

From Causal Discovery to Risk Warning: An Interpretable Heterogeneous Graph Neural Networks for Railway Dispatching Safety

Siyu Cao, Xuelei Meng, Hao Peng, Qianyu Wang, Doudou Wang, Tianshu Qi

Abstract—To improve the railway dispatching safety capabilities and enable early warning of sudden risk events, this paper proposes a linear non-Gaussian acyclic model with heterogeneous graph neural networks (LiNGAM-HetGNN). The proposed method addresses the issues of the complex interactions of multi-source heterogeneous data and the lack of causal interpretability in railway transport. Firstly, the model extracts asymmetric causal relationships from railway incident data among entities based on LiNGAM. A causal weight matrix is constructed to quantify causal relationships, and we develop the causal graph to show the propagation pathways of risk events. Secondly, we construct the Causal GAN to effectively tackle the problem of data scarcity associated with long-tail risk events. Then, we design an interpretable HetGNN and utilize GCN to capture spatial dependencies of stations, while GAT to model the interactions of trains and signals. The causal weight matrix is embedded into the convolutional layers of GCN and the attention mechanism of GAT, enabling causality-driven graph learning. Finally, a hierarchical risk classifier is developed to assess the risk level of each node in the network. Experimental validation using the U.S. Federal Railroad Administration accident dataset demonstrates the effectiveness of the method. The model achieves the precision, recall, and F1-score of 98.1%, 98.7%, and 98.5%, respectively. Ablation studies further confirm the essential role of the causal discovery in the model. This method uncovers causal relationships among various risk events, supports early warning and causal traceability in railway dispatching progress, and provides practical value for intelligent railway transport development.

Index Terms—Railway dispatching safety; Causal discovery; LiNGAM; HetGNN; Risk warning

Manuscript received June 10, 2025; revised August 14, 2025.

This work was supported in part by the Science and Technology Plan Funding of Gansu Province(24JRRA865) and Central Government Guided Local Science and Technology Development Fund Project(25ZYJA015).

Siyu Cao is a postgraduate student at the School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China (e-mail: 11230018@stu.lzjtu.edu.cn).

Xuelei Meng is a professor of Traffic and Transportation School, Lanzhou Jiaotong University, Lanzhou 730070, China (corresponding author, phone: +86-15117071015; fax: 86-0931-4938025; e-mail: mxl@mail.lzjtu.cn).

Hao Peng is a postgraduate student at the School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China (e-mail: 11230016@stu.lzjtu.edu.cn).

Qianyu Wang is a postgraduate student at the School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China (e-mail: 12231050@stu.lzjtu.edu.cn).

Doudou Wang is a postgraduate student at the School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China (e-mail: 12231060@stu.lzjtu.edu.cn).

Tianshu Qi is a postgraduate student at the School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China (e-mail: 12230991@stu.lzjtu.edu.cn).

I. INTRODUCTION

As the backbone of modern integrated transport systems, railway transport plays a critical role in maintaining network stability and ensuring the safety of both passenger and freight operations [1]. In recent years, the rapid expansion of high-speed rail network and the large-scale deployment of trains have significantly increased the complexity of railway dispatching system. The system is now characterized by tightly coupled entities, multi-dimensional heterogeneity, and intricate risk propagation pathways. This complexity is further exacerbated by the rising demand for coordinated control across multiple stations and lines, placing increasing pressure on dispatching operations [2,3]. However, current risk management methods primarily rely on rule-based systems and expert expertise, which struggle to provide the real-time responsiveness and interpretability required for effective risk warning [4]. Specifically, within high-density railway operations, traditional methods fail to accurately identify potential risks and provide timely early warnings, resulting in dispatching decisions may exceed allowable safety response times [5].

Railway dispatching faces multiple challenges, including mixed operation of various train types, network collaborative control, and environmental influence [6]. During operations, many potential risks remain latent and difficult to detect. In scenarios involving frequent interactions among railway entities, incidents such as dispatching conflicts, track occupancy violations, and signal control failures become particularly prominent. These risks are especially severe at convergence points and on shared track segments, where multiple trains operate simultaneously. In such situations, dispatching delays or signal anomalies could rapidly escalate into serious incidents. Risk events tend to propagate through causal pathways, potentially resulting in equipment damage, service interruptions, and cascading system failures, posing a significant challenge to the safety and reliability of railway transport [7]. Consequently, accurately identifying critical risk points, assessing the severity, and issuing timely risk warning have become priorities in railway dispatching safety research. Sari et al. [8] employed the FAHP to identify risk factors in urban rail and formulated resource allocation plans. Yan et al. [9] developed a risk identification matrix based on WBS-RBS to assess the risks in urban rail construction. Recently, Lin et al. [10] investigated the safety issues on shared passenger and freight corridors, and proposed a semi-quantitative risk analysis framework to evaluate the probability and impact of incidents on adjacent tracks.

In recent years, data-driven methods have been increasingly applied in the railway transport [11]. In particular, machine learning (ML) has demonstrated superior performance in tasks such as risk identification and fault diagnosis [12,13]. For instance, Su et al. [14] employed a multi-layer perceptron with supervised learning to extract associations between vehicle failure instances and optimal solutions for different failure scenarios. Similarly, Liu et al. [15] focused on detection and warning of abnormal passenger flow in urban rail transit by utilizing a depth-first search algorithm to identify and respond to emergencies.

GNN has shown strong capabilities in processing complex graph-structured data. Xue et al. [16] proposed a deep learning framework to predict risk pathways and utilized graph features to address safety issue related to operational interruptions caused by fault events. Bi et al. [17] studied the flood resilience of urban rail transit by incorporating physical network characteristics, flood scenarios, and recovery resources into a complex network model, and analyzed operational performance under different disaster risk conditions. More recently, HetGNN has been introduced to better adapt to the actual scenario involving multiple types of nodes and edges. Jiao et al. [18] developed a disaster chain network for urban rail transit based on network vulnerability characteristics of the nodes and edges, and proposed targeted chain-breaking strategies for disaster mitigation.

However, traditional ML methods primarily relies on statistical correlations, limiting the ability to uncover causal relationships among variables. This lack of causal insight reduces the model interpretability and effectiveness of decision-making. In contrast, causal analysis facilitates the identification of causal dependencies within data, enabling the discovery of root causes of risk events and improving model transparency. Kim et al. [19] proposed a system for analyzing human errors in railway operations by utilizing predefined links between contextual information and causal factors to identify multi-level error sources and their causal relationships. Hadj-Mabrouk et al. [20] proposed a feedback-driven approach to extract and analyze incident data, aiming to identify safety incidents and their underlying causes.

Causal analysis is particularly effective for handling complex variables, as it can identify causal relationships that are difficult to discover through statistical methods alone. In addition to identifying these dependencies, causal discovery reveals the underlying mechanism of risk propagation and provides an interpretable model. This approach facilitates root cause tracing and the mapping of risk propagation pathways, thereby enhancing early risk warning. Shi et al. [21] investigated risk propagation pathways, trigger probability and risk level of urban rail transit, proposed a risk chains mining method to identify the association between risk sources and chains based on path search theory. Belhour et al. [22] conducted a post-fault analysis in railway maintenance to uncover incident causes. Wu et al. [23] analyzed causal pathways originating from root driving factors and proposed strategies to manage the interactions among various risk factors in urban rail transit.

Causal model with structural interpretability could effectively uncover the underlying mechanism of risk events and their systemic impact, supporting accurate risk warning and root cause analysis. Specifically, by constructing a causal

graph, quantifying the influence of risk events on railway dispatching system. Cao et al. [24] proposed a YOLOv8n-LiteCBAM model that integrated a lightweight DepthStackNet backbone with pruning techniques and BiCBAM to accelerate reasoning for defect detection. Wang et al. [25] proposed a method to identify urban rail transit incident factors and improve safety management capabilities by constructing an incident semantic network of the risk control chain to evaluate the impact degree of critical items.

Existing research on railway risks has made progress, however, challenges still remain in handling multi-source heterogeneous data and the causal propagation of risks. To address research gap, this paper proposes a method that integrates causal discovery with HetGNN, which provides a deep representation of risk propagation in dispatching system and addresses the interpretability limitations of traditional ML. The main contributions can be summarized as follows:

(1) Proposing a causal model based on LiNGAM for railway dispatching risk events. By constructing the causal weight matrix, the model uncovers the asymmetric causal dependencies among entities, enabling accurate detection risk events and the propagation pathways, and provides a causal foundation for risk tracing and early warning.

(2) Designing the HetGNN tailored to represent dispatching entities as heterogeneous nodes. By integrating topological structures and feature information, the network employs GCN and GAT to model the interactions among different entities types, thereby capturing the characteristics and structural of risk propagation pathways.

(3) The causal weight matrix is embedded into HetGNN to guide the convolution operations of GCN and the attention mechanism of GAT. This integration allows the model to prioritize node relationships with causal relationships during training, thereby enhancing the ability to identify critical risk events and improving the interpretability of the model.

II. PROBLEM ANALYSIS

As the core of railway transport, dispatching depends on the coordinated operation of multiple critical entities (e.g., stations, trains, and signals), as well as their interactions (e.g., dispatch commands, track occupancy, and signal transmission). This process involves a variety of entity types, intricate interaction structures, and heterogeneous data attributes, all of which contribute to the risk coupling and causal propagation of incidents. Specifically, risks often originate from localized incidents, gradually accumulate, and propagate through risk propagation pathways. These risks may ultimately evolve into systemic incidents through causal pathways. Therefore, effective early warning of risks could accurately identify potential risks amid high traffic density and operational complexity, providing timely warning by quantifying causal influence.

To model the complex interactions among multiple entities in railway dispatching, this study represents the railway transport network as a heterogeneous graph $G = (V, E, A)$, where $V = \{v_s, v_t, v_x\}$ denotes the set of nodes. Station nodes v_s are classified into passenger and freight stations. Passenger stations manage the arrival, departure, and boarding of passenger trains, with attributes such as yard capacity and train frequency. Freight stations handle the

loading and unloading of freight trains, with attributes such as loading capacity and operational duration. Train nodes v_t represent individual train entities, the attributes include train speed, train type, etc. Signal nodes v_x control train movement permissions, the attributes include signal status and control range, etc. Incorrect commands issued by signals can result in dispatching conflicts, or even collisions.

In railway transport network, trains operate between stations and signals, and frequently interacting with both. $E = \{e_{s-s}, e_{s-t}, \dots, e_{x-s}\}$ denotes the set of edges. Specifically, physical connection edges e_{s-s} represent railway tracks that connect stations, forming the fundamental topological structure of the network. Dispatch coordination edges e_{s-t} represent interactions between stations and trains, including train arrivals, departures, and dwelling time, driven by dispatching demands. Resource contention edges e_{t-t} denote conflicts between trains competing for shared track segments or signal-controlled blocks, implicitly dispatch priority competition. Control dependency edges e_{t-x} define the regulatory relationship between trains and signals, whether a train is accessed to a track segment depends on the status of the preceding signal. Signal relay edges e_{x-x} represent the information transmission between signals to maintain consistency and continuity in dispatch commands. Dispatch interaction edges e_{x-s} denote the coordination between stations and signals to manage train entry and departure sequences, aligning dispatch instructions with station resource availability. The connectivity among entities is represented by an adjacency matrix A .

$$A = \begin{cases} a_{ij} = 1 \\ a_{ij} = 0 \end{cases} \quad (1)$$

$X_i^v (i = 1, 2, \dots, m)$ denote the feature vector of node V , and $X_i^e (i = m + 1, m + 2, \dots, n)$ denote the feature vector of edge E . We define a risk level function $R = f(V, E, A, X_i^v, X_i^e)$, which integrates structural properties and attribute features of the nodes and edges, as illustrated in Table I. This level reflects the degree of risk associated with nodes and edges within the dispatching process.

Table I

RISK LEVEL CLASSIFICATION

Risk Level	Illustrate
High	Risks with significantly impact on dispatching operations and are deemed operationally unacceptable. Immediate mitigation measures are required.
Medium	Risks with moderate impact on dispatching operations. These require further assessment based on operational cost and potential hazards to determine whether they are acceptable.
Low	Risks with minimal impact on dispatching operations. No additional intervention is necessary.

To illustrate the application of heterogeneous graph in railway transport network, Figure 1 shows the interaction relationships among entities in a dispatching scenario. Red circular nodes (PS1-PS8) represent passenger stations, while orange circular nodes (FS1-FS5) denote freight stations. Blue circular nodes (RS1-RS13) correspond to signals. Yellow train icons (Train 1 and Train 2) represent freight trains, and green train icons (Train 3 and Train 4) represent passenger trains. Edges are colored differently to distinguish different types of interactions. Black edges represent physical railway tracks connecting adjacent stations (e.g., PS1-PS2, PS6-PS8), forming the core topological structure. Brown edges represent signal transmission between stations and signals (e.g., PS2-RS1). Gray edges represent control dependencies between trains and signals (e.g., Train 1-RS9). Dark blue edges represent dispatching relationships between trains and stations (e.g., Train 2-RS8).

During operation at a time step $[T_1, T_2]$, dynamic changes in train states and scheduling conditions may expose the system to the following three major categories of risk: (1) Route conflict: A potential route conflict may occur at the convergence node PS2 between Train 2 and Train 3. If scheduling priorities are misassigned or signal responses are delayed, the system faces a heightened risk of train conflicts, including rear-end collisions. (2) Station overcapacity: Each station has a limited yard capacity. If Train 2 and Train 4 arrive simultaneously at the same station (e.g., PS1) without timely adjustments, track occupancy conflicts may arise, leading to train delays and scheduling bottlenecks. (3) Signal control failure: Malfunctions in signals (e.g., RS6 or RS7) may cause Train 4 entering unauthorized or restricted sections. Such failures increase the risk of derailments or other critical incidents.

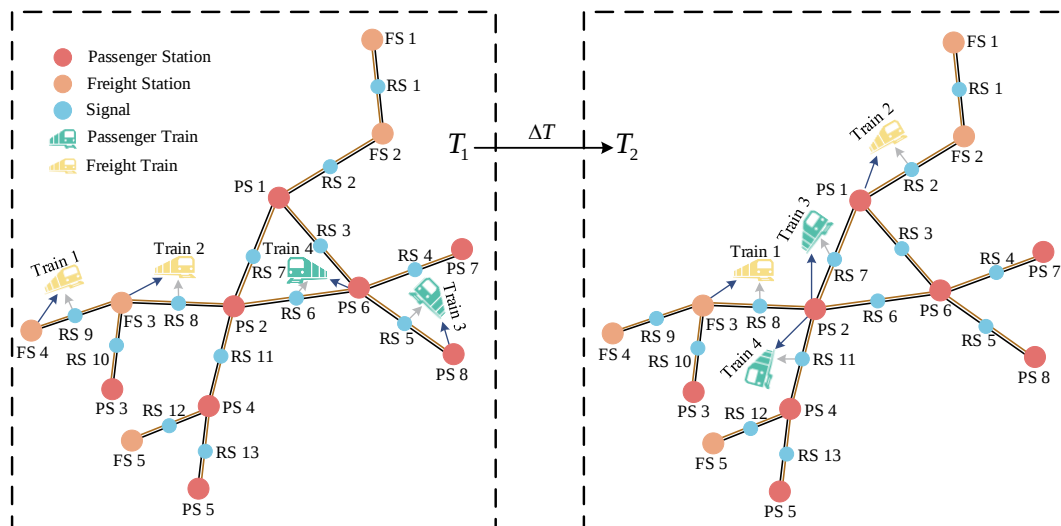


Fig.1. Train operation process of railway network

III. METHODS

A. Model Architecture

To enable effective risk warning in railway diapatching, this paper proposes a three-layer HetGNN model based on causal discovery, as illustrated in Figure 2. The model comprises three components: the causal discovery layer, the HetGNN layer, and the risk warning layer.

In causal discovery layer, the model extracts risk-related variables from incident data and applies LiNGAM algorithm to identify causal relationships among entities. These relationships are used to construct a directed acyclic graph (DAG), along with the causal weight matrix that quantifies causal influence. In HetGNN layer, the railway transport network is represented as a heterogeneous graph consisting of diverse node and edge types. This layer integrates GCN and GAT to model structural interactions among entities, incorporating topological structure and attribute features. In risk warning layer, the model utilizes the causal information to evaluate the risk level of each entity, thereby enabling timely detection and early warning for potential risk events throughout the dispatching process.

B. Risk Indicators

The data used in this study for railway dispatching are collected from multiple sources, as summarized in Table II. To systematically characterize risk events in the dispatching process, a comprehensive risk indicator system is developed based on publicly available accident statistics reported by the U.S. Federal Railroad Administration (FRA) [26], as shown in Table III. This indicator system categorizes risks into six primary domains, including train-related risks, station-related risks, and signal-related risks, etc. Each category is further decomposed into a set of quantifiable secondary risk indicators, providing a structured and detailed framework for analyzing and assessing risk factors in railway dispatching.

Table II
RAILWAY DISPATCHING DATA SOURCES

Data source	Data categories
Dispatching log	Train arrival and departure time, station throughput, signal status transition, etc.
Train operation	Train speed, acceleration, trajectories, headway intervals, etc.
Equipment monitoring	Signal status, track occupancy status, communication system failure, etc.
External environment	Weather conditions, construction activity, unexpected events, etc.

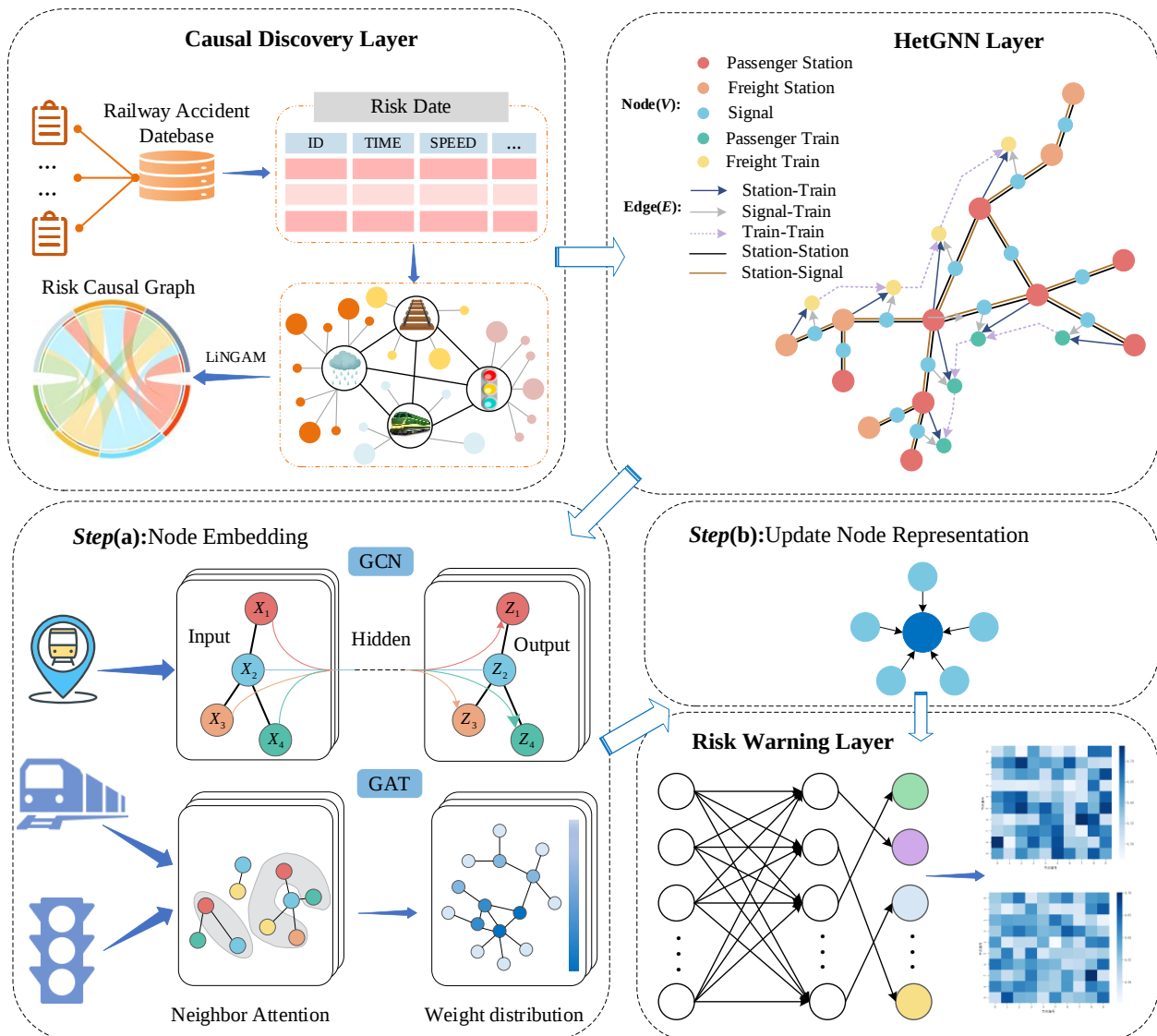


Fig.2. Architecture of the LiNGAM-HetGNN

Table III
Risk Indicator System for Railway Dispatching

Risk category	Key risk events
Train risks A	Train derailment A_1
	Train collision A_2
	Departure delay A_3
	Arrival delay A_4
	Abnormal wheel-rail contact A_5
	Brake system malfunction A_6
	Overspeed operation A_7
Station risks B	Station equipment failure B_1
	Passenger fall from platform B_2
	Improper handling of hazardous materials B_3
	Delay in passenger boarding/alighting B_4
	Delay in freight loading/unloading B_5
	Obstruction in station passageway B_6
	Inadequate worksite protection B_7
Signal equipment risks C	Signal response delay C_1
	Signal loss C_2
	Absence of signal display C_3
	Signal display error C_4
	Signal power outage C_5
	Signal interference C_6
	Signal obstruction C_7
Train-signal interaction risks D	Delay in signal confirmation feedback D_1
	Signal interference at track junction D_2
	Signal priority conflict D_3
	Failed to stop at signal D_4
	Inability to confirm train position D_5
	Train communication interruption D_6
	Delay in signal switching D_7
Train-station interaction risks E	Track occupancy conflicts at station E_1
	Excessive train-platform gap E_2
	Multi-train convergence congestion E_3
	Improper speed adjustment on arrival E_4
	Train stopping position deviation E_5
	Incorrect track assignment E_6
	Insufficient clearance near trackside equipment E_7
Systemic risks F	Power supply instability F_1
	Command issuance error F_2
	Extreme weather impact F_3
	System communication failure F_4
	Network transmission delay F_5

C. LiNGAM

Railway dispatching risks are characterized by structural complexity and dynamic temporal evolution. These risks typically arise from the interaction of multiple factors, with propagation mechanism exhibit strong causal dependencies. Consequently, a localized risk event may escalate through causal pathways among entities, leading to cascading failures and systemic disruptions. Therefore, accurately identifying these causal relationships is essential for improving risk warning capabilities and optimizing dispatching strategies.

To address the challenge of uncovering risk propagation in railway dispatching system, this study employs LiNGAM to uncover the causal relationships, it offers several notable advantages: (1) It does not rely on external interventions, enabling causal directed edges to be inferred directly from observational data. (2) It is well-suited for high-dimensional and structurally heterogeneous systems, consistent with the multi-entity and multi-relation nature of railway dispatching. (3) By assuming non-Gaussian and independent noise in the data, LiNGAM effectively resolves the issue of causal direction symmetry and effectively avoid causal inversion.

The proposed model is constructed based on the following assumptions:

- (1) Risk events in railway dispatching exhibit linear causal relationships.
- (2) The error terms associated with risk variables follow a non-Gaussian distribution.

Specifically, for an observed dataset $X = [X_1, X_2, \dots, X_d]$ containing d risk variables, their causal dependencies can be modeled using a linear structural equation model defined as:

$$X = BX + E \quad (2)$$

where $B \in \mathbb{R}^{d \times d}$ is the causal weight matrix, constrained to be strictly lower triangular (i.e., $b_{ij} = 0$ for $i \leq j$), ensuring each variable is influenced only by variables with smaller indices. E represents independent non-Gaussian noise terms.

LiNGAM aims to estimate B from X , thereby uncovering the directed causal structure among the risk events. This method employs independent component analysis (ICA) for causal discovery. The procedure steps are as follows.

To mitigate the influence of system drift and external disturbances, the raw observations data are mean-centered so that each variable has zero mean:

$$X' = X - E(X) \quad (3)$$

where X' is the centered value of variable X , and $E(X)$ denotes its sample mean.

Secondly, to determine the directionality of the causal relationships, ICA is applied to decompose X' into a set of non-Gaussian independent components E . This process obtains the transformation matrix C .

$$X' = (I - B)^{-1} E = CE \quad (4)$$

Assuming linear relationships among components and leveraging their non-Gaussianity, the causal directions can be identified. The causal weight matrix B is computed by inverting C .

$$B = I - C^{-1} \quad (5)$$

where I is the identity matrix.

To estimate the specific causal effects among variables X_i , a linear regression model is constructed for each risk variable to calculate its influence strength by the variable X_j .

$$X_i = \sum_{j < i} b_{ij} X_j + e_i \quad (6)$$

where b_{ij} denotes the causal effect coefficient of X_j on X_i , and e_i is an independent error term.

Given the high dimensionality and complex dependencies of risk variables in railway dispatching, LASSO regularization is applied to matrix B to obtain a sparse causal weight matrix B' . This approach mitigates overfitting and enhances the interpretability of the causal structure.

$$B' = \arg \min_B \left(\sum_{i=1}^d \left\| X_i - \sum_{j < i} b_{ij} X_j \right\|^2 + \lambda \sum_{i,j} |b_{ij}| \right) \quad (7)$$

where λ is the regularization parameter that controls the sparsity level of the matrix B' .

The B' is a strictly lower triangular matrix, ensuring that the causal structure conforms to the properties of a DAG. Risk propagation often involves complex interactions and cascading effects among multiple risk variables. These causal

relationships not only influence the evolution of individual events but can also escalate into system-wide incidents, affecting overall operational stability. To model these interdependencies, this study constructs a risk causal graph based on matrix B' . This graph represents the indirect risk propagation pathways among risk events, enabling the identification of critical risk nodes that serve as hubs in the risk propagation process.

Assume an initial risk event X_d triggers an intermediate variable X_e , which subsequently leads to a high-risk event X_f . This cascade can be represented as a causal pathway.

$$X_d \xrightarrow{b_{dc}} X_c \xrightarrow{b_{ce}} X_e \quad (8)$$

In practice, risk propagation is often driven by multiple concurrent pathways originating from diverse sources. To evaluate the influence of each propagation pathway and quantify its importance within the causal graph, we define the risk pathway propagation strength P_{Strength} . This metric measures the cumulative impact of causal interactions along a specific pathway and is calculated as the product of causal weights of edges in the pathway.

$$P_{\text{Strength}} = \prod_{(i,j) \in \text{Path}} b_{ij} \quad (9)$$

For the given risk propagation pathway $X_d \rightarrow X_e \rightarrow X_f$, its $P_{\text{Strength}} = b_{de} \times b_{ef}$, which reflects the degree to which risk is amplified or weakened during its propagation. This measure reveals the cumulative effect of causal interactions within the pathway, allowing for the track of risk propagation mechanism through the complex network.

To explicitly model the cascading effects of interacting risk events, we construct a DAG based on matrix B' . In this causal graph, nodes $\{d_1, d_2, \dots, d_n\}$ represent risk variables, edges b_{ij} represent the causal dependencies, and the non-zero elements of B' determine the presence and direction of the edge in causal graph.

Assume a submatrix of B' is given by:

$$B'_{\text{sub}} = \begin{bmatrix} 0 & 0 & 0 \\ 0.6 & 0 & 0 \\ 0.4 & 0.8 & 0 \end{bmatrix}$$

The corresponding risk propagation pathway is:

$$X_3 \xrightarrow{0.8} X_2 \xrightarrow{0.6} X_1 \quad (10)$$

According to matrix B' , the direct causal weight from X_3 to X_1 is 0.4. However, when considering the risk propagation pathways, with intermediate causal weights of 0.6 and 0.8, the pathway strength is $P_{\text{Strength}} = 0.6 \times 0.8 = 0.48$.

This result indicates that the cumulative influence of the indirect causal pathway exceeds the direct effect ($0.48 > 0.4$). Therefore, relying solely on direct causal weight may underestimate the true potential for risk propagation. To more accurately characterize the risk propagation mechanism, we propose a weighted average causal matrix $\hat{B}$ that incorporates both direct and indirect causal weight. The comprehensive causal weight matrix is defined through α that balance the direct and indirect pathways:

$$\hat{B} = \alpha b_{ij} + (1 - \alpha) P_{\text{Strength}} \quad (11)$$

D. Causal GAN

Railway dispatching data is often scarce and incomplete, limiting the effectiveness of data-driven methods for risk warning. However, GAN fail to account for causal relationships among variables, resulting in data that lacks interpretability and undermines the reliability of models.

To address this issue, this study proposes a Causal GAN guided by the causal graph derived from LiNGAM. By embedding the causal mechanisms of railway risks into prior constraints of the generator, the model produces high-quality synthetic data that adheres to underlying causal structure. Causal GAN not only enhances the accuracy of risk event modeling, but also uncovers causal chains within multi-source heterogeneous data, improving the interpretability of the model and the reliability of dispatching decisions.

The Causal GAN comprises two components: a generator and a discriminator. The innovation lies in incorporating the causal graph as a prior constraint on the generator. This ensures that the generated data conforms to the causal dependencies inherent in railway dispatching risk scenarios. The overall architecture is illustrated in Figure 3.

The generator is designed to synthesize high-quality data that conforms to the underlying causal structure. Unlike conventional generators that produce data directly from random noise, the Causal GAN generator leverages the causal graph to constrain the noise propagation, thereby ensuring consistency in the causal relationships in the generated data. The input to the generator consists of noise variables and a causal controller. Through the edge weights from the causal graph, the generator enforces structural constraints on the dependency patterns during data generation, ensuring that the output data adheres conforms to the causal pathways defined by the causal graph.

$$\tilde{X} = \hat{B}\tilde{X} + G(Z) + E \quad (12)$$

where, $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)^T$ denotes the generated variables, $Z = (z_1, z_2, \dots, z_n)^T$ represents an independent random noise vector, and $G(Z)$ is a nonlinear transformation function.

Subsequently, a MLP is employed to extract features from the input, with a causal weighting mechanism used to adjust the connection weights across layers. This design ensures that the generator not only synthesizes data resembling actual observations but also adheres to the risk propagation mechanisms inherent in the causal structure of the railway dispatching system. The final output is a set of multi-dimensional synthetic data samples representing railway dispatching risk events, which preserve the causal structure and distributional characteristics of real data.

To enforce consistency with causal priors, we incorporate an L1 regularization term into loss function of the generator:

$$L_{causal} = \lambda \sum_{i,j} \|b_{ij} (X_{gen,i} - f_j(X_{gen}))\|_1 \quad (13)$$

Here, $X_{gen,i}$ denotes the i variable in the generated sample, $f_j(\cdot)$ represents the causal function determining X_j based on its direct causes, and λ is the regularization coefficient.

The overall loss function as follows:

$$L_G = -E_{z \sim p(z)} [\log(D(X_{gen}))] + L_{causal} \quad (14)$$

The primary function of the discriminator is to distinguish between real and synthetic data, while also verifying the causal consistency of the generated samples to ensure the outputs align with the causal structure of railway dispatching risk. The discriminator receives both real samples X_{real} and generated samples X_{gen} , and utilizes the CNN to model the dependencies within the data. In addition, the discriminator incorporates a causal consistency checking function that computes the causal weight matrix $\hat{B}$ for the generated samples X_{gen} . The final output is a probability score $D(X) \in [0,1]$, indicating the authenticity and causal validity of the input data.

Causal consistency f is central to the design of the discriminator. To ensure the generated data reflects the causal relationships inherent in actual dispatching risks, we introduce a causal consistency loss function. This quantifies the discrepancy between the causal dependency matrix $Causal(X_{gen})$ estimated from the generated samples and the true causal influence matrix $\hat{B}$:

$$L_{consistency} = \mu E_{X_{gen}} [\|Causal(X_{gen}) - \hat{B}\|_F] \quad (15)$$

where, $Causal(X_{gen})$ denotes the causal matrix estimated from the generated data, μ is a weight coefficient, and $\|\cdot\|_F$ represents the Frobenius norm.

The loss function of the discriminator is defined as:

$$L_{Del} = E_{X_{real}} [\log D(X_{real})] + E_{X_{gen}} [\log(1 - D(X_{real}))] + L_{consistency} \quad (16)$$

where $E_{X_{real}} [\log D(X_{real})]$ reflects the ability of the discriminator to distinguish real from generated samples, and $E_{X_{gen}} [\log(1 - D(X_{real}))]$ reflects the effectiveness in detecting generated samples.

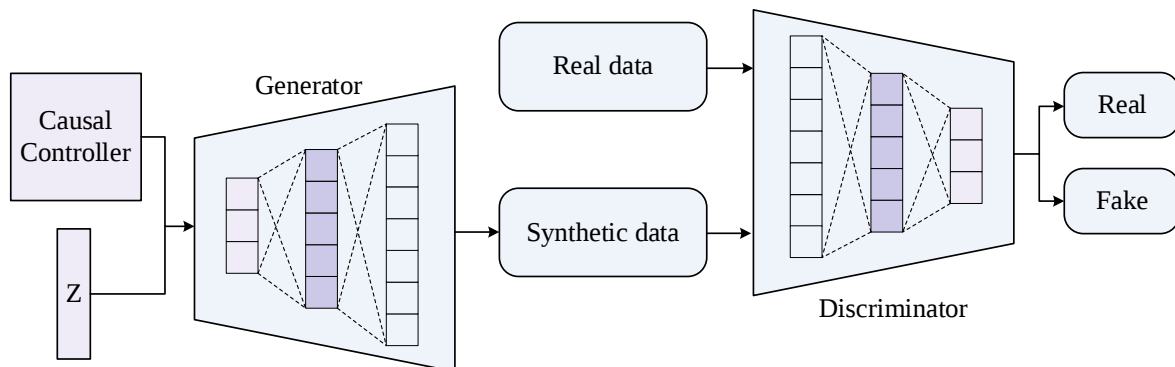


Fig.3. Architecture of Causal GAN

E. HetGNN

In railway dispatching network, entities engage in frequent multi-source information exchanges, characterized by significant structural heterogeneity. Meanwhile, various risk events exhibit complex causal dependencies. To address the challenges posed by diverse risk types and heterogeneity incident data, this study employs HetGNN to model the dispatching network, as illustrated in Figure 4. Furthermore, to improve the accuracy of risk propagation pathways modeling and causal influence analysis, the causal weight matrix $\hat{B}$ is integrated into the message-passing process of HetGNN. By embedding causal learning into the HetGNN architecture, the proposed model could capture the intricate associations among risk events within the railway dispatching and enhance risk detection and early warning capabilities for potential risk events.

E.1 GCN

In railway network, station nodes possess well-defined spatial positions and structural attributes, with adjacency relationships are determined by geographic layout and track connectivity. To capture spatial dependencies among stations, the GCN is employed to model the information transmission between station nodes. Specifically, GCN updates each node representations through layer-wise aggregation of features from its neighboring nodes, enabling each node to integrate information from an increasingly broader range of the network. For node embedding update at layer k , the features of each station node v_i neighbor set $\mathbb{N}(i)$ are normalized and weighted before aggregation. This process allows the model to progressively incorporates information from more distant neighbors and generates updated node representations through a nonlinear activation function.

To incorporate causal dependencies among risk events, we embed the causal weight matrix $\hat{B}$ into the standard adjacency structure. Since the diagonal elements of $\hat{B}$ are

zero (i.e., self-influences are not included), we add the identity matrix I to retain self-loop information, resulting in the enhanced adjacency matrix $\tilde{A} = \hat{B} + I$. The augmented matrix enables GCN to model not only topological connectivity but also the causal influences among risk variables during the propagation process, thereby improving the accuracy and interpretability of risk propagation.

$$h_i^{(k)} = \sigma \left(\sum_{j \in \mathbb{N}(i)} \frac{\tilde{A}_{ij}}{c_{ij}} W^{(k)} h_j^{(k-1)} \right) \quad (17)$$

where, $h_i^{(k)}$ denotes the embedding of station node v_i at layer k , σ an activation function, $\mathbb{N}(i)$ represents the set of neighbors of node v_i , c_{ij} is the normalization coefficient, and $W^{(k)}$ is the learnable weight matrix at layer k .

E.2 GAT

In contrast to station nodes, the interactions between trains and signals are more dynamic and complex. To effectively model these heterogeneous interactions, we adopt the GAT, which allows the model to adaptively learn the relative importance of neighboring nodes with respect to the target node, thereby improving the feature aggregation process.

In GAT, information propagation is governed by attention weights α_{ij} , which quantify the influence of the neighboring node v_j on the target node v_i . To better capture the underlying causal relationships in risk propagation, we incorporate the matrix $\hat{B}$ into the attention mechanism. Specifically, the enhanced causal matrix $\tilde{A}$ is used to reweight the attention coefficients, allowing the model to prioritize neighbors based on the causal influence. The modified attention mechanism is defined as follows:

$$\alpha_{ij} = \frac{\exp \left(\text{LeakyReLU} \left(a^T \left[W^{(k)} h_i^{(k-1)} \parallel W^{(k)} h_j^{(k-1)} \right] + \mu \tilde{A}_{ij} \right) \right)}{\sum_{k \in \mathbb{N}(i)} \exp \left(\text{LeakyReLU} \left(a^T \left[W^{(k)} h_i^{(k-1)} \parallel W^{(k)} h_k^{(k-1)} \right] + \mu \tilde{A}_{ik} \right) \right)} \quad (18)$$

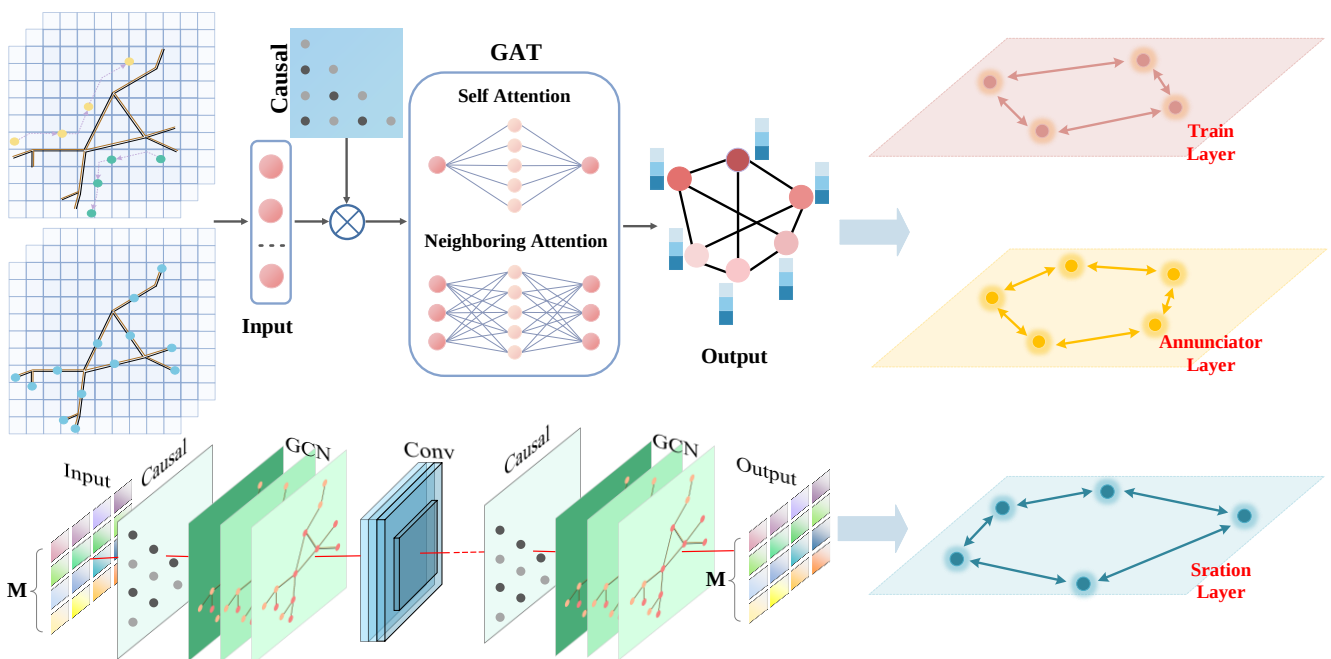


Fig.4. Railway network heterogeneity graph

where a^T denotes the learnable weight vector, $W^{(k)}$ is the feature transformation matrix, $\parallel$ represents the concatenation operation, LeakyReLU is the non-linear activation function, μ serves as a causal influence factor.

Finally, the updated representation of node v_i is obtained by aggregating the transformed features of its neighbors, weighted by the causally informed attention coefficients.

$$h_i^{(k)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W^{(k)} h_j^{(k-1)} \right) \quad (19)$$

F. Risks Level Assessment

Each node in the network may be associated with multiple risk events. To estimate the likelihood of these events, we calculate the occurrence probabilities $P(R_k | v_i)$ for a set of risk events $R(x_i) = \{R_1, R_2, \dots, R_k\}$ associated with node v_i . Specifically, nodes with causal feature $h_i^{(k)}$ are input into a multi-layer perceptron (MLP) consisting of two hidden layers. The first layer projects the input features into 128 dimensions, while the second layer compresses the representation to 64 dimensions. Both layers utilize the ReLU activation function to introduce non-linearity.

To obtain the final event probability distribution, the 64 dimensions feature vectors are passed through the Softmax layer, producing a normalized probability vector over K predefined risk event categories. To enhance generalization and reduce overfitting, Dropout regularization is applied to the hidden layers during training.

$$h_i^{(1)} = \text{ReLU}(W_1 H + b_1) \quad (20)$$

$$h_i^{(2)} = \text{Dropout}(\text{ReLU}(W_2 h_i^{(1)} + b_2)) \quad (21)$$

$$P(R_k | v_i) = \text{Softmax}(W_3 h_i^{(2)} + b_3) \quad (22)$$

where $W_1 \in R^{128 \times d}$, $W_2 \in R^{64 \times d}$, $W_3 \in R^{3 \times 64}$ denote the weight matrices, b_1 , b_2 , b_3 denote the corresponding bias terms for each layer.

To mitigate the impact of low-probability events on risk level assessment, we introduce a probability threshold θ . Only risk events with occurrence probabilities exceeding this threshold are retained for further analysis.

$$R^*(v_i) = \{R_k \in R(v_i) | P(R_k | v_i) \geq \theta\} \quad (23)$$

To comprehensively assess node risk level, we calculate the composite risk score R_i by averaging the top three highest probabilities among the retained risk events associated with node v_i . This score reflects risk level of the node based on its most significant potential events.

$$R_i = \frac{1}{k} \sum_{k=1}^3 P(R_k | v_i) \quad (24)$$

IV. EXPERIMENTS

A. Datasets

To assess the effectiveness of the proposed LiNGAM-HetGNN model for railway dispatching risk warning, we conduct experiments using a publicly available dataset from 2014 to 2024 released by the FRA. This dataset contains records of operational conditions, safety incidents, and dispatching anomalies. It integrates heterogeneous data, including train operation logs, equipment status reports, weather conditions, and maintenance records.

In this study, we construct a representative railway transport network consisting of 17 stations, 17 signals, and 6 trains, as illustrated in Figure 5. Table IV presents a sample of the raw operational data, including train IDs, station names, scheduled and actual arrival/departure times, train types, and corresponding weather conditions. The dataset encompasses the diverse range of safety-related events such as scheduling conflicts, operational delays, and equipment failures, supporting the robustness and generalizability of the proposed model under complex dispatching scenarios.

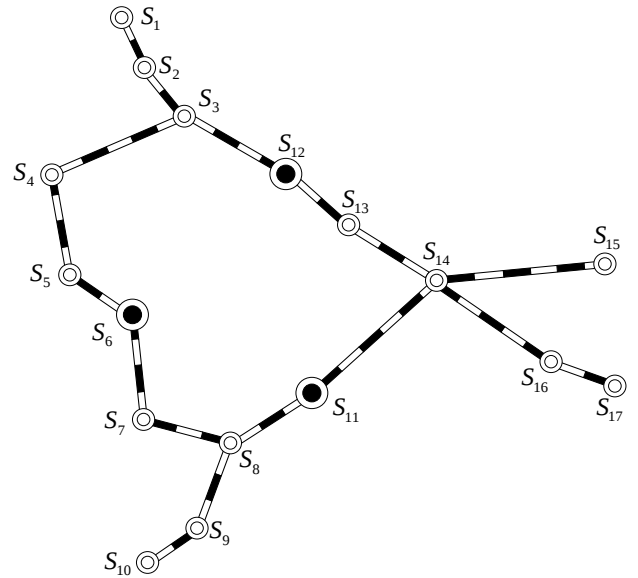


Fig.5. Railway line structure

B. Experimental Environment Settings

The experiments are conducted in a Python 3.12 environment using the PyTorch framework on a personal computer equipped with an NVIDIA RTX 3090 GPU.

The model is trained with the following hyperparameters: a learning rate of 0.001, a batch size of 64, the θ is 0.1, and the maximum training epochs of 100. To enhance generalization, the dropout rate is 0.5, utilizing the Adam optimizer, and L2 regularization is incorporated to mitigate overfitting. The loss function is set to cross-entropy.

Table IV
RAILWAY OPERATION DATA

Train number	Station	Actual arrival	Actual departure	Scheduled arrival	Scheduled departure	Train type	Weather
T_1	S_1	08:20	08:23	08:23	08:25	Passenger	Rainy
T_2	S_2	09:10	09:12	09:03	09:06	Passenger	Sunny
T_3	S_3	10:45	10:22	10:40	11:00	Freight	Cloudy
T_4	S_3	11:30	11:35	11:30	11:35	Passenger	Cloudy
T_4	S_3	11:30	11:35	11:30	11:35	Passenger	Sunny

To mitigate the risk of overfitting in causal modeling, this study adopts a sparsity control method to select the optimal regularization parameter λ . This approach determines the value of λ by targeting a desired sparsity level in the causal weight matrix, thereby promoting a parsimonious and interpretable causal graph structure. Specifically, during the LASSO regression for each risk variable, a range of candidate λ values is evaluated. For each candidate, the proportion of non-zero elements in the causal weight matrix $\hat{B}$ is calculated. The optimal λ is selected as the smallest value that ensures the proportion of non-zero elements in $\hat{B}$ meets a predefined sparsity threshold.

$$\frac{\|\hat{B}\|_0}{d(d-1)} \leq \rho \quad (25)$$

where $\|\hat{B}\|_0$ represents the number of non-zero elements, and d is the total number of risk variables. The sparsity threshold ρ is set to 0.3

C. Results Analysis

Based on LiNGAM, we construct a causal weight matrix for 40 representative risk events. To visualize the strengths of these causal influences, a lower triangular heatmap is generated, as shown in Figure 6, highlighting the irreversibility and hierarchical structure of risk relationships.

To emphasize the significant causal influence, only events pairs with weights exceeding a predefined threshold are displayed. For instance, the weight from B_7 to B_6 is 0.92, indicating insufficient construction protection has a significant causal influence on obstruction in station passageways, suggesting that B_7 serves as a source risk node, and should be prioritized for monitoring and intervention. Notably, the high-weight edges ($b_{ij} > 0.8$), shown in bright yellow, are primarily concentrated among train-related risks (Risk A), signal-related risks (Risk C), and train-signal interaction risks (Risk D), suggesting these risks form a tightly coupled risk cluster.

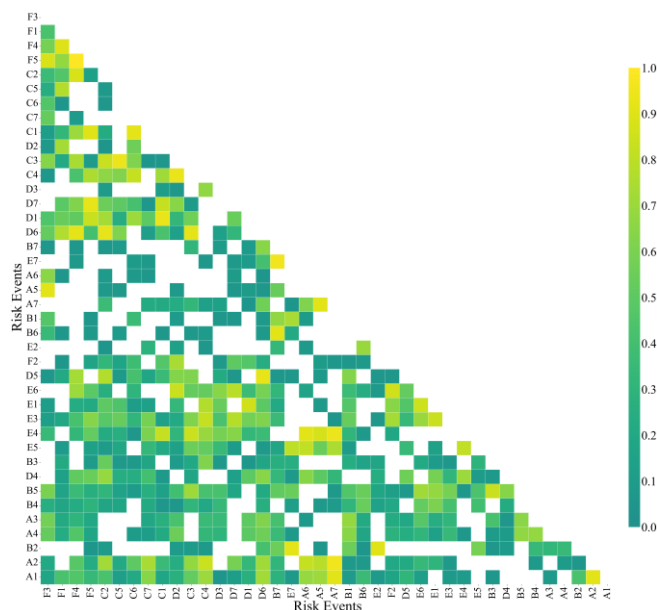


Fig.6. Causal influence of risk events

Additionally, clusters with moderate weights (0.2~0.5), such as $F_4 - C_6$ and $E_4 - E_6$, shown in light green, indicate structural causal dependencies among risk C, F and E. Blocks with consistently yellowish coloration, such as $A_6 - A_7$ and $A_1 - A_2$, with weights 0.6 to 1.0, reflect the internal propagation and linkage within risk A, implying that occurrence of one event is likely to trigger cascading failures among related risks. Moreover, top-ranked nodes such as F_3 , F_1 , and F_4 exhibit high out-degrees, indicating them as critical initiators in the causal propagation pathways, with the potential to trigger system-wide chain reactions. In contrast, nodes like B_1 , A_2 , and A_1 have higher in-degrees, situating them at the receiving end of risk propagation pathways and indicating them as vulnerable nodes under systemic stress.

Figure 7 illustrates the risk causal graph in railway dispatching, where each node represents a distinct risk event, and directed edges denote causal relationships. The color intensity of edges reflects the strength of the causal influence, while node size is proportional to out-degree, indicating potential of the node to impact other risk events. Meanwhile, to avoid result deviation caused by excessively long paths, the maximum propagation pathways for acyclic paths P_{ij} is set to 3. At the local structural level, the subgraph formed by nodes B_1 through B_7 exhibits dense connections and darker edge coloring, indicating strong coupling among station risks. This high-density cluster reveals a tightly interlinked of station risks with significant mutual influence.

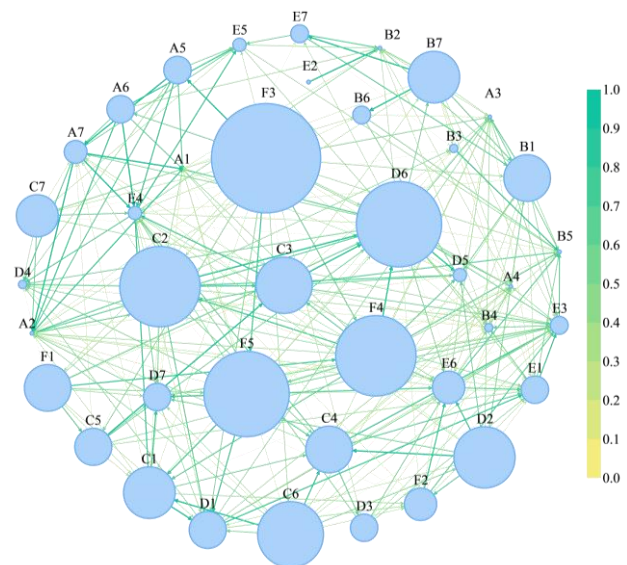


Fig.7. Risk events causal graph

From a global topological perspective, centrally located nodes in the causal network, such as C_2 , D_6 , and F_3 , exhibit larger sizes, indicating their high out-degrees and substantial causal influence. These nodes often serve as primary initiators within the dispatching system. Once disturbed, they may propagate risks through multiple pathways and various risk categories, potentially triggering cascading failures. Therefore, these central nodes should be prioritized for monitoring and intervention as critical control points.

By leveraging the topological structure of the causal graph, both backward and forward tracing of risk propagation pathways becomes possible. This enables the identification

of root causes and potential impacts of high-risk events, thereby uncovering the risk propagation pathways within the system. Furthermore, by calculating and ranking pathway strength based on causal weights, the cascading influence of globally influential risk events can be quantified. This method provides a basis for identifying vulnerable nodes and critical trigger points, supporting more effective system-level risk mitigation strategies.

D. Risks Assessment

Based on the risk scores produced by evaluation function, all nodes are classified into three risk levels: low, medium, and high, as defined in Table V. The statistical distribution of risk event probabilities across various entity nodes is presented in Table VI.

Table V
RISK LEVEL DETERMINATION CRITERIA

Risk Level	Illustrate
High	$R_i < 0.3$
Medium	$0.3 \leq R_i < 0.6$
Low	$R_i \geq 0.6$

Among the high-risk nodes, multiple events associated with nodes such as S_{13} , X_4 , X_7 , and T_1 , exhibit scores exceeding 0.6. This indicating the significant triggering potential, suggesting these nodes serve as potential risk sources. Medium-risk nodes, such as X_6 , X_9 , T_4 , and T_6 , show dominant event probabilities in the range of 0.5 to 0.6, implying a moderate risk propagation capability and suggesting them as secondary transmission hubs. In contrast, low-risk nodes like X_{11} , X_{15} , and T_3 demonstrate below 0.3,

reflecting limited influence and a relatively low risk profile.

Notably, strong central nodes such as D_1 , B_4 , and C_4 , which consistent with the propagation pathways in Figure 7. These nodes not only occupy core positions within the network topology but also exhibit strong influence in risk propagation. Therefore, they are regarded as critical risk hubs in railway dispatching system.

From the perspective of risk event probability distribution, nodes of risk B , C , and D occur frequently across multiple target nodes and are consistently associated with higher probabilities. For instance, B_1 , B_2 , B_4 , and B_5 repeatedly observed at various station nodes, with probability exceeding 0.6, indicating strong propagation capabilities. Similarly, node C_4 exhibits probabilities above 0.65 in multiple train nodes, such as T_1 , T_2 , and T_6 , underscoring its bridging role between dispatching and physical entities.

To comprehensively evaluate the effectiveness of the proposed method, we employ three standard metrics: Accuracy, Recall, and F1-score.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (26)$$

$$Recall = \frac{TP}{TP + FN} \quad (27)$$

$$Precision = \frac{TP}{TP + FP} \quad (28)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (29)$$

where TP, TN, FP, FN represent true positives, true negatives, false positives and false negatives, respectively

Table VI
PROBABILITY OF RISK EVENTS

Node	Probability	Node	Probability
S_1	$\{B_4 : 0.45, B_1 : 0.35, E_1 : 0.40\}$	X_4	$\{C_7 : 0.85, C_5 : 0.59, F_5 : 0.58\}$
S_2	$\{E_2 : 0.50, F_1 : 0.42, B_2 : 0.23\}$	X_5	$\{D_3 : 0.72, D_2 : 0.70, F_2 : 0.38\}$
S_3	$\{E_3 : 0.55, B_7 : 0.48, F_4 : 0.36\}$	X_6	$\{D_4 : 0.57, D_5 : 0.32, A_5 : 0.26\}$
S_4	$\{B_3 : 0.60, B_5 : 0.52, E_6 : 0.52\}$	X_7	$\{D_7 : 0.74, C_1 : 0.70, D_7 : 0.60\}$
S_5	$\{E_5 : 0.73, B_6 : 0.40, F_2 : 0.35\}$	X_8	$\{C_5 : 0.63, C_3 : 0.57, D_5 : 0.50\}$
S_6	$\{E_4 : 0.65, E_7 : 0.44, B_4 : 0.44\}$	X_9	$\{C_4 : 0.84, C_6 : 0.72, D_3 : 0.39\}$
S_7	$\{E_1 : 0.70, E_6 : 0.64, B_1 : 0.48\}$	X_{10}	$\{D_2 : 0.78, C_7 : 0.45, D_1 : 0.41\}$
S_8	$\{E_3 : 0.75, F_4 : 0.60, B_4 : 0.42\}$	X_{11}	$\{D_7 : 0.33, D_4 : 0.25, F_2 : 0.16\}$
S_9	$\{B_5 : 0.63, A_1 : 0.65, A_5 : 0.42\}$	X_{12}	$\{D_3 : 0.72, D_4 : 0.50, D_6 : 0.30\}$
S_{10}	$\{B_2 : 0.49, E_2 : 0.45, E_6 : 0.40\}$	X_{13}	$\{D_5 : 0.81, D_7 : 0.75, F_1 : 0.57\}$
S_{11}	$\{B_3 : 0.82, E_5 : 0.65, F_1 : 0.42\}$	X_{14}	$\{C_1 : 0.63, C_2 : 0.60, C_4 : 0.45\}$
S_{12}	$\{B_4 : 0.80, E_7 : 0.70, E_4 : 0.52\}$	X_{15}	$\{C_3 : 0.56, C_4 : 0.52, F_4 : 0.22\}$
S_{13}	$\{E_1 : 0.88, B_1 : 0.74, B_5 : 0.56\}$	X_{16}	$\{C_6 : 0.72, C_5 : 0.68, F_5 : 0.55\}$
S_{14}	$\{E_3 : 0.78, B_7 : 0.67, B_6 : 0.65\}$	X_{17}	$\{D_1 : 0.80, D_6 : 0.53, D_5 : 0.38\}$
S_{15}	$\{E_2 : 0.46, F_4 : 0.45, F_2 : 0.42\}$	T_1	$\{A_4 : 0.88, D_6 : 0.65, A_1 : 0.42\}$
S_{16}	$\{E_6 : 0.62, B_3 : 0.58, B_6 : 0.50\}$	T_2	$\{A_3 : 0.70, D_4 : 0.65, A_2 : 0.42\}$
S_{17}	$\{E_5 : 0.75, E_4 : 0.68, E_5 : 0.63\}$	T_3	$\{A_3 : 0.42, E_2 : 0.35, E_5 : 0.33\}$
X_1	$\{C_1 : 0.86, C_2 : 0.68, F_1 : 0.54\}$	T_4	$\{E_3 : 0.58, A_3 : 0.65, A_4 : 0.42\}$
X_2	$\{C_3 : 0.55, C_6 : 0.48, D_4 : 0.42\}$	T_5	$\{A_3 : 0.7, A_6 : 0.65, E_7 : 0.42\}$
X_3	$\{C_4 : 0.73, D_1 : 0.66, F_4 : 0.53\}$	T_6	$\{A_7 : 0.7, E_4 : 0.65, E_5 : 0.42\}$

To verify the rationality of the parameter α in the comprehensive weight, we test the impact of different α values on the model performance, as shown in Table VII.

Table VII
COMPARISON INDICATORS OF DIFFERENT WEIGHT

α	Accuracy	Recall	F1-score
0.3	0.92	0.93	0.93
0.4	0.98	0.99	0.99
0.5	0.94	0.90	0.92
0.6	0.89	0.87	0.87

To further verify the effectiveness of the proposed LiNGAM-HetGNN method in risk warning, we compare its performance with three traditional ML methods: Logistic Regression (LR), Support Vector Machine (SVM), and Random Forest (RF), as well as two widely adopted graph learning models: GCN and GAT. The experimental results are presented in Table VIII.

Table VIII
COMPARISON INDICATORS OF DIFFERENT METHODS

Model	Accuracy	Recall	F1-score
LR	0.64	0.90	0.87
RF	0.61	0.87	0.75
SVM	0.81	0.97	0.81
GCN	0.83	0.85	0.83
GAT	0.80	0.80	0.79
LiNGAM-HetGNN	0.98	0.99	0.99

As shown in Figure 8, LiNGAM-HetGNN achieves the

best performance across three evaluation metrics, with Accuracy of 0.98, Recall of 0.99, and F1-score of 0.99, significantly outperforming all baseline methods. Notably, compared to the second-best GCN, our method improves F1-score by 16%. Although SVM achieves a comparable Recall, its F1-score remains relatively low at 0.81, indicating a higher false positive rate. In contrast, our method achieves high recall while demonstrates superior accuracy.

The performance improvement can be attributed to two critical design components. First, the LiNGAM-based causal structure enables the model to explicitly capture the directed dependencies among risk events, thereby identifying potential propagation pathways and improving global reasoning capability of the model. Second, The HetGNN framework integrates features from heterogeneous entities types and relationships, thereby enriching node-level semantic representations and enhancing the accuracy and robustness of risk identification and assessment.

Figure 9 compares the differences in node risk scoring across various methods. ML methods such as LR in Figure 9(b) and RF in Figure 9(c) show dispersed score distributions, primarily concentrated in the range of 0.3 to 0.6. These methods fail to form distinct high-risk clusters, indicating limited ability to exploit the spatial correlations inherent in graph structures. Although SVM in Figure 9(d) demonstrates partial risk aggregation at specific nodes, the overall lack of spatial continuity indicates inadequate modeling of the topological dependencies.

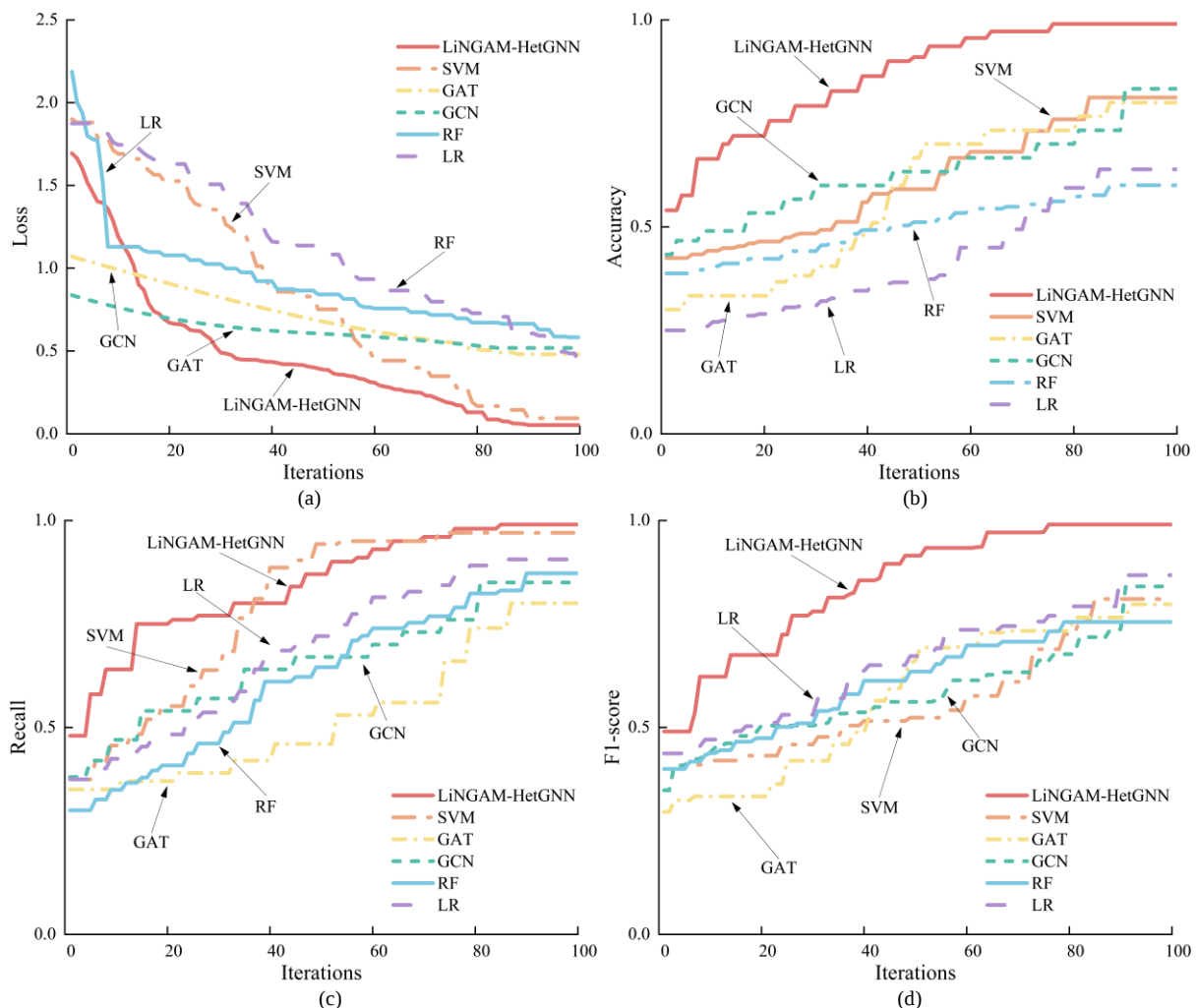


Fig.8. Iteration curves of different methods

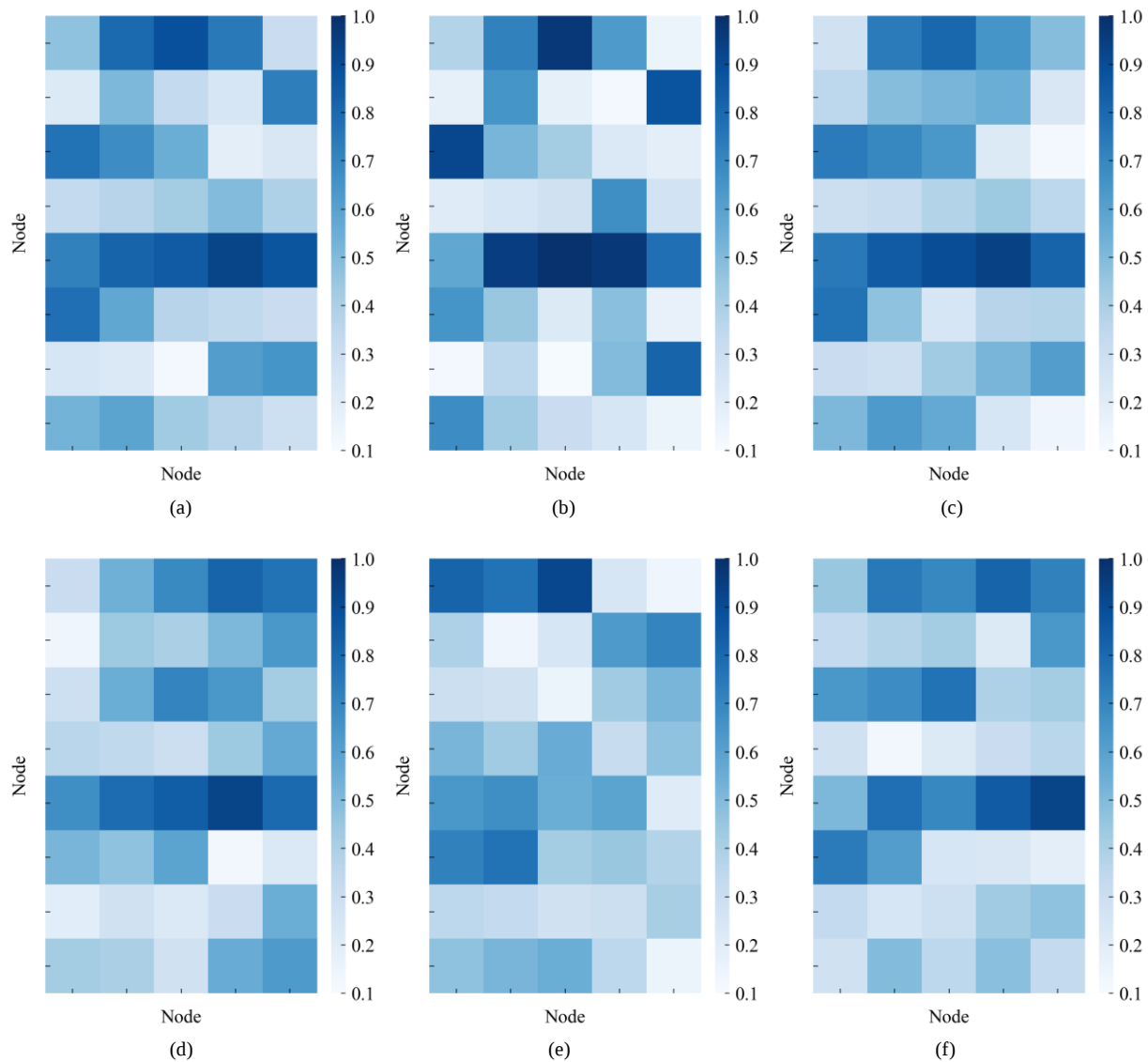


Fig.9. Risk level scores by different methods

As shown in Figure 9(a), the proposed method integrates the causal discovery with heterogeneous graph learning, significantly improving structural sensitivity and precision of node-level risk scoring. High-risk areas appear as well-connected clusters within the causal graph, accurately revealing potential risk propagation pathways in railway dispatching system. More importantly, the high-risk nodes identified by the model show strong spatial alignment with historical dispatching incidents, validating the interpretability and practical applicability of the model.

E. Ablation Studies

To validate the effectiveness of each core component in the proposed method, we conducted ablation studies to evaluate the impact of individual components on overall performance. The experimental results are summarized in Table IX.

When the causal learning component is removed, the Recall and F1-score drop to 0.94 and 0.97, respectively, representing decreases of 5.1% and 2.0% compared to the full model. This decline indicates that the absence of causal discovery impairs the ability of model to identify risk events, making it difficult to accurately capture potential causal pathways in the dispatching process. By uncovering the structural causal dependencies among variables, LiNGAM

provides more interpretable reasoning pathways, thereby improving the accuracy and credibility of risk warning.

Components	Accuracy	Recall	F1-score
No-LiNGAM	0.97	0.94	0.97
No-HetGNN	0.72	0.75	0.61
Full	0.98	0.99	0.99

The exclusion of HetGNN component leads to an even more significant impact, as shown in Figure 10. Without the ability to incorporate heterogeneous information and model the graph structure information, the model relies solely on local node features. Consequently, Accuracy and Recall drop by 26% and 24%, respectively, and the F1-score plummets to 0.61. These results underscore the essential role of HetGNN in integrating multi-source heterogeneous data and modeling risk propagation pathways across railway network. Given the inherent graph structure nature of railway dispatching system, where topological connections reflect geographic adjacency and operational routes, HetGNN enables global awareness of risk propagation by aggregating information from each node and its neighbors.

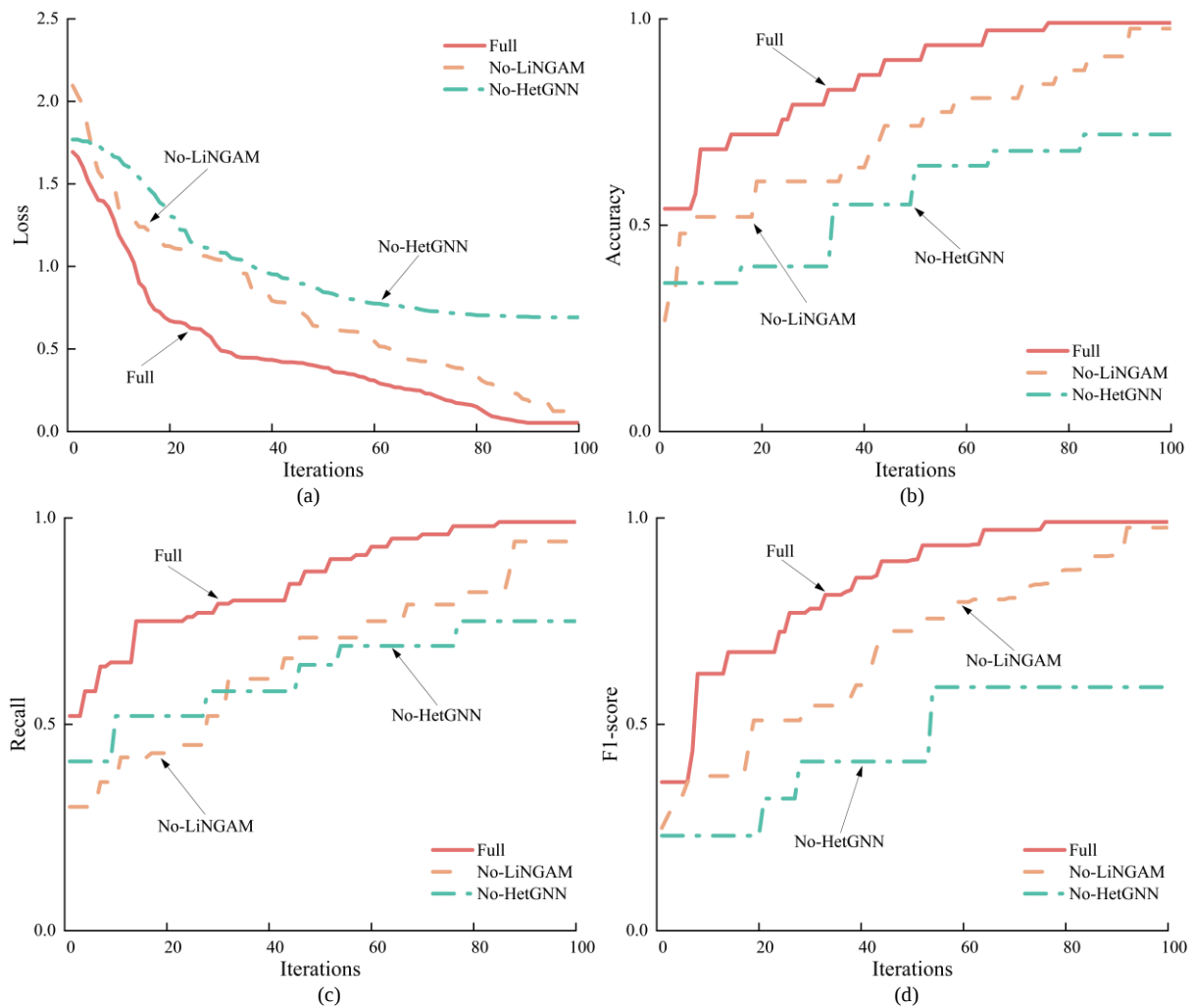


Fig.10. Iteration curve of ablation study

In conclusion, the complete LiNGAM-HetGNN model integrates causal discovery with HetGNN. The former enhances causal reasoning and interpretability of the model, while the latter strengthens the representation of graph-structured data. This synergy leads to more accurate and reliable risk warning for railway dispatching system.

The risk level scores are visualized in Figure 11. The complete model in Figure 11(a) exhibits a clear spatial clustering pattern, with risk scores highly correlated across

nodes and demonstrating strong structural connectivity. This facilitates a distinct separation between high-risk and low-risk regions. Additionally, several clusters of critical nodes form continuous risk propagation pathways, indicating strong effectiveness of the model in causal pathways reasoning and structural pattern recognition. These findings confirm the effectiveness of the model in uncovering the underlying risk propagation mechanisms within the railway transport network.

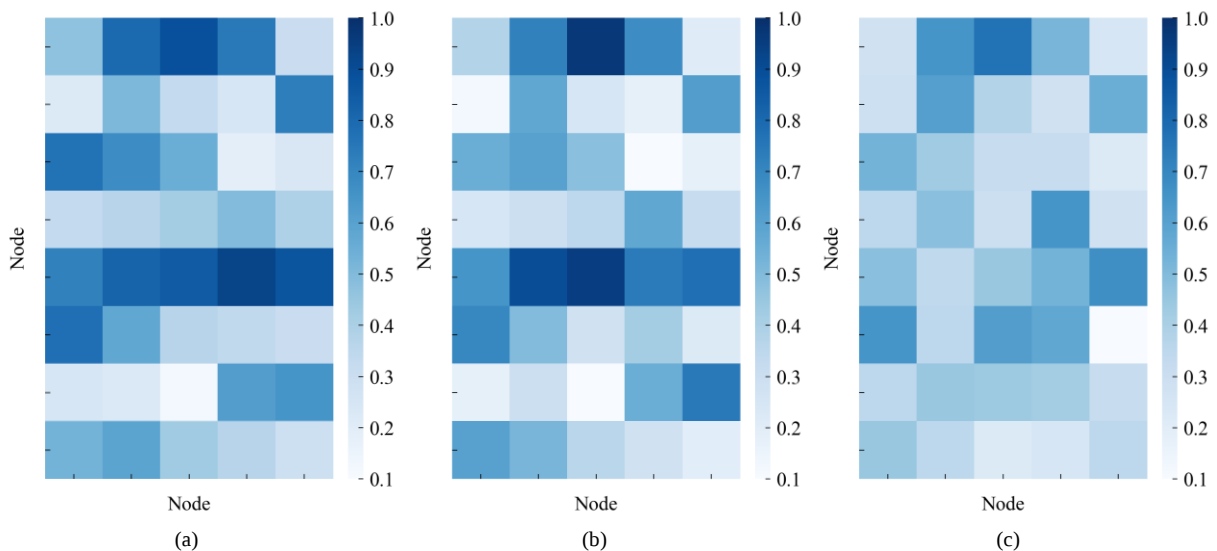


Fig.11. Risk level scores of ablation study

When the causal learning component is removed in Figure 11(b), the spatial coherence of risk scores is noticeably reduced. Connections among high-risk nodes become sparse, local connectivity weakens, and the continuity of risk propagation pathways is significantly disrupted. These results indicate that without causal structure modeling, the model struggles to accurately infer causal relationships among variables. Relying solely on HetGNN for features aggregation is insufficient to capture the deeper causal mechanisms underlying complex dispatching progress.

Similarly, when removing HetGNN component in Figure 11(c), the risk level scores map becomes highly fragmented. Inter-node risk correlations diminish significantly, resulting in scattered high-risk points and indistinct high-risk regions. Recognizable propagation pathways are nearly absent. These outcomes underscore the essential role of structural information modeling. Without leveraging the heterogeneity and topology of the railway network, risk identification performance of the model degrades substantially. In contrast, integrating HetGNN enables the model to capture global risk aggregation patterns by modeling interactions among diverse entity types and relationships, thereby improving structural sensitivity and interpretability.

V. CONCLUSION

This paper proposes a LiNGAM-HetGNN method to enhance early risk warning in railway dispatching system. By incorporating LiNGAM algorithm, the model effectively identifies asymmetric causal relationships among critical entities in incident data. It constructs a causal weight matrix of risk events, uncovering causal structures and modeling the risk propagation pathways within railway transport network. Compared with traditional ML methods, this method leverages data-driven causal discovery to uncover deep structural dependencies, thereby providing support for risk tracing and propagation analysis.

In terms of model design, this work innovatively integrates GCN and GAT to construct the HetGNN framework tailored for multi-source and multi-typed data in railway dispatching. By capturing complex interactions among diverse node types and integrating causal discovery into HetGNN, the model significantly improves the identification of risk propagation pathways. This design enhances the interpretability and accuracy of risk assessment, while also facilitates an intuitive understanding of risk propagation mechanism. Furthermore, experimental results confirm that the proposed method consistently outperforms baseline models in terms of risk recognition, causal interpretability, and generalization, as validated through risk scoring and ablation studies. These advantages contribute to improved safety and more efficient decision-making in railway operations.

Despite its demonstrated performance, this method has certain limitations. First, its effectiveness relies on the quality of input data, which may limit adaptability in complex or uncertain environments. Future research will explore the reinforcement learning to improve robustness to noise and data variability. Second, the computational cost of causal discovery and heterogeneous graph modeling limits real-time response. To address this, future work will focus on employing lightweight architectures to enhance deployment efficiency in large-scale systems.

REFERENCES

- [1] Nikhil Bugalia, Yu Maemura, and Kazumasa Ozawa, "Demand risk management of private high-speed rail operators: A review of experiences in Japan and Taiwan," *Transport Policy*, vol. 113, pp. 67-76, 2021.
- [2] Ting Yuan, Pengcheng Xiang, Huaiyin Li, and Lei Zhang, "Identification of the main risks for international rail construction projects based on the effects of cost-estimating risks," *Journal of Cleaner Production*, vol. 274, 2020.
- [3] Michelle Nolan-McSweeney, Brendan Ryan, and Sue Cobb, "Interviews with rail industry leaders about systems thinking in the management of organisational change and risk management," *Safety Science*, vol. 164, 2023.
- [4] Mingchang Zhu, Xiaoqing Zeng, Peiran Ying, and Lixia Bao, "Network-based resilience assessment of an urban rail transit infrastructure with a multi-dimensional performance metric," *Physical A: Statistical Mechanics and its Applications*, vol. 656, 2024.
- [5] Jinghua Song, Jianfeng Ding, Xuechen Gui, and Yuyi Zhu, "Assessment and solutions for vulnerability of urban rail transit network based on complex network theory: A case study of Chongqing," *Heliyon*, vol. 10, no. 5, 2024.
- [6] Adrian V. Gheorghe, Jürg Birchmeier, Dan Vamanu, Ioannis Papazoglou, and Wolfgang Kröger, "Comprehensive risk assessment for rail transportation of dangerous goods: a validated platform for decision support," *Reliability Engineering & System Safety*, vol. 88, no. 3, pp. 247-272, 2005.
- [7] Dheeraj Joshi, Wataru Takeuchi, Nirmal Kumar, and Ram Avatar, "Multi-hazard risk assessment of rail infrastructure in India under local vulnerabilities towards adaptive pathways for disaster resilient infrastructure planning," *Progress in Disaster Science*, vol. 21, 2024.
- [8] Irem Ucal Sari, Hulya Behret, and Cengiz Kahraman, "Risk governance of urban rail systems using fuzzy AHP: the case of Istanbul," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, no. 1, pp. 67-79, 2012.
- [9] Hongyan Yan, Ce Gao, Hazem Elzarka, Kareem Mostafa, and Wenbin Tang, "Risk assessment for construction of urban rail transit projects," *Safety Science*, vol. 118, pp. 583-594, 2019.
- [10] Chunyu Lin, M. Rapik Saat, and Christopher P.L. Barkan, "Semi-quantitative risk assessment of adjacent track accidents on shared-use rail corridors," *Journal of Rail Transport Planning & Management*, vol. 24, 2022.
- [11] Kyuetaek Oh, Mintaek Yoo, Nayoung Jin, Jisu Ko, Jeonguk Seo, Hyojin Joo, and Minsam Ko, "A review of deep learning applications for railway safety," *Applied Sciences*, vol. 12, no. 20, 2022.
- [12] Yingying Liao, Lei Han, Haoyu Wang, and Hougui Zhang, "Prediction models for railway track geometry degradation using machine learning methods: a review," *Sensors*, vol. 22, no. 19, 2022.
- [13] Nader Azad, Elkafi Hassini, and Manish Verma, "Disruption risk management in railroad networks: An optimization-based methodology and a case study," *Transportation Research Part B: Methodological*, vol. 85, pp. 70-88, 2016.
- [14] Boyi Su, Andrea D'Ariano, Shuai Su, Zhikai Wang, and Tao Tang, "A data-driven mixed-integer linear programming approach for real-time rescheduling of urban rail transit under rolling stock faults," *Transportation Research Part C: Emerging Technologies*, vol. 169, 2024.
- [15] Huiran Liu, Zheng Wang, and Zhiming Fang, "Integrating hybrid deep learning and path allocation for real-time inbound passenger flow prediction and anomaly detection in urban rail transit," *Information Sciences*, vol. 692, 2025.
- [16] Gang Xue, Shifeng Liu, Long Ren, and Daqing Gong, "A data aggregation-based spatiotemporal model for rail transit risk path forecasting," *Reliability Engineering & System Safety*, vol. 239, 2023.
- [17] Wei Bi, Jennifer Schooling, and Kristen MacAskill, "Assessing flood resilience of urban rail transit systems: Complex network modelling and stress testing in a case study of London," *Transportation Research Part D: Transport and Environment*, vol. 134, 2024.
- [18] Liudan Jiao, Qiudie Luo, Hao Lu, Xiaosen Huo, Yu Zhang, and Ya Wu, "Research on the urban rail transit disaster chain: Critical nodes, edge vulnerability and breaking strategy," *International Journal of Disaster Risk Reduction*, vol. 102, 2024.
- [19] Dong San Kim, Dong Hyun Baek, and Wan Chul Yoon, "Development and evaluation of a computer-aided system for analyzing human error in railway operations," *Reliability Engineering & System Safety*, vol. 95, no. 2, pp. 87-98, 2010.
- [20] Habib Hadj-Mabrouk, "Analysis and prediction of railway accident risks using machine learning," *AIMS Electronics and Electrical Engineering*, vol. 4, no. 1, pp. 19-46, 2020.
- [21] Gan Shi, Xiaobing Ding, Chen Hong, Zhigang Liu, and Lu Zhao,

- “Research on key risk chain mining method for urban rail transit operations: A new approach to risk management,” *International Journal of Transportation Science and Technology*, vol. 13, pp. 29-43, 2024.
- [22] Samira Belhour, Rachid Chaib, Rachid Kahoul, and Ion Verzea, “For an effective disruptions management in rail transport: Case study tramway of Constantine, Algeria,” *Procedia Structural Integrity*, vol. 48, pp. 288-295, 2023.
- [23] Ping Wu, Qian Chen, Yinqi Chen, Shuai Chen, and Jie Zou, “ISM-MICMAC based safety risk sources analysis and control measures for underground engineering of urban rail transit projects,” *Journal of Engineering Research*, vol. 11, no. 3, pp. 40-50, 2023.
- [24] Yuan Cao, Yue Liu, Yongkui Sun, Shuai Su, and Feng Wang, “Enhancing rail safety through real-time defect detection: A novel lightweight network approach,” *Accident Analysis & Prevention*, vol. 203, 2024.
- [25] Yanhui Wang, Kexin Sheng, Penghua Niu, Chenhong Chu, Man Li, and Limin Jia, “A comprehensive analysis method of urban rail transit operation accidents and safety management strategies based on text big data,” *Safety Science*, vol. 172, 2024.
- [26] <https://railroads.dot.gov/safety-data/forms-guides-publications/forms/archived-forms>