Mitigating Noise and Imbalance in Social Governance Graphs for Multi-Type Risk Assessment

Di Jin[1,2], Haotian Zhao[1], Xiaobao Wang[1*], Fengyu Yan[1], Dongxiao He[1]

[1]Tianjin Key Laboratory of Cognitive Computing and Application, College of Intelligence and Computing, Tianjin University, Tianjin, China

[2]Key Laboratory of Artificial Intelligence Application Technology, Qinghai Minzu University, Xining, China

{jindi, zhaohaotian, wangxiaobao, fengyuyan, hedongxiao}@tju.edu.cn

Abstract

Heterogeneous graphs are widely used to model real-world systems with diverse entity types and relational structures, and existing methods have shown promising performance in various applications. However, most current models assume balanced and semantically aligned features across nodes, which rarely holds in practice. In scenarios such as social risk governance, node types often exhibit severe feature imbalance, making it difficult for standard aggregation mechanisms to extract meaningful signals. This imbalance leads to three key challenges: inaccurate neighbor weighting, noise propagation, and biased representations skewed toward text-rich nodes. To address these issues, we propose HeCoGNN, a collaborative and adaptive aggregation framework that jointly performs neighbor filtering and relation-aware message calibration, enabling robust representation learning under semantic disparity. Experiments on real-world social governance graphs show that HeCoGNN consistently outperforms state-of-the-art baselines, particularly in handling underrepresented and noisy node types.

Introduction

With the increasing complexity of real-world data, many applications such as knowledge graphs (Wang et al. 2025c, 2023b), recommender systems (Jin et al. 2023), financial risk control (Yang et al. 2020), gtaph attacks (Zhu et al. 2024b; He et al. 2025; Zhu et al. 2024a; Wang et al. 2025b; Yan et al. 2025) and social networks (Hui et al. 2025; Lu, Chen, and Li 2024; Hu et al. 2023) now involve diverse types of entities and complex relationships. Traditional graphs (Wang et al. 2024a) are often insufficient for modeling such multifaceted structures, making heterogeneous graphs essential for representing multi-type nodes and relations. These graphs provide a unified framework for encoding rich semantics and intricate associations, and have become a fundamental tool for modeling information flow across diverse domains.

Heterogeneous graph learning methods are generally categorized into metapath-based and metapath-free approaches. Metapath-based models rely on predefined semantic paths to capture meaningful structural and semantic dependencies within the graph (Wang et al. 2019; Dong, Chawla, and Swami 2017; Fu et al. 2020), and were widely adopted in early research. In contrast, metapath-free approaches eliminate the need for manual path engineering by learning node representations through direct modeling of neighbors and relations, leading to a new generation of adaptive heterogeneous graph neural networks (Li et al. 2025; Lv et al. 2021; Zhu et al. 2019).

Despite their effectiveness, these methods face serious challenges when applied to real-world heterogeneous graphs, which often suffer from imbalanced node feature distributions and the presence of noisy or weakly informative edges. These issues are particularly evident in social risk governance scenarios, where the graph typically includes various entity types such as persons, locations, and events, along with diverse relations including participation and occurrence. In such graphs, event nodes tend to be rich in high-dimensional textual or domain-specific information, whereas person and location nodes often contain only limited structured attributes due to privacy constraints or data collection limitations. This creates a strong semantic disparity between nodes, which hinders the ability of conventional graph learning models to generate effective representations, especially for underrepresented node types.

To quantify this issue, we conduct node classification experiments on real-world social risk case graphs, evaluating three node types (person, event, location) using three representative models: MLP, HGT (Hu et al. 2020), and RGCN (Schlichtkrull et al. 2018). As shown in the left of Figure 1, event nodes, which are text-rich, achieve significantly higher classification accuracy (71.17%) compared to person (54.63%) and location (40.50%) nodes under an MLP baseline. Incorporating graph structure through HGT and RGCN does not improve performance for event nodes and provides only marginal gains for text-less nodes. These results indicate that existing heterogeneous graph models struggle to generalize under severe feature imbalance.

As shown in the right of Figure 1, we argue that feature imbalance fundamentally disrupts the neighbor aggregation mechanism in heterogeneous graphs. This disruption degrades node representations and leads to lower classification performance. Effective representation learning depends on accurately capturing semantically relevant signals from neighboring nodes. However, when semantically rich and sparse nodes coexist, conventional aggregation strategies

*Corresponding author

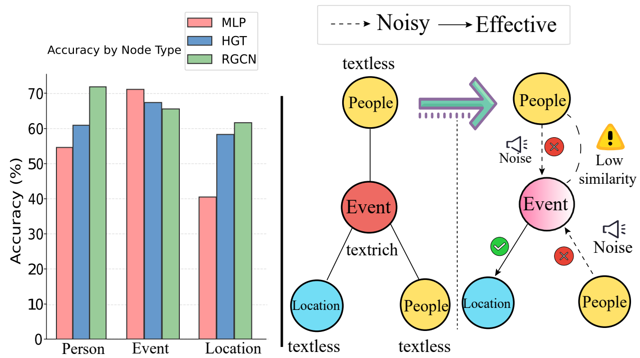

Figure 1: Feature imbalance and its impact on aggregation. Left: Node classification accuracy for Person, Event, and Location nodes using MLP, HGT, and RGCN. Right: Schematic of message passing and node interactions in a heterogeneous graph under feature imbalance.

such as attention mechanisms, mean pooling, or relation-specific weighting often fail to correctly distinguish between informative and uninformative neighbors. This results in a cascade of representational failures. ***Informative neighbors with low surface-level similarity may be underweighted or even ignored, while noisy or weakly informative nodes may be mistakenly emphasized***. These aggregation mismatches destabilize the message-passing process and are particularly detrimental for nodes that lack sufficient features and rely on neighbors for semantic enrichment. Collectively, these issues reflect three interconnected challenges: incorrect neighbor weighting, noise propagation, and representation bias toward semantically rich nodes. Together, they significantly impair the generalization ability of existing heterogeneous graph learning methods.

To address these issues, we draw inspiration from the decentralized and modular coordination principles in Federated Learning (Zou et al. 2024; Fu et al. 2025; Wang et al. 2025a), which has proven effective in handling non-IID data across distributed clients. In federated settings, clients often possess data with distinct feature distributions and must perform local, context-aware updates to avoid optimization bias. This setting is analogous to the heterogeneous graph scenario, where different node types exhibit varying levels of semantic richness and structural connectivity. Similar to clients in Federated Learning, nodes in heterogeneous graphs require adaptive, localized strategies rather than uniform aggregation mechanisms to achieve balanced and robust representation learning. To this end, we propose Heterogeneous Collaborative Graph Neural Network (HeCoGNN), a collaborative and adaptive edge attention framework that integrates dynamic neighbor selection and relation-aware message aggregation in a unified manner. Specifically, HeCoGNN employs a type- and context-sensitive gating mechanism to filter out noisy or weakly correlated neighbors, effectively mitigating the risk of noise propagation and incorrect neighbor weighting. On the refined topology, a relation-specific attention module further calibrates aggregation weights to amplify semantically relevant yet underrepresented signals, addressing representation bias induced by feature disparity. By jointly optimizing edge selection and message weighting, HeCoGNN dynamically restructures the propagation process based on local semantics and structural context, enabling robust and balanced representation learning across heterogeneous node types.

Our contributions are summarized as follows:

- We analyze the fundamental reasons why conventional heterogeneous graph neural networks fail in imbalanced and noisy scenarios.
- We propose HeCoGNN, a collaborative and adaptive aggregation framework that mitigates semantic mismatch, noise propagation, and representation bias.
- Extensive experiments on real-world social risk governance graphs demonstrate the effectiveness of our method, showing significant improvements over state-of-the-art baselines.

Related Work

Heterogeneous graph neural networks (HGNNs) are essential for modeling the complex relationships in real-world multi-type data. Early works such as HAN (Wang et al. 2019) introduced hierarchical attention to aggregate information from different node and meta-path types, enabling flexible fusion of heterogeneous semantics. HGT (Hu et al. 2020) further advanced this direction with dynamic attention (Vaswani et al. 2017) allocation for fine-grained modeling of node and relation types.Beyond attention, various aggregation strategies have been proposed to enhance semantic expressiveness. RGCN (Schlichtkrull et al. 2018) implemented edge-specific transformations to handle the varying importance of neighbors. HetGNN (Zhang et al. 2019) and MAGNN (Fu et al. 2020) leverage meta-paths and structure-aware mechanisms to capture richer semantics and higher-order connectivity, while HetSANN (Hong et al. 2020) employs multi-level attention for flexible message aggregation in complex topologies.Recent studies explore alternative architectures and theoretical approaches. SeHGNN (Yang et al. 2023) abandons neighbor attention in favor of single-layer long-metapath fusion for efficiency and accuracy, and PSHGCN (He et al. 2024) introduces spectral-based HGNNs with positive semidefinite heterogeneous graph filters, achieving strong performance and scalability. HGRAN (Wang et al. 2024b) proposes a novel heterogeneous graph representation learning model using a hybrid-attention mechanism, combining relation attention and node attention to effectively capture complex heterogeneous information and structural relationships, outperforming state-of-the-art methods on multiple benchmark datasets, while HetTree (Guan et al. 2025) employs subtree attention on semantic trees to capture hierarchical relationships often missed by previous models. Despite these advances, most HGNNs still struggle with feature imbalance, noisy edges, and representation bias, especially when node types differ greatly in attribute richness. This motivates the development of more adaptive and robust frameworks, as proposed in this work, to better address the challenges of real-world heterogeneous data.

Preliminaries

A heterogeneous graph is defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the sets of nodes and edges, respectively. $\mathcal{A}$ and $\mathcal{R}$ represent the sets of node types and edge types, with mapping functions $\phi : \mathcal{V} \rightarrow \mathcal{A}$ and $\psi : \mathcal{E} \rightarrow \mathcal{R}$. Each node $v \in \mathcal{V}$ is associated with a feature vector $\mathbf{x}_v \in \mathbb{R}^{d_{\text{in}}}$ representing its attributes. In heterogeneous graphs, the rich semantics of different types of entities and relations introduce complexity in information propagation, requiring sophisticated models to effectively capture both local and global dependencies across types.

We consider a multi-type node classification task on a heterogeneous graph. Let $\mathcal{T}_V \subseteq \mathcal{A}$ denote the set of node types with classification targets. For each type $t \in \mathcal{T}_V$ and each node $v \in \mathcal{V}$ such that $\phi(v) = t$, the objective is to predict a label $y_v \in \mathcal{Y}_t$, where $\mathcal{Y}_t$ is the type-specific label space with cardinality $C_t = |\mathcal{Y}_t|$. The goal is to learn unified node representations that jointly model node attributes and the heterogeneous structural context defined by the edge types in $\mathcal{R}$. Node types outside $\mathcal{T}_V$ serve only as auxiliary information and are not assigned classification targets.

Method

Overall Framework

The proposed HeCoGNN framework addresses three core challenges in heterogeneous graph learning: incorrect neighbor weighting, noise propagation, and representation bias toward semantically rich nodes. As shown in Figure 2, the architecture consists of two collaborative modules: **GateNet** and **ContextNet**. **GateNet** adaptively selects which edges participate in message passing based on node and neighbor types. This dynamic gating suppresses noisy or weak connections, retaining essential information flow. **ContextNet** assigns relation-aware attention weights to the filtered edges, enabling feature-sparse nodes to effectively absorb contextual information from rich neighbors and mitigating representation bias. By combining discrete edge selection (GateNet) with fine-grained attention (ContextNet), HeCoGNN flexibly controls information propagation, resulting in more balanced and robust node representations for real-world heterogeneous graphs.

Feature Projection and Initialization

To enable joint message passing across heterogeneous node types, we first project the raw features of each type into a common hidden space. Specifically, for each node type $a \in \mathcal{A}$, we learn a type-specific linear transformation:

$$\mathbf{H}^{(a)} = \mathrm{Proj}^{(a)}(\mathbf{X}^{(a)}) = \mathbf{X}^{(a)}\mathbf{W}^{(a)} + \mathbf{b}^{(a)}, \tag{1}$$

$\mathbf{X}^{(a)} \in \mathbb{R}^{N_a \times d_{\text{in}}}$ denotes the input feature matrix for node type a, with N_a nodes and d_{in} feature dimension.$\mathbf{W}^{(a)} \in \mathbb{R}^{d_{\text{in}} \times d_h}$ and $\mathbf{b}^{(a)} \in \mathbb{R}^{d_h}$ are learnable parameters, representing the projection weights and bias for node type a. $\mathbf{H}^{(a)} \in \mathbb{R}^{N_a \times d_h}$ is the resulting projected feature matrix in the unified hidden space of dimension d_h.

The purpose of this step is to align all node types into a shared semantic space, so that nodes of different types can be aggregated and compared in subsequent layers. This also mitigates the feature dimension disparity between textrich and textless nodes, laying the foundation for effective heterogeneous message passing.

GateNet

Following CoGNN (Zou et al. 2024) GateNet is designed to dynamically regulate the structural flow of information in heterogeneous graphs by learning edge-wise participation probabilities conditioned on both source and target node types. This bidirectional gating mechanism ensures that only mutually consented, semantically coherent connections are preserved for aggregation, effectively reducing structural noise and suppressing irrelevant interactions. Unlike simple aggregation schemes, GateNet adaptively filters out spurious or redundant neighbors based on node-specific features and type semantics, thereby addressing the core limitations of uniform message passing and mitigating the impact of noise propagation in heterogeneous settings.

Node-Type-Specific MLP Gating. For each node v of type $a = \phi(v)$, GateNet uses two independent multilayer perceptrons (MLPs) to generate logits for "in" and "out" edge gates:

$$\mathbf{z}_{in}^{(a)}(v) = \mathrm{MLP}_{in}^{(a)}(\mathbf{h}_v), \quad \mathbf{z}_{out}^{(a)}(v) = \mathrm{MLP}_{out}^{(a)}(\mathbf{h}_v), \tag{2}$$

where $\mathbf{h}_v \in \mathbb{R}^{d_h}$ is the hidden representation of node v after feature projection. Each $\mathrm{MLP}_{in}^{(a)}$ and $\mathrm{MLP}_{out}^{(a)}$ outputs a 2-dimensional logit vector, encoding the selection probability for each gate (*select*, *not select*).

Gumbel-Softmax Sampling. Given the MLP-produced logits $\mathbf{z}_{in}$ and $\mathbf{z}_{out}$ for each node, we employ the Gumbel-Softmax trick to obtain discrete but differentiable gating decisions for edge activation.

Specifically, the logits are perturbed with Gumbel noise and converted to probabilities as:

$$g_i = -\log(-\log(U_i)), \quad U_i \sim \mathrm{Uniform}(0,1) \tag{3}$$

$$p_i = \frac{\exp\left((z_i + g_i)/\tau\right)}{\sum_j \exp\left((z_j + g_j)/\tau\right)}, \tag{4}$$

where τ is a temperature hyperparameter controlling the "hardness" of the sample.

The resulting probabilities $\mathbf{p}_{in}, \mathbf{p}_{out} \in [0,1]^2$ provide a continuous and differentiable approximation to one-hot selection (as $\tau \rightarrow 0$, this approaches a hard $\arg\max$). During training, we use the straight-through estimator: the forward pass uses hard samples, and the backward pass uses the soft version for gradient computation.

Direct use of $\arg\max$ is non-differentiable and breaks gradient flow. Gumbel-Softmax (Potapczynski, Loaiza-Ganem, and Cunningham 2020) enables end-to-end optimization of the MLP-based gating network, translating learned scores into sparse, data-driven edge activations, and offering a principled, flexible alternative to fixed thresholds or static edge selection. This is crucial for precise structural control and stable training in the presence of severe feature imbalance and structural noise.

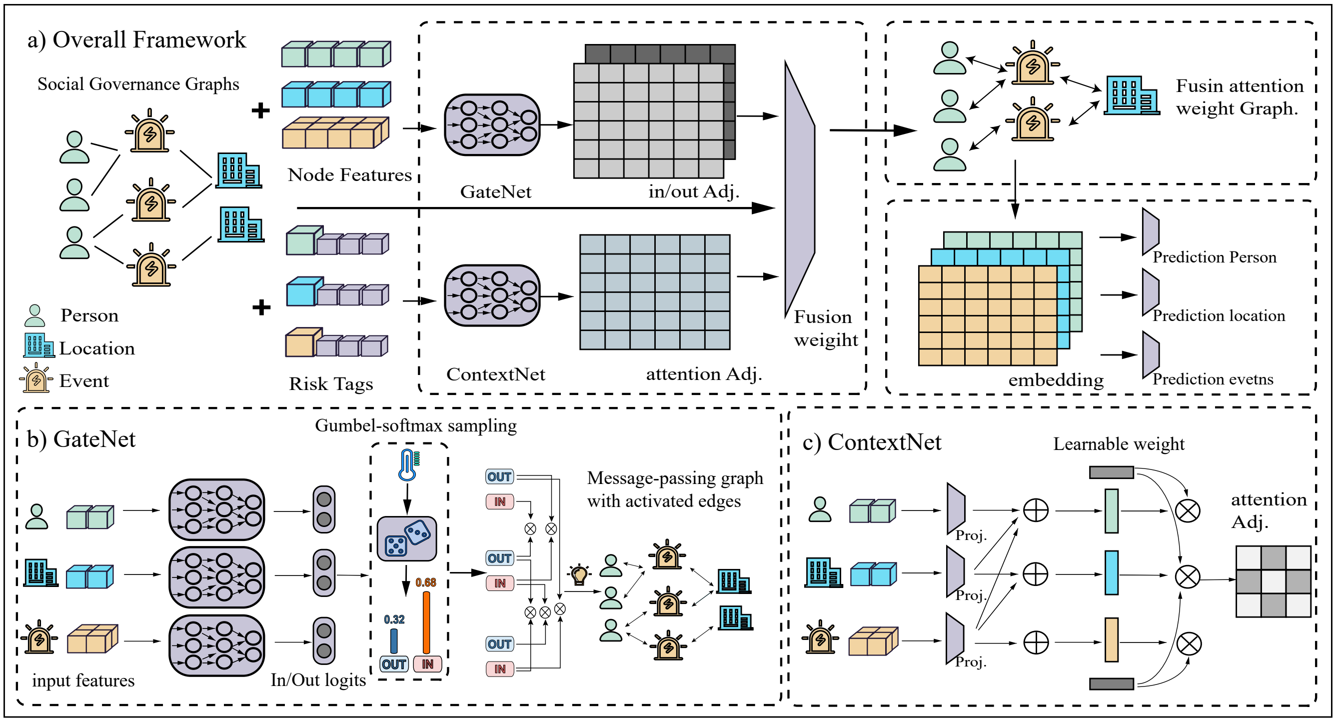

Figure 2: The overall architecture of HeCoGNN: (a) Overall pipeline; (b) GateNet for in/out probability generation; (c) ContextNet for attention adjacency computation.

Dynamic Edge Gate Construction. After obtaining the in- and out- selection probabilities for each node from the GateNet module, we construct the dynamic gating score for every edge in the graph. Specifically, for each edge (u, r, v), u is the source node, v is the target node, and r is the relation type, we combine the out-gating probability of the source node u and the in-gating probability of the target node v as follows:

$$s_{(u,r,v)} = [\mathbf{p}_{out}^{(\phi(u))}(u)]_1 \cdot [\mathbf{p}_{in}^{(\phi(v))}(v)]_1, \qquad (5)$$

where $[\cdot]_1$ denotes the probability corresponding to the "select" gate in the Gumbel-Softmax output, and $\phi(\cdot)$ is the node type function.

This design embodies the principle of mutual agreement: an edge is strongly activated only if both the source node is willing to send information (high out-probability) and the target node is willing to receive it (high in-probability). If either node "refuses" to participate, the gating score will be low, effectively reducing or blocking message propagation along that edge.

By constructing the edge gating mechanism in this multiplicative and mutually dependent manner, we ensure that each edge activation is not only node-specific but also deeply context-aware, allowing it to dynamically adapt to the evolving representational state and semantic landscape of every node in the heterogeneous graph. This mechanism enables the model to effectively suppress irrelevant, spurious, or noisy connections, while simultaneously preserving and amplifying information flow along the most critical, semantically aligned, and high-value paths. As a result, it directly addresses the dual challenges of rigid aggregation and uncontrolled noise propagation, thereby laying a solid foundation for more precise, robust, and interpretable message passing in the downstream relation-aware aggregation stage handled by the ContextNet module.

ContextNet

Motivated by the semantic sparsity of nodes such as Person and Location, ContextNet is designed to amplify weak but informative signals through relation-specific, multi-head attention, thereby facilitating semantic completion for textless nodes. Building upon the dynamic edge pruning performed by GateNet, ContextNet assigns fine-grained, context-aware attention weights to the remaining neighbors, allowing flexible and relation-sensitive information aggregation. This mechanism plays a critical role in mitigating representation bias and enhancing the expressive power of structurally or semantically underrepresented node types.

Enhancing Representation of Textless Nodes. Textless nodes (e.g., person, location) typically lack rich semantic features, making them vulnerable to being overwhelmed during message aggregation. ContextNet addresses this by computing context-aware, heterogeneous attention: for each target node, attention weights (Vaswani et al. 2017) over its neighbors are conditioned not only on neighbor features but also on the relation type and the target node's context. This enables textless nodes to selectively amplify signals from informative textrich neighbors, effectively achieving semantic completion and improving downstream classification.

Multi-Head Relation-Specific Attention. ContextNet utilizes multi-head, relation-specific attention to model the complex semantic patterns in heterogeneous graphs. For each edge (u, r, v) and each head h:

$$\alpha_{(u,r,v)}^{(h)} = \mathrm{softmax}\left(\mathbf{a}_r^{(h)\top} \left[\mathbf{h}_u \parallel \mathbf{h}_v\right]\right), \quad (6)$$

where $\mathbf{a}_r^{(h)}$ is a trainable vector for relation r and head h. This design allows the model to attend to diverse relational patterns and semantic perspectives, ensuring adaptive aggregation tailored to each edge type.

Attention Computation and Integration. For each edge, the raw attention score is:

$$\alpha_{(u,r,v)} = \mathbf{w}_r^\top [\mathbf{h}_u \| \mathbf{h}_v] + b_r, \quad (7)$$

and the normalized attention is:

$$a_{(u,r,v)} = \frac{\exp(\alpha_{(u,r,v)})}{\sum_{u' \in \mathcal{N}_r(v)} \exp(\alpha_{(u',r,v)})}. \quad (8)$$

The final edge weight is obtained by fusing attention with GateNet's gate:

$$\tilde{a}_{(u,r,v)} = a_{(u,r,v)} \cdot s_{(u,r,v)}. \quad (9)$$

This captures both "who can speak" (GateNet) and "who should be listened to" (ContextNet), achieving robust, balanced, and interpretable message passing in highly heterogeneous and imbalanced graphs.

Collaborative Evolutionary Message Passing

A central innovation of HeCoGNN is the collaborative and evolutionary integration of GateNet and ContextNet within each message passing layer of the architecture. Rather than performing neighbor selection and attention aggregation independently, our model enables these two modules to interact and co-evolve seamlessly throughout the learning process, dynamically shaping both the structure and intensity of information flow in the heterogeneous graph.

At each layer, GateNet first computes dynamic, type- and context-aware gating probabilities for all edges, sparsely selecting which neighbors are eligible for message passing. ContextNet then assigns fine-grained, relation-specific attention weights to these activated edges, further modulating the strength of information aggregation based on the semantic context of the connections. The final message passing weights are obtained by fusing GateNet's gates with ContextNet's attention scores via element-wise multiplication:

$$\tilde{a}(u, r, v) = a(u, r, v) \cdot s_{(u,r,v)}. \quad (10)$$

As node representations are iteratively updated, both gating and attention strategies are jointly adapted in a layer-wise, co-evolutionary manner, enabling HeCoGNN to dynamically restructure information flow and achieve robust, balanced representation learning.

Through this collaborative evolutionary message passing process, HeCoGNN can dynamically and flexibly restructure the effective graph at each layer, thereby enabling richer, more discriminative, and more balanced node representations for downstream tasks.

Optimization

HeCoGNN is optimized for multi-type node classification over the set $\mathcal{T}_V$. For each type $t \in \mathcal{T}_V$, a dedicated classifier head $f^{(t)} : \mathbb{R}^{d_h} \to \mathbb{R}^{C_t}$ outputs logits $\mathbf{o}_v$ for nodes of type t. During training, we compute the cross-entropy loss for each type separately using the labeled nodes S_t, and then average the losses across all types:

$$\mathcal{L}_{\text{cls}} = \frac{1}{|\mathcal{T}_V|} \sum_{t \in \mathcal{T}_V} \frac{1}{|S_t|} \sum_{v \in S_t} \mathrm{CrossEntropy}(\mathbf{o}_v, Y_v), \quad (11)$$

where Y_v is the ground-truth label for node v. This approach ensures that each node type is supervised according to its own label space, and prevents imbalance among types. All model parameters—including feature projection, GateNet, ContextNet, and classifier heads—are updated jointly using the Adam optimizer (Kingma and Ba 2015).

Experiment

Datasets

Overview To evaluate HeCoGNN in realistic settings, we construct two heterogeneous graph datasets from real-world police case records, referred to as Social Governance I (SG-I) and Social Governance II (SG-II). These datasets are generated by parsing and extracting structured and unstructured information from approximately 16,000 and 1,000 incident reports, respectively. Each dataset includes four types of nodes: (1) Person, denoting individuals involved in cases, characterized by identity attributes and brief textual profiles; (2) Event, denoting case incidents enriched with structured keywords and detailed narratives; (3) Location, referring to geographical entities extracted from incident records; (4) EventType, representing abstract event categories that support semantic generalization during aggregation. The statistics of both datasets are summarized in Table 1.

Data Filtering To ensure dataset quality and representativeness, we implement a rigorous data filtering pipeline. The raw police records contain noisy, redundant, and incomplete entries. We first remove records missing key attributes (identity, location, or event descriptions) or with invalid risk labels. For person, event, and location nodes, only instances reliably matched to at least one incident are retained, ensuring every node is involved in the graph. Textual attributes undergo normalization, keyword extrgate, and deduplication; short or uninformative descriptions are excluded. For event nodes, we discard samples with ambiguous or overly generic keywords. Only records with sufficient semantic and structural information are preserved. After filtering, we obtain two benchmark heterogeneous graph datasets: a small set (1,000 high-quality samples) and a large set (15,000 samples). The details of node feature construction and label annotation are described below.

Node Feature Construction For each node, descriptive information is processed using regular expressions, segmentation, and keyword extrgate. The resulting keywords are concatenated and encoded into a unified 384-dimensional semantic vector using a pre-trained SentenceTransformer

Dataset	SG-I	SG-II
Node Types		
Person	12,024	1,051
Event	15,277	1,092
Location	2,931	301
EventType	163	46
Edge Types		
(Person, involved, Event)	15,277	1,092
(Location, occurred, Event)	14,508	1,092
(Event, belong_to, EventType)	15,277	1,092
(Event, rev_involved, Person)	15,277	1,092
(Event, rev_occurred, Location)	14,508	1,092
(EventType, rev_belong_to, Event)	15,277	1,092

Table 1: The statistics of both datasets

model, ensuring all node types are represented in a common embedding space suitable for message passing.

Label Annotation Risk levels (*low*, *medium*, *high*) are initially assigned to each node in the raw dataset based on its contextual semantics, and subsequently mapped to integer labels, which serve as supervised targets for the node classification tasks during training.

Experimental Setups

Baselines We compare our proposed HeCoGNN model with several state-of-the-art graph learning methods, including both heterogeneous and homogeneous approaches. The heterogeneous methods include HAN (Wang et al. 2019), HGT (Hu et al. 2020), RGCN (Schlichtkrull et al. 2018) and HeteroSAGE (Hamilton, Ying, and Leskovec 2017), HGATE (Wang et al. 2023a), HGRAN (Wang et al. 2024b), while the homogeneous methods include GCN (Kipf and Welling 2017), GAT (Velickovic et al. 2018) and a multilayer perceptron (MLP) using only node features. All models are implemented with the same node features and edge structures for a fair comparison, and hyperparameters are tuned based on standard practices. Additionally, all models use the same input features and label splits as HeCoGNN to ensure consistency in the evaluation.

Experimental Settings We address a multi-type node classification problem on heterogeneous graphs, where each graph contains Person, Event, and Location nodes. The objective is to train a unified model that predicts the risk level (low, medium, or high) for each node type simultaneously. Our approach fully leverages both the node attributes and the graph structure to optimize classification accuracy across all target types.For evaluation, nodes of each type are randomly split into 80% for training and 20% for testing, and results are averaged over five random splits to ensure statistical robustness. All models are implemented using PyTorch Geometric. We use a hidden dimension of 256, with 3 network layers and a dropout rate of 0.2. The Adam optimizer is used with a learning rate of 1×10^{-3}, and all models are trained for 300 epochs. The Gumbel-Softmax temperature parameter is set to 1.0 throughout all experiments.

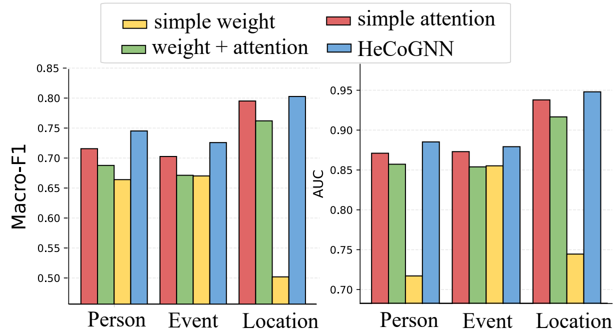

Figure 3: Parameter study: Macro-F1 and AUC on different node types for various GateNet and ContextNet.

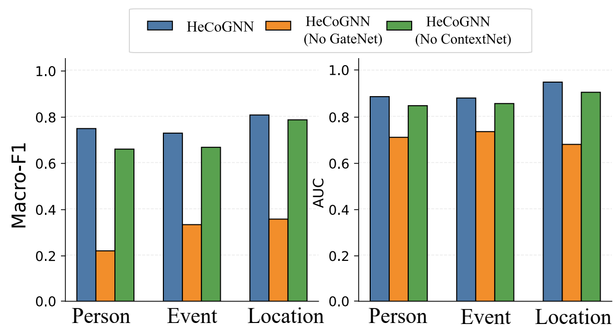

Figure 4: Ablation study: accuracy of each model variant on Person, Event, and Location nodes.

Experimental Results

Node Classification We evaluate the performance of HeCoGNN and all baselines on the joint multi-type node classification task, using Macro-F1, AUC, and Accuracy as metrics. The detailed results are presented in Table 2. Our main findings are as follows: HeCoGNN consistently achieves the best performance across all node types. Benefiting from collaborative dynamic gating and relation-aware attention, it yields substantial improvements, especially for textless node types (Person, Location), compared to all baselines. Heterogeneous GNNs (HAN, HGT, RGCN) outperform homogeneous models (GCN, GAT), highlighting the importance of modeling node and edge type heterogeneity in social risk graphs. MLP performs significantly worse than all GNN-based methods, underscoring the value of incorporating structural information. Feature imbalance across node types is evident. Models with static or non-adaptive aggregation struggle with sparse-feature nodes, whereas HeCoGNN's adaptive mechanism enhances performance on these challenging types. These results confirm HeCoGNN's effectiveness in addressing feature imbalance and noise in heterogeneous graphs. Furthermore, HeCoGNN achieves competitive Macro-F1 and Micro-F1 scores on the public ACM dataset, as shown in Table 3, demonstrating strong generalization across different graph benchmarks.

Parameter Study To investigate the individual contributions of GateNet and ContextNet, we conduct a detailed

Method	Person			Event			Location		
	Macro-F1	AUC	ACC	Macro-F1	AUC	ACC	Macro-F1	AUC	ACC
Social Governance I									
HAN	0.5037	0.7497	0.5541	0.6152	0.8402	0.6867	0.6404	0.8537	0.7543
HGT	0.2181	0.7092	0.4214	0.3315	0.7339	0.5349	0.3554	0.6766	0.6894
RGCN	0.6874	**0.8600**	0.6892	0.6866	0.8737	0.7286	0.6898	0.8886	0.7733
HeteroSAGE	0.5403	0.8081	0.5844	0.5521	**0.8962**	0.7074	0.5906	0.8726	0.7782
HGATE	0.2789	0.5106	0.7190	0.2132	0.6444	0.4495	0.2222	0.6274	0.5029
HGRAN	0.5007	0.7206	0.7024	0.7098	0.8375	0.7105	0.5040	0.7294	0.6000
MLP	0.3181	0.5606	0.4214	0.6415	0.8299	0.6799	0.2631	0.5116	0.6519
GCN	0.6462	0.8353	0.6571	0.6431	0.8324	0.6914	0.5148	0.8844	0.7532
GAT	0.5341	0.7384	0.5620	0.6334	0.8275	0.6900	0.6032	0.8068	0.7338
HeCoGNN	**0.7471**	0.8452	**0.7384**	**0.7273**	0.8792	**0.7483**	**0.8058**	**0.9498**	**0.8686**
Social Governance II									
HAN	0.3693	0.7039	0.7048	0.6619	0.8572	0.6927	0.5237	0.7735	0.6167
RGCN	0.4962	0.8035	0.6810	0.6700	**0.8676**	0.6743	0.5634	0.7781	0.6333
HGT	0.2864	0.6348	0.6619	0.3921	0.7943	0.5642	0.2717	0.6855	0.5000
HeteroSAGE	0.4874	0.7271	0.6429	0.6571	0.8387	0.6697	0.5787	0.7678	0.5667
HGATE	0.2549	0.5309	0.619	0.1933	0.6471	0.4083	0.2271	0.5697	0.5167
HGRAN	0.3202	0.619	0.6753	0.4002	0.7798	0.5810	0.4367	0.7215	0.5652
GCN	0.5161	0.7368	0.6333	0.6185	0.8522	0.6789	0.5675	0.8172	0.5500
GAT	0.5467	0.6830	0.6810	0.6188	0.8227	0.6560	0.5913	0.7995	0.5833
MLP	0.3125	0.5142	0.5667	**0.6791**	0.8529	**0.7110**	0.3505	0.5500	0.5500
HeCoGNN	**0.6310**	**0.8170**	**0.7238**	0.6774	0.8378	0.6881	**0.6889**	**0.8812**	**0.7333**

Table 2: Node classification performance (%). The best results are shown in bold, and the second best is underlined.

Model	Macro-F1	Micro-F1	AUC
HeCoGNN	0.8324	**0.8416**	**0.9258**
HGT	0.8264	0.8157	0.8906
RGCN	**0.8325**	0.8305	0.9125

Table 3: Performance comparison on the ACM dataset.The best results are shown in bold.

parameter study by replacing their respective architectures with alternative designs, including a Dynamic Attention-based Model, a Fixed Weight Aggregation Model, and a simple MLP. As shown in Figure 3, the configuration that employs MLP-based GateNet and Dynamic Attention-based ContextNet (**HeCoGNN (GateNet MLP, ContextNet Dynamic Attention)**) consistently outperforms other variants across all metrics and node types. In contrast, substituting either component with fixed or static counterparts leads to noticeable performance degradation, especially for feature-sparse node types such as Person and Location. This demonstrates the necessity of type-sensitive gating for suppressing irrelevant information and the importance of relation-aware attention in capturing nuanced semantic dependencies. Overall, the results underscore the critical role of both modules in enabling effective and balanced representation learning in heterogeneous graphs.

Ablation Study To further assess the effectiveness of each component within HeCoGNN, we perform an ablation study using three model variants: **HeCoGNN (GateNet + ContextNet + Fusion)**, which represents the full model with dynamic gating and relation-aware attention; **w/o GateNet (Only ContextNet)**, which removes the gating mechanism and relies solely on attention-based aggregation; and **w/o ContextNet (Only GateNet)**, which omits attention and uses only hard gating for message passing. As illustrated in Figure 4, both modules are indispensable. Removing either one leads to a substantial drop in classification accuracy, particularly for nodes with limited features. The performance degradation is most prominent on Person and Location nodes, indicating their greater reliance on structural and contextual cues. In contrast, the full HeCoGNN model consistently achieves the highest accuracy, validating the effectiveness of the collaborative design in mitigating imbalance and noise in real-world heterogeneous graphs.

Conclusion

In this paper, we propose HeCoGNN to tackle the critical challenges of severe feature imbalance and noise propagation in real-world heterogeneous social risk graphs. By introducing the collaborative mechanisms of GateNet and ContextNet, our model achieves adaptive edge selection and relation-aware information aggregation, effectively mitigating representation bias and enhancing the expressiveness of textless nodes. Extensive experiments on two real social governance datasets demonstrate that HeCoGNN significantly outperforms existing heterogeneous GNN baselines, especially in terms of balanced performance across different node types and robustness to structural noise. These results highlight the practical value of our framework for complex multi-type risk assessment tasks in heterogeneous networks.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2023YFC3304503), the National Natural Science Foundation of China (No. 62422210, 92370111, 62302333, 62272340, 62276187) and Hebei Natural Science Foundation (No. F2024202047).

References

Dong, Y.; Chawla, N. V.; and Swami, A. 2017. metapath2vec: Scalable Representation Learning for Heterogeneous Networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 135–144.

Fu, X.; Chen, Z.; He, Y.; Wang, S.; Zhang, B.; Chen, C.; and Li, J. 2025. Virtual Nodes Can Help: Tackling Distribution Shifts in Federated Graph Learning. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence*, 16657–16665.

Fu, X.; Zhang, J.; Meng, Z.; and King, I. 2020. MAGNN: Metapath Aggregated Graph Neural Network for Heterogeneous Graph Embedding. In *WWW '20: The Web Conference 2020*, 2331–2341.

Guan, M.; Stokes, J. W.; Luo, Q.; Liu, F.; Mehta, P.; Nouri, E.; and Kim, T. 2025. Heterogeneous Graph Neural Network on Semantic Tree. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence*, 16924–16932.

Hamilton, W. L.; Ying, Z.; and Leskovec, J. 2017. Inductive Representation Learning on Large Graphs. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, 1024–1034.

He, M.; Wei, Z.; Feng, S.; Huang, Z.; Li, W.; Sun, Y.; and Yu, D. 2024. Spectral Heterogeneous Graph Convolutions via Positive Noncommutative Polynomials. In *Proceedings of the ACM on Web Conference 2024, WWW 2024*, 685–696.

He, M.; Zhu, P.; Tang, K.; and Guo, Y. 2025. Hypergraph attacks via injecting homogeneous nodes into elite hyperedges. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 282–290.

Hong, H.; Guo, H.; Lin, Y.; Yang, X.; Li, Z.; and Ye, J. 2020. An Attention-Based Graph Neural Network for Heterogeneous Structural Learning. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020*, 4132–4139.

Hu, Y.; Zhang, Y.; Wang, Y.; and Work, D. B. 2023. Detecting Socially Abnormal Highway Driving Behaviors via Recurrent Graph Attention Networks. In *Proceedings of the ACM Web Conference 2023, WWW 2023*, 3086–3097.

Hu, Z.; Dong, Y.; Wang, K.; and Sun, Y. 2020. Heterogeneous Graph Transformer. In *WWW '20: The Web Conference 2020*, 2704–2710.

Hui, Y.; Zwetsloot, I. M.; Trimborn, S.; and Rudinac, S. 2025. Domain-Informed Negative Sampling Strategies for Dynamic Graph Embedding in Meme Stock-Related Social Networks. In *Proceedings of the ACM on Web Conference 2025, WWW 2025*, 518–529.

Jin, D.; Wang, L.; Zheng, Y.; Song, G.; Jiang, F.; Li, X.; Lin, W.; and Pan, S. 2023. Dual intent enhanced graph neural network for session-based new item recommendation. In *Proceedings of the ACM web conference 2023*, 684–693.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015,*.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *5th International Conference on Learning Representations, ICLR 2017, Conference Track Proceedings*.

Li, R.; Jin, D.; Wang, X.; He, D.; Feng, B.; and Wang, Z. 2025. Single-Node Trigger Backdoor Attacks in Graph-Based Recommendation Systems. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2025, Montreal, Canada, August 16-22, 2025*, 3072–3080.

Lu, Y.; Chen, C.; and Li, C. 2024. Dual Graph Networks with Synthetic Oversampling for Imbalanced Rumor Detection on Social Media. In *Companion Proceedings of the ACM on Web Conference 2024, WWW 2024*, 750–753.

Lv, Q.; Ding, M.; Liu, Q.; Chen, Y.; Feng, W.; He, S.; Zhou, C.; Jiang, J.; Dong, Y.; and Tang, J. 2021. Are we really making much progress?: Revisiting, benchmarking and refining heterogeneous graph neural networks. In *KDD '21: The 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1150–1160.

Potapczynski, A.; Loaiza-Ganem, G.; and Cunningham, J. P. 2020. Invertible Gaussian Reparameterization: Revisiting the Gumbel-Softmax. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020*.

Schlichtkrull, M. S.; Kipf, T. N.; Bloem, P.; van den Berg, R.; Titov, I.; and Welling, M. 2018. Modeling Relational Data with Graph Convolutional Networks. In *The Semantic Web - 15th International Conference, ESWC 2018, Proceedings*, 593–607.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, 5998–6008.

Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; and Bengio, Y. 2018. Graph Attention Networks. In *6th International Conference on Learning Representations, ICLR 2018, Conference Track Proceedings*.

Wang, J.; Li, Y.; Shao, Y.; Xue, Z.; Guan, Z.; Li, A.; and Ye, G. 2025a. Reinforcement Active Client Selection for Federated Heterogeneous Graph Learning. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence*, 21117–21125.

Wang, L.; Zheng, Y.; Jin, D.; Li, F.; Qiao, Y.; and Pan, S. 2024a. Contrastive graph similarity networks. *ACM Transactions on the Web*, 18(2): 1–20.

Wang, W.; Suo, X.; Wei, X.; Wang, B.; Wang, H.; Dai, H.; and Zhang, X. 2023a. HGATE: Heterogeneous Graph Attention Auto-Encoders. *IEEE Trans. Knowl. Data Eng.*, 35: 3938–3951.

Wang, X.; Deng, W.; Meng, Z.; and Chen, D. 2024b. Hybrid-attention mechanism based heterogeneous graph representation learning. *Expert Syst. Appl.*, 250: 123963.

Wang, X.; Dong, Y.; Jin, D.; Li, Y.; Wang, L.; and Dang, J. 2023b. Augmenting Affective Dependency Graph via Iterative Incongruity Graph Learning for Sarcasm Detection. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, 4702–4710.

Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; and Yu, P. S. 2019. Heterogeneous Graph Attention Network. In *The World Wide Web Conference, WWW 2019*, 2022–2032.

Wang, X.; Sun, R.; Zhang, Y.; Feng, B.; He, D.; Wang, L.; and Jin, D. 2025b. Stealthy Yet Effective: Distribution-Preserving Backdoor Attacks on Graph Classification.

Wang, X.; Wang, Y.; He, D.; Yu, Z.; Li, Y.; Wang, L.; Dang, J.; and Jin, D. 2025c. Elevating Knowledge-Enhanced Entity and Relationship Understanding for Sarcasm Detection. *IEEE Trans. Knowl. Data Eng.*, 37(6): 3356–3371.

Yan, F.; Wang, X.; He, D.; Wang, L.; Dang, J.; and Jin, D. 2025. HeterGP: Bridging Heterogeneity in Graph Neural Networks with Multi-View Prompting. In *AAAI-25, Sponsored by the Association for the Advancement of Artificial Intelligence, February 25 - March 4, 2025, Philadelphia, PA, USA*, 21895–21903. AAAI Press.

Yang, S.; Zhang, Z.; Zhou, J.; Wang, Y.; Sun, W.; Zhong, X.; Fang, Y.; Yu, Q.; and Qi, Y. 2020. Financial Risk Analysis for SMEs with Graph-based Supply Chain Mining. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*, 4661–4667.

Yang, X.; Yan, M.; Pan, S.; Ye, X.; and Fan, D. 2023. Simple and Efficient Heterogeneous Graph Neural Network. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023*, 10816–10824.

Zhang, C.; Song, D.; Huang, C.; Swami, A.; and Chawla, N. V. 2019. Heterogeneous Graph Neural Network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019*, 793–803.

Zhu, P.; Pan, Z.; Liu, Y.; Tian, J.; Tang, K.; and Wang, Z. 2024a. A general black-box adversarial attack on graph-based fake news detectors. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 568–576.

Zhu, P.; Pan, Z.; Tang, K.; Cui, X.; Wang, J.; and Xuan, Q. 2024b. Node injection attack based on label propagation against graph neural network. *IEEE Transactions on Computational Social Systems*, 11(5): 5858–5870.

Zhu, S.; Zhou, C.; Pan, S.; Zhu, X.; and Wang, B. 2019. Relation Structure-Aware Heterogeneous Graph Neural Network. In *2019 IEEE International Conference on Data Mining, ICDM 2019*, 1534–1539.

Zou, Z.; Liu, Z.; Shan, J.; Li, Q.; Xu, K.; and Xu, M. 2024. CoGNN: Towards Secure and Efficient Collaborative Graph Learning. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security, CCS 2024*, 4032–4046.