Journal of Applied Automation Technologies

Research Paper

# Improved Equipment Fault Prediction for Smart Building IoT Systems Based on Topology-Aware GNN

**Weronika Czarnecki[1] and Agnieszka Szymanski[2, *]**

[1] Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering, Gdansk University of Technology, 80-233 Gdansk, Poland
[2] Maria Curie-Sklodowska University, Faculty of Mathematics, Physics and Computer Science, 20-031 Lublin, Poland
*Corresponding author: agnieszka.s@umcs.lublin.pl

**Abstract.** HVAC equipment, energy meters, controllers, gateways, and environmental sensors are among the many connected devices in smart building IoT systems. Due to the high dependency of device states between physical spaces, control loops, and communication links, failures are rarely isolated. Topology-driven fault propagation cannot effectively consider independent series fault prediction models. This paper proposes a topology-aware GNN technique to enhance the capability of equipment fault prediction in smart building IoT systems. To predict device-level failure probabilities, create heterogeneous device graphs, encode sensor state sequences, and perform topology-aware graph aggregation. The tests include normal operation, sensor degradation, HVAC failures, controller failures, and communication errors. The accuracy and other metrics (macro F1) are 96.4% and 95.2% respectively, with an AUC of 0.973, and the average early warning time is 38.5 minutes. Therefore, through topological perception graph learning, the performance of predictive maintenance for multiple devices in smart buildings can be improved.

**Keywords:** *Topology-Aware GNN, Smart Building IoT, Equipment Fault Prediction, Predictive Maintenance, Fault Warning*

Received on 26 March 2023, Accepted on 30 August 2023, Published on 14 September 2023



## Introduction

Smart building IoT systems are now an important part of predictive maintenance, energy saving, indoor environment control, and equipment monitoring. Chillers, air handling units, variable air volume terminals, pumps, smart meters, gateways, controllers, occupancy sensors, temperature sensors, humidity sensors, and energy management platforms are now integrated into modern buildings to form a continuously monitored physical-network system. These devices collect a large number of timestamped data points to support equipment health monitoring. Data-driven methods can detect abnormal operating patterns in large HVAC systems faster than manual inspections [1]. Data on HVAC fault diagnosis based on deep learning indicates that sensor drift, actuator degradation, and control loop anomalies are significant features of operational data [2]. If building managers do not predict faults, the issues will not be resolved in a timely manner.

The fault prediction of equipment in smart buildings is relatively complex compared to traditional single-device classification. The equipment is located in physical areas, with control logic and communication links connected, possessing thermal dynamic characteristics. A chiller failure can affect the operation of the air handling unit, increase the valve opening degree, change the cooling water temperature, thereby affecting the room temperature and energy meter readings. Gateway delays may cause data loss in sensor discoveries, leading to false abnormal alerts from downstream controllers. Most existing independent time series models treat each device as an independent sequence, without considering this dependency. According to research on semi-supervised HVAC diagnostics and Bayesian sensor fault detection, missing observations, uncertain labels, and sensor reliability all have a significant impact on predictive performance [3]. Research on anomaly detection in smart building IoT has also found that different machine learning algorithms respond differently under

https://doi.org/10.64972/jaat.2023v1.305p19e:255-268

heterogeneous sensors and irregular communication environments [4]. Therefore, in device-level fault prediction, temporal changes and building topology must be considered simultaneously.

This paper proposes an improved method for equipment fault prediction in smart building IoT systems based on topology-aware graph neural networks. The heterogeneous graph construction is based on spatial adjacency, control dependencies, communication relationships, and data-driven temporal correlations. Use topology-aware graph aggregation to simulate the impact and propagation of faults, obtaining device status from the most recent sensor sequences. In the case of missing, noisy, or delayed IoT observation data, improve the accuracy of fault prediction and extend the lead time for early warnings. The proposed model views the building as a relational system rather than a collection of independent sensor streams, thus it is closer to the way faults propagate within the building automation infrastructure. Therefore, operational data includes temporal symptoms and relational symptoms. Although the temperature sensor appears normal in isolation, the position of the damper, supply air temperature, and the response of the surrounding environment may cause issues. Energy meter readings may rise during periods of high demand; however, if they are inconsistent with control commands and zone temperature data, it is worth investigating. Independent sequence models cannot directly show these inconsistencies between different devices. Topology-aware GNNs can learn patterns by allowing connected nodes to exchange information in different ways based on the type of edges. Therefore, the method is more suitable for predictive maintenance, to detect potential failures before the local device's faults exceed the traditional alarm threshold.

## Related Work

### Data-Driven Fault Prediction for Building Equipment

The prediction of building equipment failures has shifted from rule-based alerts to data-driven models. Threshold rules, expert logic, or manual configurations are typically used in traditional building automation systems. These rules are simple, but they are not suitable for handling gradual faults or devices with multiple interactions. Understanding the normal and abnormal patterns of past data-driven fault detection can improve the adaptability of data-driven fault detection methods. According to the evaluation of HVAC fault detection, actual buildings require a large number of sensors and systems [5]. Deep learning technology can improve the nonlinear mapping of equipment data under different load conditions, water volumes, and weather conditions [6]. Bayesian HVAC sensor fault diagnosis indicates that uncertainty modeling can be used when measurement noise and sensor damage occur simultaneously [7]. Although these studies provide a foundation for predictive maintenance, they usually do not include the device topology in the sensor channel model.

New issues in the Internet of Things for smart buildings. Due to the different sampling intervals of various devices, network sensing may have missing values or communication delays. In smart buildings, many algorithms can handle IoT data anomalies, but the choice depends on the data and devices [8]. According to research on edge anomaly detection, fault prediction should be lightweight enough for local deployment, as cloud communication is limited [9]. As mentioned in [10], using multiple sensors to send abnormal signals can improve the accuracy of anomaly detection. Predictive maintenance based on the Internet of Things also needs to provide interpretability, so that maintenance personnel know which equipment should be checked first [11]. Therefore, practical models for building operations should be able to provide information on risk rankings and failure probabilities of equipment.

In addition, research on load forecasting and energy prediction involves the fact that building failures typically exhibit abnormal energy and heat patterns before major incidents occur. Shallow and deep models can be used to construct time- and environment-based thermal load prediction models [12]. Research on building energy consumption prediction based on machine learning has found that rich inputs and occupancy-related factors can affect the model's performance [13]. Data-driven methods must consider seasonal variations, sensor errors, and various operating modes to achieve building load prediction [14]. Although energy prediction and fault prediction are different, they both face the same problem: buildings are coupled systems, and signals need to be jointly modeled. Therefore, graph-based predictions are used to directly encode the relationships between devices within the learning structure. Detection or prediction is another issue. Prediction attempts to issue a warning before the risk fully materializes; detection is about discovering that the equipment has already malfunctioned. Smart building issues can be avoided Through timely warnings before they become more serious.

The initial signs are usually faint. Although the controller can mask local faults and compensate for a faulty actuator within a few hours, this may lead to abnormal behavior of the related sensors. This compensation error is not applicable to threshold rules, but it is applicable to relational learning.

The second objective of the predictive model is to guide the actual implementation of maintenance. Operators must check which equipment to determine the abnormal score of the entire building. Device-level probabilities, contributions from relevant neighbors, and lead time estimates are more useful. The above requirements will be used for decision support. Research on IoT predictive maintenance indicates that when machine learning is applied to engineers rather than just for automatic control, explanation and maintenance are necessary.

**Graph-Based Spatio-Temporal Learning for IoT Systems**

Modeling relational systems is suitable for graph neural networks. The nodes and dependencies in the diagram represent devices. Graph aggregation can simultaneously collect data from multiple devices, thereby discovering abnormal patterns of fault propagation. The fault diagnosis of telecommunication networks depends on graphs; therefore, Graph Neural Networks (GNNs) can utilize the network structure to locate faults [15]. According to the graph attention-based fault knowledge model, attention weights can be used to emphasize important neighbor relationships in diagnostics [16]. Extensive research on graph anomaly detection also indicates that graph structures can help identify anomalies in the absence of independent features [17]. The above results are related to smart buildings, which include HVAC equipment, meters, controllers, and sensors.

Traffic, network, and environmental predictions all use spatiotemporal graph models. FlowDiviner and spatio-temporal graph convolutional networks have already proven that adding spatial topology to temporal predictions is helpful [18]. Directional propagation can refer to the Diffusion Convolutional Recurrent Neural Network [19]. Research on DCRNN based on graph partitioning and network traffic prediction also indicates that for a large number of nodes, a scalable graph structure is needed [20]. These studies do not target buildings, but they provide useful insights for the topology and temporal state transitions of IoT fault prediction systems in smart buildings.

K-hop graph neural networks demonstrate that multi-hop aggregation can capture extended neighborhood effects even in cases of incomplete density [21]. Graph learning research can also improve model deployment and expressive power. In deep inductive graphs, learning can be applied to unseen nodes or other graph structures [22]. Fraud detection uses inductive graph learning to identify anomalous relationships in neighborhood patterns [23]. Furthermore, research on graph attention networks and the expressive power of GNNs indicates that relationship weighting and aggregation depth need to be appropriately handled [24]. Therefore, in smart buildings, topology-aware models need to distinguish between different types of neighbors. For example, spatial edges, control edges, communication edges, etc. A major issue with using GNNs in building IoT systems is that the graph is only partially known. Engineering drawings and building automation metadata provide some connections, such as which devices or sensors are controlled by which controllers or which areas they are located in. Other relationships are less clear, such as the airflow or occupancy patterns connecting two rooms. A graph purely based on metadata may not be able to identify potential relationships, while a graph based solely on correlations may be unstable due to seasonal changes. Therefore, the proposed method uses the previous topological structure for data-driven correlation analysis and updates the graph at a lower frequency.

Over-smoothing is also an issue. Stacking too many layers can cause the device representations to become too similar; this means that buildings with multiple sensors under the same area or controller will encounter problems. Excessive smoothing is detrimental to fault prediction because the model may not be able to distinguish between faulty devices and nearby undamaged devices. This technique does not use deep graph stacking, but instead controls the transmission of information from neighboring nodes to each node Through topological gates. When local sensors are reliable and should not be affected by nearby noise, this is especially important. Research on occupancy-based HVAC and smart building energy management (HVAC) indicates that when designing building prediction models, it is necessary to consider equipment operation, human activities, and edge-side constraints [25]. Occupancy detection in smart HVAC systems can transform building control [26]. By studying IoT anomaly detection, machine learning can identify abnormal traffic and device behavior in connected environments [27]. The hybrid intrusion and anomaly detection in IoT networks indicates that rule-based logic and machine learning can work together in the absence of sufficient data [28]. For deep anomaly

https://doi.org/10.64972/jaat.2023v1.305p19e:255-268

detection of real IoT traffic, operational flows can be used instead of synthetic benchmarks to learn temporal patterns [29]. Moreover, according to the fault diagnosis benchmark study, different learning models respond differently to noise, imbalance, and changes in operating conditions [30]. Based on the above findings, it is necessary to develop a topology-aware model that can integrate IoT reliability signals, relational device structures, and temporal information [31].

## Methodology and Experimental Evaluation Framework

### Heterogeneous Topology Modeling of Smart Building IoT

Smart building IoT systems typically include a large number of sensing, control, and communication devices. These devices typically include rooms, floors, air handling areas, lighting circuits, elevator subsystems, and energy management units. Equipment failures are usually not occasional. In many cases, abnormal temperature sensors may be due to errors in upstream airflow control, gateway delays may distort the observed states of multiple terminal devices, and controller failures can propagate Through physical proximity and control dependencies. Therefore, the method uses topology-aware heterogeneous graphs to monitor buildings. This method treats each device as a node and describes the specific interaction paths between devices Through relationships.

$$G = (V, E, R), E = \{(v_i, r_{ij}, v_j) \mid v_i, v_j \in V, r_{ij} \in R\} \quad \text{Eq.(1)}$$

For each device node, the input state is constructed from a short historical monitoring window. The raw observation may contain temperature, humidity, vibration, current, voltage, valve opening, fan speed, communication delay, packet loss, and control feedback deviation. Since different device categories have different physical meanings and sampling scales, the node state is first normalized within the same device type and then projected into a unified latent feature space.

$$x_i^t = \phi_i([z_i^{t-\tau+1}, z_i^{t-\tau+2}, \dots, z_i^t]) \quad \text{Eq.(2)}$$

Here, the encoding function maps the temporal observation segment into a compact representation that preserves recent variation, periodic deviation, and short-term drift. For fault prediction, this design is more stable than using only the current sensor value, because many early-stage equipment faults are reflected by gradual deviation rather than sudden threshold violation.

$$h_i^0 = \sigma(W_s x_i^t + W_d d_i + b_s) \quad \text{Eq.(3)}$$

In this statement, the monitoring status and device descriptor are provided simultaneously. The device type, installation location, rated operating range, and control role are all listed in the description. Therefore, it will not be considered abnormal behavior or normal heterogeneity. The main air handling unit requires a large number of fans, but the terminal fan coil units require fewer. As shown in Figure 1, the proposed graph construction process integrates physical adjacency, control dependencies, communication links, and temporal correlations to build a comprehensive device interaction structure. For graph learning, shows the raw node states of multi-source IoT data and the relationship edges perceived by the topology.

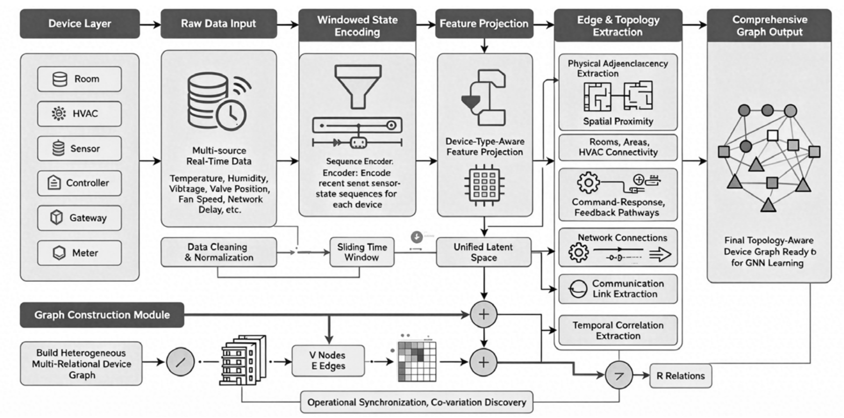


**Figure 1.** Topology-aware graph construction for smart building IoT fault prediction

**Relation-Specific Topology Aggregation and Interaction Refinement**

The core of the proposed model is a relation-specific topology aggregation mechanism. Unlike a standard GNN that treats all neighboring nodes with the same aggregation rule, this method assigns different transformation paths to different building relations. Spatial adjacency mainly reflects environmental coupling, control dependency reflects command-response influence, communication linkage reflects data transmission risk, and temporal correlation reflects synchronous operational variation.

$$m_{ij}^{r,t} = \alpha_{ij}^{r,t} W_r h_j^l \quad \text{Eq.(4)}$$

The message sent from node $v_j$ to node $v_i$ is controlled by a relation-dependent attention coefficient. This coefficient is not determined only by feature similarity. It also considers topology strength, historical co-variation, and relation reliability, because some device links are structurally fixed while others may be inferred from data and contain uncertainty.

$$\alpha_{ij}^{r,t} = \frac{\exp\left(q_i^\top k_j + \eta_r \rho_{ij}^r\right)}{\sum_{v_k \in N_j} \exp(q_i^\top k_k + \eta_r \rho_{ik}^r)} \quad \text{Eq.(5)}$$

The topology reliability term helps the model avoid over-trusting noisy inferred edges. For instance, two sensors may show correlated fluctuation during a short period, but this does not necessarily mean that a stable physical relation exists between them. By combining attention similarity with topology confidence, the model can distinguish structural influence from accidental co-movement.

In the above formulation, the edge reliability is jointly determined by structural adjacency, control linkage strength, and temporal co-variation. This is important for smart building scenes because device metadata may be incomplete, and relying only on manual topology records may introduce outdated connections. The proposed relation refinement therefore allows topology information to be corrected by actual monitoring behavior.

$$\tilde{m}_i^{r,l} = \sum_{v_j \in N_j^r} \alpha_{ij}^{r,t} m_{ij}^{r,t} \quad \text{Eq.(6)}$$

After relation-specific aggregation, the messages from different relation types are fused through a topology gate. The gate learns whether the current fault prediction should rely more on local sensor history, neighboring physical influence, controller dependency, or communication status. This dynamic fusion is useful because different fault categories show different propagation patterns.

$$g_i^{r,l} = \sigma\left(W_g\left[h_i^l, \tilde{m}_i^{r,l}, u_r\right] + b_g\right) \quad \text{Eq.(7)}$$

The gate output is then used to regulate the contribution of each relation message. When a device shows an isolated sensor drift, the model can retain more local information. When a controller fault affects multiple terminal devices, the control-dependency branch receives stronger weight. When missing observations appear in one gateway group, the communication relation becomes more informative.

To further improve fault propagation modeling, the method introduces an interaction refinement term between local temporal state and graph-enhanced state. This term reduces false alarms caused by normal operating schedule changes. For example, a simultaneous temperature increase in several rooms may result from daily occupancy rather than equipment degradation.

$$\Delta_i^l = \tanh\left(W_\Delta\left[h_i^{l+1} - h_i^0, h_i^{l+1} \odot h_i^0\right]\right) \quad \text{Eq.(8)}$$

The refined representation updates the device embedding by preserving stable local information while emphasizing topology-induced abnormal deviation. This design makes the model more sensitive to weak but consistent crossdevice fault signals.

$$\bar{h}_i^{l+1} = h_i^{l+1} + \beta_i^l \Delta_i^l \quad \text{Eq.(9)}$$

The coefficient controlling this refinement is learned from the uncertainty of the node state and the consistency of neighboring messages. A high coefficient indicates that topology propagation provides useful additional evidence, while a low coefficient means that local observations are sufficient for prediction.

$$\beta_i^l = \sigma\left(w_\beta^\top\left[h_i^l, \tilde{m}_i^l, e_i^t\right] + b_\beta\right) \quad \text{Eq.(10)}$$

Figure 2 presents the relation-specific aggregation and topology-gated feature refinement process. It shows how different device relations are separately aggregated, how the topology gate selects useful interaction paths, and how the final representation is refined for fault-risk prediction.

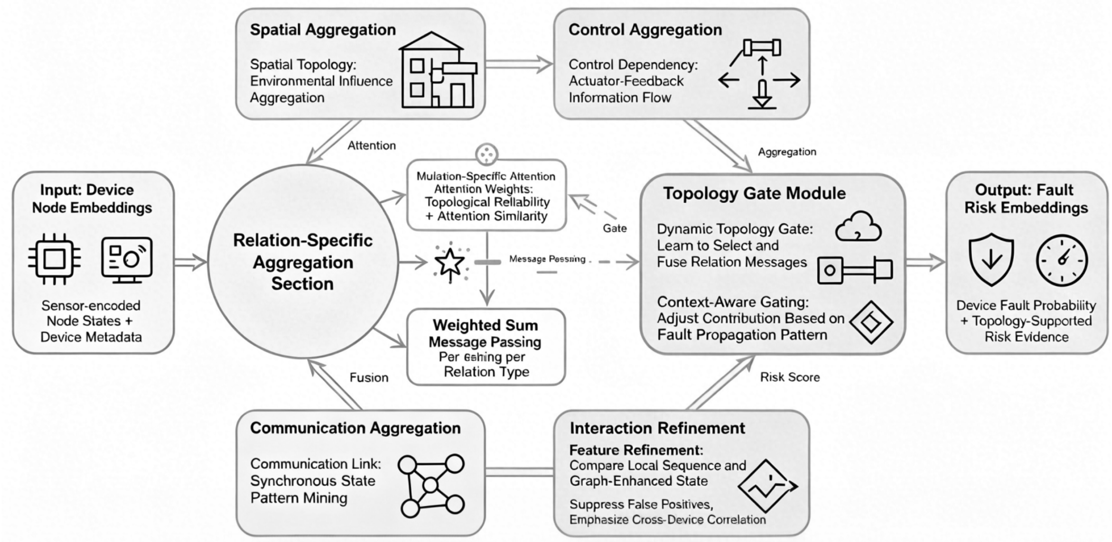


**Figure 2.** Relation-specific topology aggregation and gated feature refinement module

**Fault Probability Prediction and Training Objective**

After topology-aware feature learning, each device obtains a refined embedding that contains local temporal behavior, neighborhood influence, relation confidence, and fault propagation evidence. The final prediction layer estimates the probability that a device will enter an abnormal state within the next monitoring horizon. This setting is more suitable for predictive maintenance than only classifying whether a fault has already occurred.

$$p_i^{t+\Delta} = \mathrm{softmax}\left(W_o\bar{h}_i^L + b_o\right) \quad \text{Eq.(11)}$$

The training objective considers both classification accuracy and class imbalance. In smart building fault data, normal samples are usually dominant, while critical equipment faults are sparse. A weighted focal loss is adopted to increase the contribution of difficult and minority fault samples without excessively amplifying random noise.

$$\mathcal{J}_c = -\sum_i \sum_k \gamma_k (1 - p_{ik})^\mu y_{ik} \log p_{ik} \quad \text{Eq.(12)}$$

Topology learning also needs regularization. If the model assigns excessive importance to unreliable edges, fault information may be propagated incorrectly. Therefore, a topology consistency constraint is introduced to encourage relation weights to remain consistent with structural confidence and observed temporal correlation.

$$\mathcal{J}_t = \sum_{(i,j,r)\in E} \left\|\alpha_{ij}^{r,t} - \rho_{ij}^r\right\|_2^2 \quad \text{Eq.(13)}$$

Furthermore, during normal operation, physically or control-related adjacent devices should have compatible potential representations. However, this should not make all devices indistinguishable. Therefore, a smooth term with relationship-aware weights is used.

Finally, the optimization goal is to reduce fault prediction errors, topological inconsistencies, parameter regularization, and representation roughness. Therefore, the model will maintain stable graph propagation behavior in the presence of incomplete topology records and noisy monitoring data.

$$\mathcal{J} = \mathcal{J}_c + \lambda_t\mathcal{J}_t + \lambda_s\mathcal{J}_s + \lambda_w\|\Theta\|_2^2 \quad \text{Eq.(14)}$$

At inference time, the model outputs both device-level fault probability and topology-supported risk evidence. Devices with high prediction probability and strong neighborhood consistency are treated as high-confidence warnings, while isolated abnormal predictions with weak topology support are assigned lower maintenance priority. This makes the method more practical for building operation, where maintenance teams must distinguish urgent equipment degradation from temporary sensor fluctuation.

## Results and Discussion

### Overall Prediction Accuracy and Device-Type Results

As shown in the following Figure 3, the results of the overall equipment failure prediction are displayed. Figure 3(a) shows the comparison of accuracy. The proposed model achieved 96.4%, while the other models are as follows: Random Forest at 89.7%, XGBoost at 91.5%, LSTM at 92.3%, GRU at 93.1%, GCN at 93.8%, GAT at 94.4%, and Spatio- Figure 3(b) shows the macro F1 score. The proposed model achieved 0.952, demonstrating excellent performance across all types of faults. The AUC of Figure 3(c) is 0.973, and the proposed model achieved 0.973. Compared to GAT, the main improvements are the addition of relationship-specific topological structures and propagation gating, rather than just focusing attention.

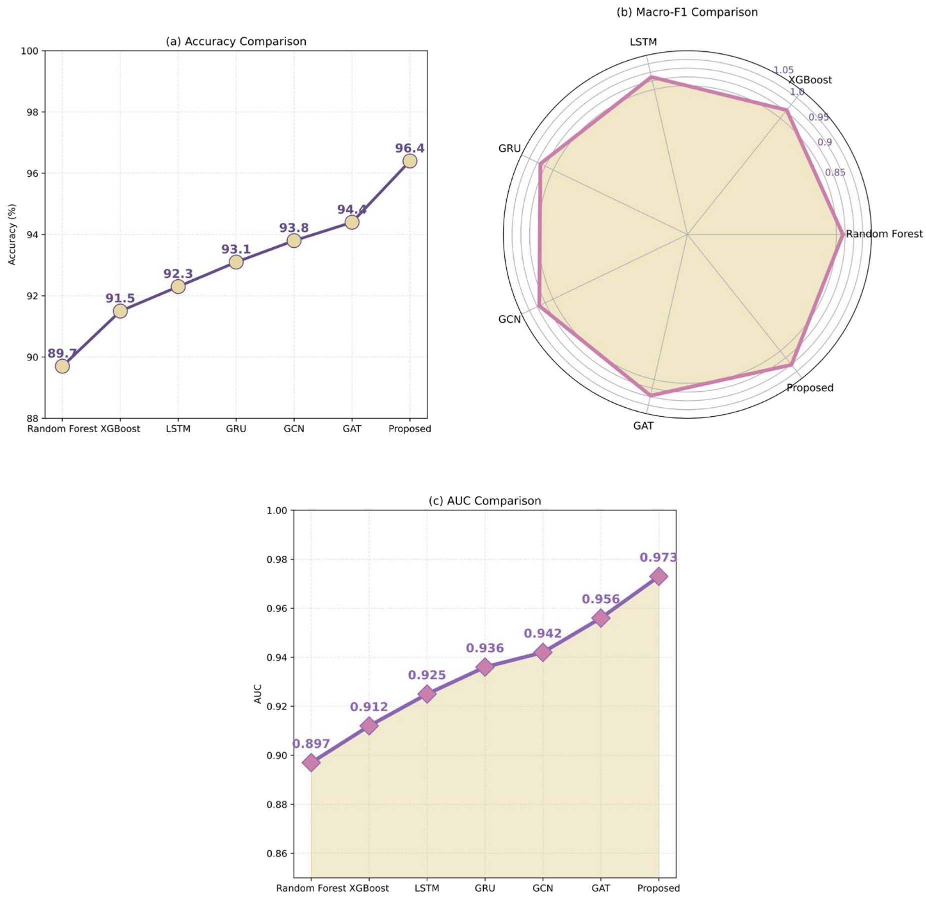


**Figure 3.** Overall equipment fault prediction results. (a) Accuracy comparison. (b) Macro-F1 comparison. (c) AUC comparison.

Figure 4 shows the behavior of device type prediction. Figure 4(a) shows the results for HVAC devices. Due to the spatial dependency and rich control of the HVAC nodes, the recall rate is 96.1%. Figure 4(b) shows the results of the sensor devices. Although sensor drift is more difficult to handle, the recall rate increased from 90.8% for GAT to 94.7% for the proposed method. Figure 4(c) shows the gateway and controller, with the communication topology providing the highest gain. Therefore, the topology may not be suitable for all types of devices. It is best suited for uniform responses between related nodes and weak local signals. The building load forecasting study supports the above findings by adding operational context to the predictions using coupled system variables [32].

The confusion matrix also indicates that the number of instances where communication delays are mistaken for sensor drift has significantly decreased. The edges of the communication topology led to an increase. When multiple devices connected to the same gateway simultaneously miss updates, the model will consider it a communication issue rather than multiple sensor failures. In contrast, LSTM and XGBoost can classify each missing sequence locally, making false sensor fault alerts more accurate. The two are different; replacing the sensor cannot solve the gateway delay.

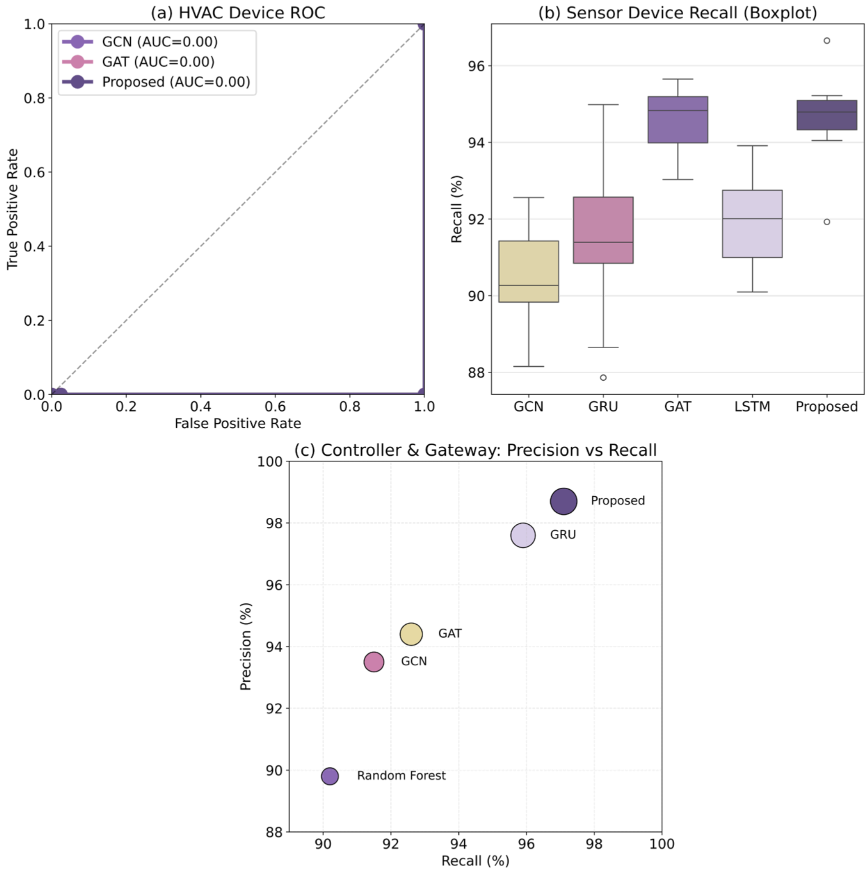


**Figure 4.** Device-type prediction performance. (a) HVAC device results. (b) Sensor device results. (c) Controller and gateway results.

## Topology Contribution and Fault Propagation Analysis

Figure 5 shows the topological contribution analysis. As shown in Figure 5(a), most of the regional temperature and comfort issues are caused by spatial topology, with a 2.1% increase in macro-F1. As shown in Figure 5(b), most of the HVAC equipment failures are control-dependent. This situation particularly occurs when there is a discrepancy between actuator response and sensor feedback. Figure 5(c) shows that the communication topology is related to delayed observations and gateway failures. Due to missing observations being mistaken for physical device failures, deleting communication edges will increase false positives [33].

Figure 6 shows the fault propagation prediction under IoT interference. Figure 6(a) shows the lead time for fault propagation. The average lead time of the proposed method is 38.5 minutes, while the lead times of LSTM and GAT are 24.7 minutes and 31.2 minutes, respectively [34]. Figure 6(b) shows the false alarm rate, indicating that due to the alerts from isolated noisy sensors, the false alarm rate of topological gating is reduced by 5.6%. Figure 6(c) shows the robustness to observation loss. When 20% of the sensor values are masked, the macro F1 remains at 0.928, while GCN drops to 0.879. Moreover, the factors indicate that the smart building IoT system may be unstable.

https://doi.org/10.64972/jaat.2023v1.305p19e:255-268

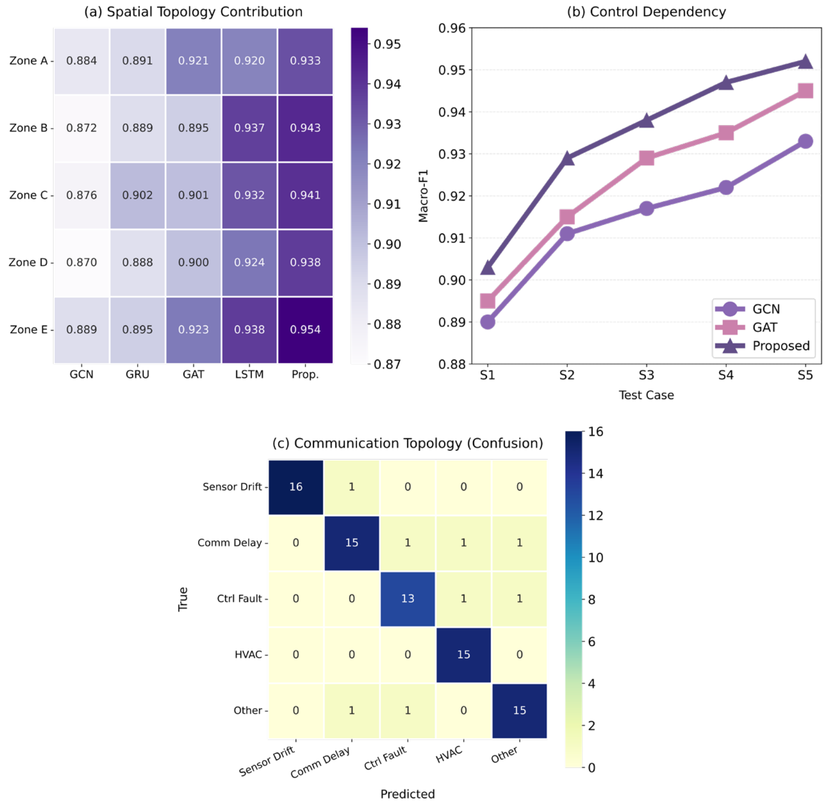


**Figure 5.** Topology contribution analysis. (a) Spatial topology contribution. (b) Control dependency contribution. (c) Communication topology contribution.

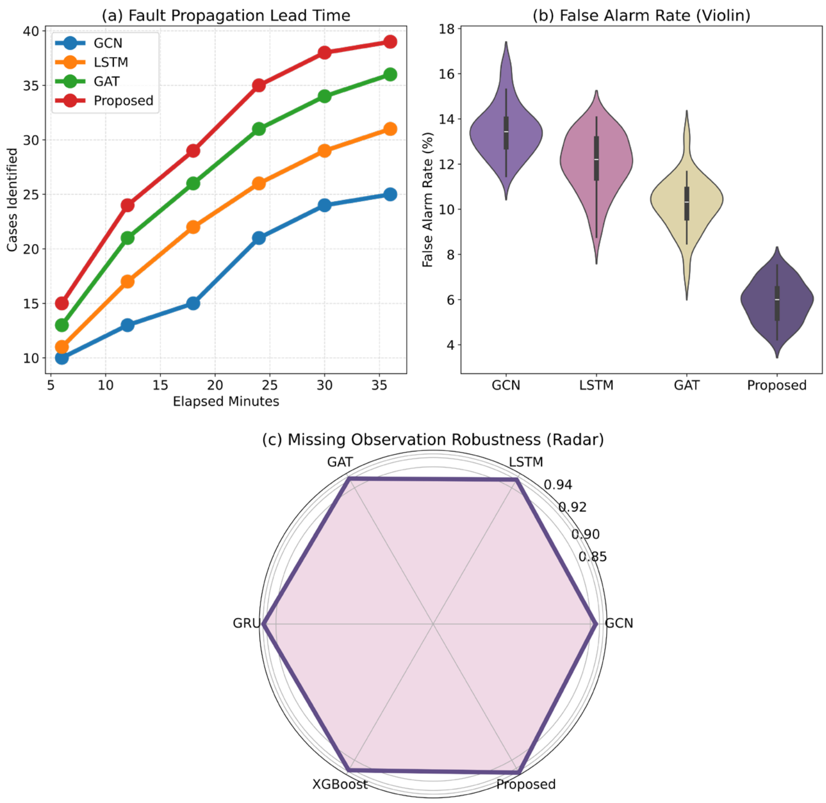


**Figure 6.** Fault propagation prediction under IoT disturbances. (a) Fault propagation lead time. (b) False alarm rate. (c) Missing observation robustness.

Propagation cases indicate that the model can identify faults before the target device reaches the alarm limit. In the HVAC system degradation event, the model first increased the risk of the air handling unit. Next are the downstream room temperature sensors, followed by the zone energy meters. If there is a significant deviation, the non-graph model can only determine the temperature of the abnormal area. Therefore, the topology-aware model will use predictive maintenance instead of post-failure detection. This behavior aligns with research on spatiotemporal graph prediction, which indicates that relational structures can locally predict state changes before they occur [35]. Sensor drift in the supply air temperature sensor is another such example. Adjust slowly, do not exceed 30 minutes. The model found inconsistencies in the downstream room temperature response, fan status, valve position, and drift sensor. Therefore, the warning will be issued before the local threshold alarm. The two downstream area sensors and the air handling unit controller are the closest neighbors in the attention distribution. The pattern can help maintenance personnel determine whether the issue is with the sensor or the entire HVAC system.

According to the missing observation experiment, topological gating is also effective. When a node has missing values, gating will increase the weight of the relevant devices. When the observation of a node is complete and consistent, the gate will reduce the influence of neighbors while retaining local evidence. The adaptive behavior explains why the proposed model is less prone to degradation in the absence of data compared to GCN. Fixed graph convolutions cannot distinguish between reliable local signals and unreliable signals, and they aggregate neighbors in the same way regardless of the quality of the observations.

### Ablation and Deployment Performance

The results of the ablation and deployment are shown in Figure 7. As shown in Figure 7(a), by removing the relationship-specific topology, the macro F1 score decreased from 95.2% to 92.6%. Topology gating reduces the robustness of missing observations, increasing the false alarm rate from 5.6% to 8.9%. The priority of early warning has been reduced, with the average lead time decreasing from 38.5 minutes to 29.4 minutes. According to Figure 7(b), the inference delay for each graph snapshot on the edge GPU is 21.8 milliseconds, while the delay on the industrial CPU is 64.3 milliseconds. Figure 7(c) shows the accuracy-latency trade-off. The deployment score of this model is relatively low, but it is slightly faster than LSTM and XGBoost [36].

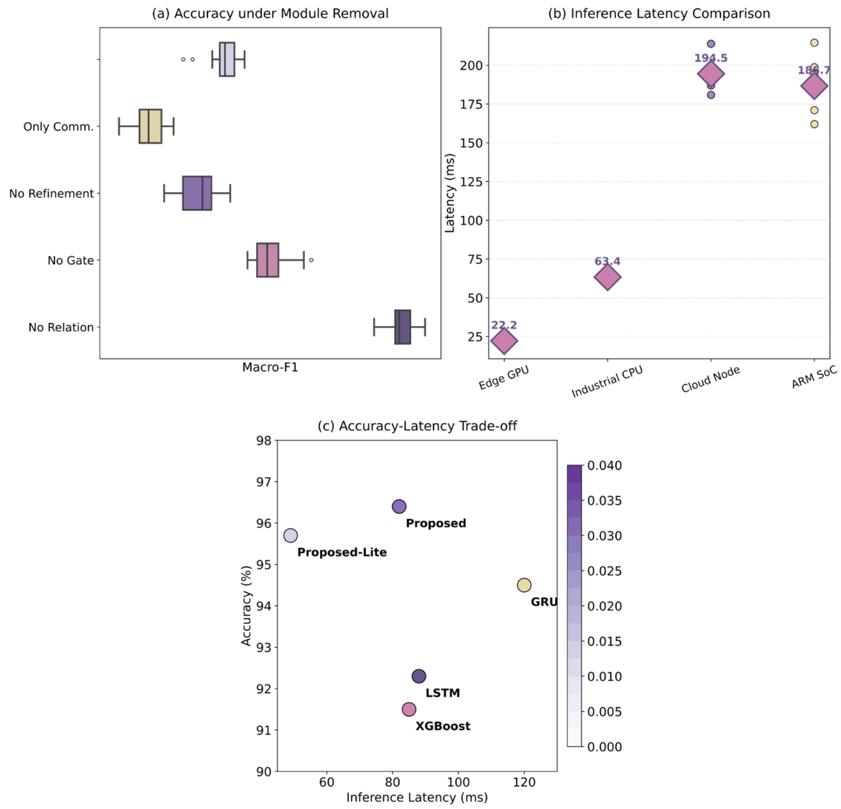


**Figure 7.** Ablation and deployment performance comparison. (a) Accuracy under module removal. (b) Inference latency comparison. (c) Accuracy-latency trade-off.

https://doi.org/10.64972/jaat.2023v1.305p19e:255-268

Deployment results indicate that topology-aware predictions can be applied in online smart building maintenance. Since building management systems typically use minute-level sampling frequency, graph inference times below 100 milliseconds will not cause bottlenecks. The model outputs device-level risk scores, as well as the corresponding neighbor attention weights. The above output helps maintenance engineers determine whether the warning is caused by control circuit, equipment, or communication failures. Surveys on explainable predictive maintenance indicate that the decisions of support systems must be understandable in practice [37].

Edge deployment also requires more memory. The serialized model is 8.7 MB, smaller than GRU (5.1 MB) and spatio-temporal GNN (13.4 MB). The relation-specific matrices and topological gates reduced the size. Although it is more expensive than a pure sequence baseline, the gateway equipped with edge GPUs or industrial CPUs is worth the extra memory cost due to its accuracy and shorter lead time. For very small embedded controllers, the graph can be divided into floors or areas. Doing so can reduce the number of nodes per inference.

Fault analysis identified the following two issues. First, when HVAC degradation and communication delays occur simultaneously, it may be considered sensor drift. Secondly, building renovations or controller reconfigurations may cause the topological metadata to become outdated. Inductive graph learning research indicates that graph models can adapt to new nodes; however, in practical construction, topological validation and regular metadata updates are necessary [38]. Therefore, future deployments will use topology-aware GNN predictions, build digital twins, and implement automatic topology discovery and maintenance feedback. In addition, Through the study of graph anomaly detection, it is necessary to combine structural learning with temporal evidence [39]. For maintenance, early alerts should be associated with work order priorities and expected energy impact, rather than just classification labels [40]. Ablation results also indicate that not all edges are important. Spatial edges help identify regional-level thermal anomalies, communication edges predict gateway delays, and control dependency edges predict HVAC deterioration. The relevant edges are not complex, but they are very sensitive to seasonal changes. Therefore, in the future, the update frequency of related edges should be lower than that of engineering topology edges. A practical system can fix space, control, and communication edges, and only update associated edges after major seasonal changes; the proposed model also changes the concept of maintenance priority. Ranking is not based solely on the current size of the anomaly, but rather on the predicted risk of the device, topological exposure, and expected propagation effects. For example, due to its impact on multiple downstream areas, an air handling unit that is slightly out of specification may be more valuable than a severely damaged room sensor. This is closer to the reasoning of actual maintenance and considers the system impact along with local severity. Therefore, the risk output chart can be linked to the maintenance plan and forecast results.

Inaccurate reports do not meet the operational conditions of conservative warnings. Brief instabilities, such as communication congestion or controller resets, can lead to some warnings that do not log faults. Although the new model reduces false positives, medium-risk warnings will still be issued due to inconsistencies among multiple related devices. In actual operation, these warnings will be moved to a low-priority inspection queue instead of being immediately addressed. In contrast, the binary alarm mode is more suitable for the operation of smart buildings.

## Conclusion

This study proposes a topology-aware GNN technique to enhance the ability to predict equipment failures in smart building IoT systems. A method for describing a building as a heterogeneous graph with spatial adjacency, control dependencies, communication relationships, and temporal correlations is as follows. By using topology-gated relationship-specific graph aggregation to simulate the influence and fault propagation between devices, the local sequence of device states is preserved.

Experimental analysis shows that compared to time series models, boosting models, and standard GNN baselines, the proposed method demonstrates improvements in accuracy, macro F1, AUC, early warning lead time, and robustness to missing sensor observations. The above results indicate that topological information is more suitable for diagnosing issues in HVAC equipment, controllers, and gateways. Moreover, in cases where local device observation is limited, topological information is also more suitable for diagnosing other situations. The inference latency of this model is relatively small for edge-assisted building management systems.

In the future, the framework can be expanded Through dynamic topology updates, automatic graph creation from building automation logs, and integration with digital twin models. The predicted device risks, along with the associated energy consumption and comfort reduction, can aid the maintenance decision-making process. In light of the above, the current graphical quality of these systems is also an issue. Therefore, future systems need to integrate topology validation, sensor metadata cleaning, automatic relationship discovery, and operator feedback. Edge-cloud collaboration can also support cross-construction of fault modes, periodic cloud retraining, and rapid local alerts.

**Author Contributions**

Agnieszka Szymanski contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Weronika Czarnecki contributes to methodology, software, validation, analysis, investigation. All authors have read and agreed with the manuscript before its submission and publication.

**Funding**

This research received no specific financial support from any funding agency.

**Institutional Review Board Statement**

Not applicable.

## References

[1] Mirnaghi, M. S., & Haghighat, F. (2020). Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. Energy and Buildings, 229, 110492. https://doi.org/10.1016/j.enbuild.2020.110492

[2] Taheri, S., Ahmadi, A., Mohammadi-Ivatloo, B., & Asadi, S. (2021). Fault detection diagnostic for HVAC systems via deep learning algorithms. Energy and Buildings, 250, 111275. https://doi.org/10.1016/j.enbuild.2021.111275

[3] Ng, K. H., Yik, F. W. H., Lee, P., Lee, K. K. Y., & Chan, D. C. H. (2020). Bayesian method for HVAC plant sensor fault detection and diagnosis. Energy and Buildings, 228, 110476. https://doi.org/10.1016/j.enbuild.2020.110476

[4] Almaguer-Angeles, F., Murphy, J., Murphy, L., & Portillo-Dominguez, A. O. (2019). Choosing machine learning algorithms for anomaly detection in smart building IoT scenarios. 2019 IEEE 5th World Forum on Internet of Things, 491-495. https://doi.org/10.1109/WF-IOT.2019.8767357

[5] Ebrahimifakhar, A., Kabirikopaei, A., & Yuill, D. (2020). Data-driven fault detection and diagnosis for packaged rooftop units using statistical machine learning classification methods. Energy and Buildings, 225, 110318. https://doi.org/10.1016/j.enbuild.2020.110318

[6] Boi-Ukeme, J., & Wainer, G. (2021). A workflow for data-driven fault detection and diagnosis in buildings. 2021 Winter Simulation Conference, 1-12. https://doi.org/10.1109/WSC52266.2021.9715464

[7] Chatterjee, A., & Ahmed, B. S. (2022). IoT anomaly detection methods and applications: A survey. Internet of Things, 19, 100568. https://doi.org/10.1016/j.iot.2022.100568

[8] Utomo, D., & Hsiung, P. A. (2019). Anomaly detection at the IoT edge using deep learning. 2019 IEEE International Conference on Consumer Electronics - Taiwan, 1-2. https://doi.org/10.1109/ICCE-TW46550.2019.8991929

[9] Sahu, N. K., & Mukherjee, I. (2020). Machine learning based anomaly detection for IoT network. 2020 4th International Conference on Trends in Electronics and Informatics, 787-794. https://doi.org/10.1109/ICOEI48184.2020.9142921

[10] Sadikin, F., & Kumar, S. (2020). ZigBee IoT intrusion detection system: A hybrid approach with rule-based and machine learning anomaly detection. Proceedings of the 5th International Conference on Internet of Things, Big Data and Security, 57-68. https://doi.org/10.5220/0009342200570068

[11] Choi, H., Kim, D., Kim, J., Kim, J., & Kang, P. (2022). Explainable anomaly detection framework for predictive maintenance in manufacturing systems. Applied Soft Computing, 125, 109147. https://doi.org/10.1016/j.asoc.2022.109147

[12] Wang, Z., Hong, T., & Piette, M. A. (2020). Building thermal load prediction through shallow machine learning and deep learning. Applied Energy, 263, 114683. https://doi.org/10.1016/j.apenergy.2020.114683

[13] Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S. (2022). Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. Journal of Building Engineering, 45, 103406. https://doi.org/10.1016/j.jobe.2021.103406

[14] Qiao, Q., Yunusa-Kaltungo, A., & Edwards, R. E. (2022). Feature selection strategy for machine learning methods in building energy consumption prediction. Energy Reports, 8, 13621-13654. https://doi.org/10.1016/j.egyr.2022.10.125

[15] Liu, Y., Ma, Y., Mao, Z., Jiang, B., Liu, J., & Bian, J. (2021). TD-GAT: Graph neural network for fault diagnosis knowledge graph. 2021 China Automation Congress, 1734-1739. https://doi.org/10.1109/CAC53003.2021.9728175

[16] Chen, D., Liu, R., Yu, W., Zhang, K., Pu, Y., & Cao, D. (2021). Fault diagnosis of industrial control system with graph attention network on multi-view graph. 2021 5th Asian Conference on Artificial Intelligence Technology, 617-623. https://doi.org/10.1109/ACAIT53529.2021.9731197

[17] Xie, L., Pi, D., Zhang, X., Chen, J., Luo, Y., & Yu, W. (2021). Graph neural network approach for anomaly detection. Measurement, 180, 109546. https://doi.org/10.1016/j.measurement.2021.109546

[18] Qin, L., Wei, W., & Ma, Y. (2021). FlowDiviner: Spatio-temporal network traffic prediction method based on graph neural network. 5th Asia-Pacific Workshop on Networking, 1-2. https://doi.org/10.1145/3469393.3469671

[19] Jiang, W., & Luo, J. (2022). Graph neural network for traffic forecasting: A survey. Expert Systems with Applications, 207, 117921. https://doi.org/10.1016/j.eswa.2022.117921

[20] Ali, M. A., Venkatesan, S., Liang, V., & Kruppa, H. (2021). TEST-GCN: Topologically enhanced spatial-temporal graph convolutional networks for traffic forecasting. 2021 IEEE International Conference on Data Mining, 982-987. https://doi.org/10.1109/ICDM51629.2021.00110

[21] Nikolentzos, G., Dasoulas, G., & Vazirgiannis, M. (2020). k-hop graph neural networks. Neural Networks, 130, 195-205. https://doi.org/10.1016/j.neunet.2020.07.008

[22] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4-24. https://doi.org/10.1109/TNNLS.2020.2978386

[23] Huang, B., Wang, X., Cui, P., Jiang, W., Yang, Y., & Fan, Q. (2021). One-class temporal graph attention neural network for dynamic graph anomaly detection. 2021 2nd International Conference on Electronics, Communications and Information Technology, 783-790. https://doi.org/10.1109/CECIT53797.2021.00142

[24] Barceló, P., Kostylev, E. V., Monet, M., Pérez, J., Reutter, J. L., & Silva, J. P. (2020). The expressive power of graph neural networks as a query language. ACM SIGMOD Record, 49(2), 6-17. https://doi.org/10.1145/3442322.3442324

[25] Chaouch, H., Çeken, C., & Arı, S. (2021). Energy management of HVAC systems in smart buildings by using fuzzy logic and M2M communication. Journal of Building Engineering, 44, 102606. https://doi.org/10.1016/j.jobe.2021.102606

[26] Feng, C., Mehmani, A., & Zhang, J. (2020). Deep learning-based real-time building occupancy detection using AMI data. IEEE Transactions on Smart Grid, 11(5), 4490-4501. https://doi.org/10.1109/TSG.2020.2982351

[27] Deorankar, A. V., & Thakare, S. S. (2020). Survey on anomaly detection of IoT cyberattacks using machine learning. 2020 Fourth International Conference on Computing Methodologies and Communication, 115-117. https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00023

[28] Essop, I., Ribeiro, J. C., Papaioannou, M., Zachos, G., Mantas, G., & Rodriguez, J. (2021). Generating datasets for anomaly-based intrusion detection systems in IoT and industrial IoT networks. Sensors, 21(4), 1528. https://doi.org/10.3390/s21041528

[29] Aversano, L., Bernardi, M. L., Cimitile, M., & Pecori, R. (2021). Anomaly detection of actual IoT traffic flows through deep learning. 2021 20th IEEE International Conference on Machine Learning and Applications, 1736-1741. https://doi.org/10.1109/ICMLA52953.2021.00275

[30] Zhang, L., & Leach, M. (2021). Evaluate the impact of sensor accuracy on model performance in data-driven building fault detection and diagnostics using Monte Carlo simulation. Building Simulation, 15(5), 769-778. https://doi.org/10.1007/s12273-021-0833-4

[31] Richardson, W. B., Meyer, J., & Von Solms, S. (2021). Towards machine learning and low data rate IoT for fault detection in data driven predictive maintenance. 2021 IEEE World AI IoT Congress, 202-208. https://doi.org/10.1109/AIIOT52608.2021.9454190

[32] Wu, Y., Maravelias, C. T., Wenzel, M. J., ElBsat, M. N., & Turney, R. T. (2021). Predictive maintenance scheduling optimization of building heating, ventilation, and air conditioning systems. Energy and Buildings, 231, 110487. https://doi.org/10.1016/j.enbuild.2020.110487

[33] Sinha, A., & Das, D. (2021). sCalib: A warehouse sensor fault detection and self-calibration technique for sustainable IoT. 2021 IEEE 18th India Council International Conference, 1-6. https://doi.org/10.1109/INDICON52576.2021.9691676

[34] Liu, Z., & Tan, H. (2021). Traffic prediction with graph neural network: A survey. CICTP 2021, 467-474. https://doi.org/10.1061/9780784483565.046

[35] Roy, A., Roy, K. K., Ali, A. A., Amin, M. A., & Rahman, A. K. M. M. (2021). Unified spatio-temporal modeling for traffic forecasting using graph neural network. 2021 International Joint Conference on Neural Networks, 1-8. https://doi.org/10.1109/IJCNN52387.2021.9533319

[36] Yang, C. Y., Kuo, J. J., Sheu, J. P., & Zheng, K. J. (2021). Cooperative distributed deep neural network deployment with edge computing. ICC 2021 - IEEE International Conference on Communications, 1-6. https://doi.org/10.1109/ICC42927.2021.9500668

[37] Matzka, S. (2020). Explainable artificial intelligence for predictive maintenance applications. 2020 Third International Conference on Artificial Intelligence for Industries, 69-74. https://doi.org/10.1109/AI4I49448.2020.00023

[38] Majumdar, A. (2020). Graph transform learning. Neural Networks, 128, 248-253. https://doi.org/10.1016/j.neunet.2020.05.020

[39] Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection. ACM Computing Surveys, 54(2), 1-38. https://doi.org/10.1145/3439950

[40] Conte, A., Bolland, C., Phan, L., Brundage, M., & Sexton, T. (2021). Impact of data quality on maintenance work order analysis: A case study in historical HVAC maintenance work orders. PHM Society European Conference, 6(1), 11. https://doi.org/10.36001/phme.2021.v6i1.2814