journal of Interdisciplinary Studies, 6(1), 394-444.

Su, D., & Wu, Y. (2026). Multilevel Growth and Social Network Modeling for Functional Communication Enhancement in Students with Intellectual Disabilities.

Moshawrab, M., Adda, M., Bouzouane, A., Ibrahim, H., & Raad, A. (2023). Reviewing Federated Learning Aggregation Algorithms; Strategies, Contributions, Limitations and Future Perspectives. Electronics (2079-9292), 12(10).

Xie, Z., Ren, X., Zheng, T., Bai, J., Fan, W., Xu, B., ... & Song, Y. (2026). A Survey on AI Agent Harness. ResearchGate Preprint. DOI, 10.

Naji, H. R., & Riyahi, H. (2025). A Survey of Load Balancing Approaches Based on Switch Migration in Software-Defined Networks. IET Networks, 14(1), NA-NA.