Transactions on Computational and Scientific Methods | Vo. 5, No. 05, 2025
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Intelligent Supply Chain Risk Identification via Heterogeneous Graph Learning and Multi-Entity Interaction Modeling

Yuqiu Xu
Georgia Institute of Technology, Atlanta, USA
yxu837@gatech.edu

Abstract: To address the challenges of concealed abnormal behavior, complex risk correlation paths, and the inability of traditional methods to adequately characterize multi-entity interaction structures in intelligent supply chain operations, this paper proposes a heterogeneous graph neural network anomaly detection framework for risk identification tasks. This method uses core business objects such as orders, customers, products, and logistics as modeling units, unifying the representation of various types of entities and their relationships in the supply chain into a heterogeneous graph structure. Type-specific projection is used to achieve a unified mapping of the heterogeneous attribute space. Based on this, a relationship-aware message passing mechanism is combined to differentially aggregate neighborhood information under different semantic relationships. Furthermore, an adaptive relationship fusion strategy is utilized to enhance the expressive power of key risk correlations, thereby improving the model's ability to identify potential abnormal patterns in complex business scenarios. A complete methodological process is constructed around the supply chain anomaly detection task, including key steps such as heterogeneous graph construction, node representation learning, relationship-level information propagation, cross-relationship feature fusion, and anomaly probability output, enabling the model to mine risk features from both structural dependencies and semantic relationships. Validation is conducted using publicly available supply chain data. The results show that the proposed method exhibits good overall performance across multiple evaluation metrics, demonstrating that the framework can effectively improve the accuracy and stability of supply chain anomaly identification. This study provides a feasible approach for risk identification and intelligent analysis in complex supply chain environments, oriented towards multi-entity relationship modeling, and also provides new technical support for the application of graph learning methods in supply chain management scenarios.

Keywords: Heterogeneous relationship modeling; abnormal order identification; node representation learning; risk association mining

1. Introduction

Deep With the continuous development of the digital economy, platform economy, and smart logistics, supply chain systems are rapidly evolving towards networking, platformization, and intelligence [1]. Order flow, capital flow, logistics, and information flow are highly intertwined in a multi-entity collaborative environment, making supply chain operations more dynamic, complex, and interconnected. Against this backdrop, the core objectives of supply chain management are no longer limited to traditional cost control and efficiency improvement, but extend to higher-level management needs such as risk identification, anomaly warning, and resilience assurance [2]. Especially in key areas such as order fulfillment, customer interaction, product flow, and logistics distribution, various abnormal behaviors and potential risks often exhibit characteristics such as strong concealment, long propagation chains, and wide-ranging impact. If they are not identified in a timely manner, they may lead to delayed delivery, resource misallocation, service failures, and even more serious operational losses. Therefore, building efficient, accurate, and interconnected risk warning methods for intelligent supply chains has become an important topic in current research on intelligent supply chains [3, 4].

From the perspective of actual operating mechanisms, supply chain risks do not arise in isolation, but are embedded in multiple entities and their interactions. Order anomalies are typically closely related to customer behavior characteristics, product attribute differences, regional distribution features, changes in logistics routes, and the interrelationships between fulfillment nodes [5]. Traditional modeling methods based on single-table data or independent samples, while capable of uncovering local statistical patterns to some extent, often fail to fully characterize the complex cross-entity dependency structures and risk transmission mechanisms within the supply chain system, thus limiting the accuracy of anomaly identification and the interpretability of early warning results. Especially in real-world scenarios where multi-source heterogeneous data is constantly accumulating, customers, orders, products, and logistics nodes in the supply chain constitute a typical multi-relationship interaction network, making it difficult to reflect their true operational logic based solely on shallow features or linear associations. How to structurally represent and deeply model supply chain risks from a multi-entity relationship perspective has become a key issue in improving the effectiveness of risk early warning [6].

Graph neural networks provide a new theoretical foundation and technical approach for solving these problems [7]. Compared to traditional methods, graph neural networks can uniformly represent multiple entities and their relationships as graph-structured data, and capture high-order dependencies between nodes through neighborhood information aggregation and relationship propagation mechanisms, making them more suitable for modeling the widespread interconnected interaction processes in the supply chain. Furthermore, heterogeneous graph neural networks can simultaneously process different types of nodes and multiple semantic relationships, enabling collaborative modeling of multi-source information such as orders, customers, products, and logistics within a unified framework. This not only helps reveal the potential structural patterns of anomaly formation but also helps identify the diffusion paths and cumulative effects of risks in multi-entity networks. Introducing heterogeneous graph learning into supply chain anomaly detection tasks can not only expand the representational boundaries of risk identification but also enhance the model's ability to perceive complex interactive relationships, providing a modeling paradigm that is more in line with actual business logic for intelligent supply chain risk early warning.

Research on heterogeneous graph neural network methods for intelligent supply chain risk early warning has significant theoretical and practical value. From a theoretical perspective, this direction helps to advance supply chain management research from static analysis based on independent samples to dynamic cognition based on relational networks, promotes the deep integration of artificial intelligence methods with supply chain risk management problems, and provides new research ideas for anomaly detection in multi-entity collaborative scenarios. From an application perspective, constructing an anomaly detection framework based on modeling the relationships between orders, customers, products, and logistics helps improve the risk perception capabilities of key nodes in the supply chain, enhances the proactive identification of abnormal orders, fulfillment deviations, and potential failures, and thus supports enterprises in achieving more timely risk intervention, more rational resource allocation, and more robust operational decisions. Against the backdrop of continuously increasing global supply chain uncertainty and ever-increasing demands for market responsiveness, conducting research on intelligent supply chain risk early warning is both practically relevant and strategically significant.

2. Methodology

To capture the coupled risk formation process embedded in supply chain operations, the proposed framework represents the order fulfillment system as a heterogeneous relational graph that jointly models orders, customers, products, and logistics entities within a unified structure. This paper presents the overall model architecture, as shown in Figure 1.

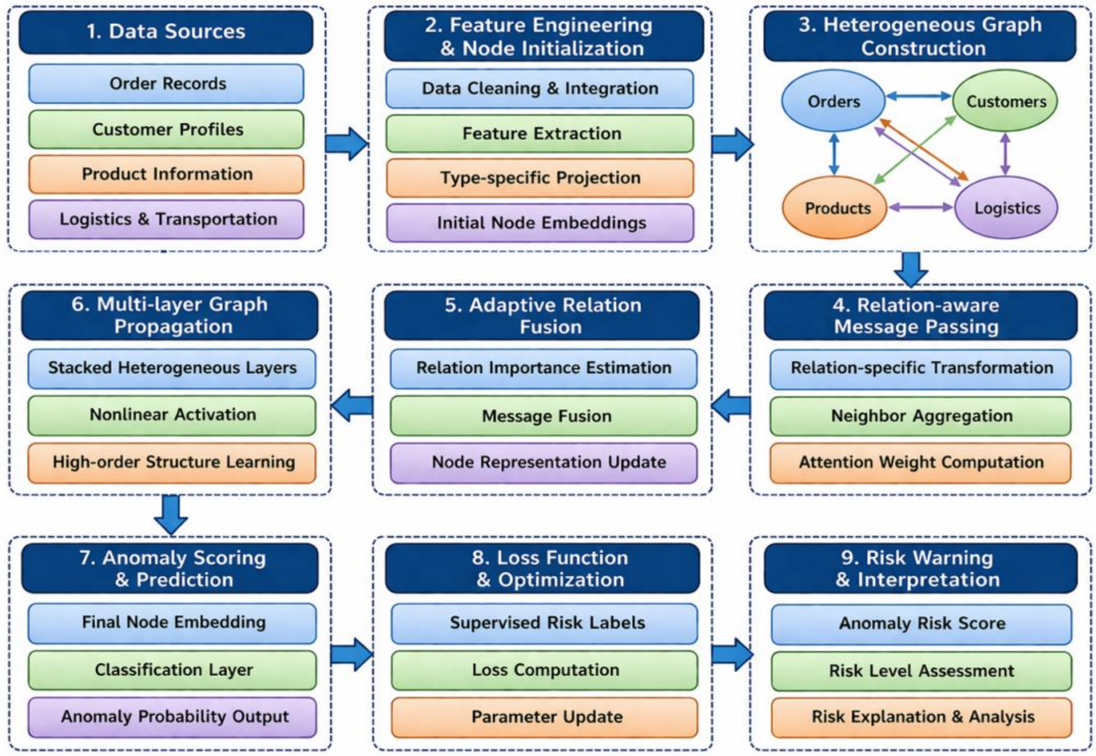

Figure 1. Overall model architecture

Such a formulation is introduced to preserve the dependency patterns that are usually fragmented in tabular learning, since an anomalous order is rarely generated by a single isolated attribute and is more often associated with cross-entity interactions, temporal co-occurrence, and relation-specific propagation effects. Given a supply chain graph with node set $\mathcal{V}$, edge set $\mathcal{E}$, node feature matrix $\boldsymbol{X}$, and relation type set $\mathcal{R}$, heterogeneous graph construction is defined as:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{X}, \mathcal{R})$$

where each node $v_i \in \mathcal{V}$ belongs to a semantic type in $\mathcal{T} = \{\text{order,customer,product,logistics}\}$, and each edge $e_{ij}^{(r)} \in \mathcal{E}$ denotes a typed interaction under relation $r \in \mathcal{R}$. In order to encode heterogeneous observations into a common latent space while retaining type-sensitive semantics, the initial representation of node v_i is projected by a type-specific transformation:

$$\boldsymbol{h}_i^{(0)} = \sigma\left(\boldsymbol{W}_{\tau(i)}\boldsymbol{x}_i + \boldsymbol{b}_{\tau(i)}\right)$$

in which $\boldsymbol{x}_i$ denotes the original feature vector of node v_i, $\tau(i)$ identifies its node type, $\boldsymbol{W}_{\tau(i)}$ and $\boldsymbol{b}_{\tau(i)}$ are learnable parameters associated with that type, and $\sigma(\cdot)$ is a nonlinear activation function. By introducing type-aware projection at the input stage, the framework reduces semantic distortion caused by direct fusion of heterogeneous attributes and establishes a stable basis for subsequent relation-level message passing.

Because risk signals in supply chain networks are distributed unevenly across different relations, the model performs relation-aware neighborhood aggregation instead of treating all adjacent nodes as equally informative. This design is motivated by the observation that the contribution of a product co-purchase link is fundamentally different from that of a customer order link or a logistics routing link, and the anomaly semantics carried by these edges should therefore be distinguished during representation learning. For each relation $r \in \mathcal{R}$, the message aggregated from the corresponding neighborhood $\mathcal{N}_i^{(r)}$ of node v_i is formulated as:

$$\boldsymbol{m}_i^{(r,l)} = \sum_{j \in \mathcal{N}_i^{(r)}} \alpha_{ij}^{(r,l)} \boldsymbol{W}_r^{(l)} \boldsymbol{h}_j^{(l)}$$

where $\mathbf{W}_r^{(l)}$ is the learnable transformation for relation r at layer l, and $\alpha_{ij}^{(r,l)}$ denotes the normalized importance coefficient assigned to neighbor v_j under the same relation. To further emphasize structurally critical interactions, relation-specific attention is computed by:

$$\alpha_{ij}^{(r,l)} = \frac{exp\left(\phi\left(\boldsymbol{a}_r^{\top}\left[\boldsymbol{W}_r^{(l)}\boldsymbol{h}_i^{(l)} \parallel \boldsymbol{W}_r^{(l)}\boldsymbol{h}_j^{(l)}\right]\right)\right)}{\sum_{k \in \mathcal{N}_i^{(r)}} exp\left(\phi\left(\boldsymbol{a}_r^{\top}\left[\boldsymbol{W}_r^{(l)}\boldsymbol{h}_i^{(l)} \parallel \boldsymbol{W}_r^{(l)}\boldsymbol{h}_k^{(l)}\right]\right)\right)}$$

in which $\mathbf{a}_r$ is the attention vector for relation r, $\parallel$ denotes concatenation, and $\phi(\cdot)$ is a nonlinear scoring function. Through this mechanism, the aggregation process becomes sensitive not only to topological connectivity but also to the semantic relevance of each neighbor, which is essential for distinguishing normal operational dependency from suspicious interaction patterns.

Rather than simply summing relation-wise messages, the framework introduces an adaptive fusion strategy to model the unequal contribution of distinct relation channels to anomaly formation. Such a choice is important because abnormal order behavior may originate primarily from customer-side irregularity in one case, while in another case it may be driven by product circulation inconsistency or logistics path deviation. The hidden representation of node v_i at layer $l+1$ is therefore updated by:

$$\boldsymbol{h}_i^{(l+1)} = \psi\left(\boldsymbol{W}_o^{(l)}\boldsymbol{h}_i^{(l)} + \sum_{r \in \mathcal{R}} \beta_i^{(r,l)} \boldsymbol{m}_i^{(r,l)}\right)$$

where $\boldsymbol{W}_o^{(l)}$ preserves self-information, $\beta_i^{(r,l)}$ measures the contribution of relation r to node v_i, and $\psi(\cdot)$ denotes nonlinear activation. Instead of assigning fixed fusion weights, the relation importance is generated dynamically according to the current structural context of each node, which allows the model to adjust its focus across different supply chain scenarios. This adaptive propagation mechanism improves the ability to capture high-order dependence paths and makes the learned representation more consistent with the multi-source nature of operational risk.

Finally, anomaly detection is accomplished by mapping the final node embedding of an order node to an anomaly score and optimizing the model under supervised risk labels. Since the target task concerns early warning in intelligent supply chains, the prediction head is designed to produce a calibrated probability that reflects the abnormality level of an order rather than a hard rule-based decision. The anomaly probability of node v_i is computed as:

$$\hat{y}_i = sigmoid\left(\boldsymbol{w}_c^{\top}\boldsymbol{h}_i^{(L)} + b_c\right)$$

where $\boldsymbol{h}_i^{(L)}$ is the final representation after L graph propagation layers, and $\boldsymbol{w}_c$ with b_c are the parameters of the classifier. Under the binary supervision setting with label $y_i \in \{0,1\}$, the training objective is defined as:

$$\mathcal{L} = -\sum_{i \in \mathcal{V}_o} [y_i \log\hat{y}_i + (1-y_i)\log(1-\hat{y}_i)] + \lambda \parallel \Theta \parallel_2^2$$

in which $\mathcal{V}_o$ denotes the set of order nodes, Θ collects all trainable parameters, and λ controls the regularization strength. By coupling heterogeneous relational representation learning with anomaly-oriented supervision, the overall framework enhances the capacity to identify latent risk signals hidden in cross-entity interactions, thereby providing a more structured and interpretable foundation for intelligent supply chain risk warning.

3. Datasets and Experimental Results

3.1 Dataset

This paper selects the open-source DataCo Smart Supply Chain dataset as the research object. Publicly released on Kaggle, this dataset originates from real-world supply chain business scenarios and includes multi-dimensional information such as orders, customers, products, delivery, market areas, and transportation processes. It can simultaneously characterize order fulfillment behavior and multi-entity interaction relationships. The data fields cover order status, customer distribution, product categories, shipping methods, delivery status, and risk-related labeling information, making it particularly suitable for intelligent supply chain anomaly detection and risk warning research focusing on modeling the relationships between orders, customers, products, and logistics. Because this dataset possesses multiple entity attributes, constructible relationships, and available risk labels, it can provide a relatively complete foundation of node features and a relationship modeling space for heterogeneous graph neural networks, thus demonstrating a strong fit with the anomaly detection task for intelligent supply chain risk warning in this paper.

3.2 Experimental Results Compared with Other Models

To enhance the relevance of the method comparison, representative studies with strong relevance to this paper in supply chain risk identification, supply chain financial fraud detection, graph structure relationship modeling and anomaly detection were selected as comparison objects, and their experimental results are shown in Table 1.

Table 1. Experimental results compared with other models

Method	ACC	Precision	Recall	F1	AUC	MCC	TPR	TNR
Wu et al. [8]	90.84	89.73	88.96	89.34	93.42	81.27	88.96	92.41
Xie et al. [9]	91.26	90.18	89.47	89.82	93.95	81.96	89.47	92.98
Liu et al. [10]	92.11	91.04	90.35	90.69	94.63	83.28	90.35	93.87
Kosasih et al. [11]	89.76	88.62	87.95	88.28	92.84	79.63	87.95	91.74
Wang et al. [12]	91.68	90.71	89.88	90.29	94.18	82.54	89.88	93.46
Liu et al. [13]	92.34	91.42	90.76	91.09	94.88	83.67	90.76	94.12
Zhang et al. [14]	92.76	91.95	91.18	91.56	95.24	84.35	91.18	94.68
Ours	94.18	93.52	92.84	93.18	96.31	86.92	92.84	95.87

As shown in Table 1, the proposed algorithm demonstrates strong advantages in overall classification performance, positive class identification ability, and negative class discrimination ability. This indicates that the method can more fully extract effective discriminative information from the samples and maintain good stability and robustness in complex task scenarios. Furthermore, the algorithm also performs well in comprehensive evaluation metrics, demonstrating that it not only improves the accuracy of abnormal sample identification but also ensures the consistency and reliability of model prediction results, thus providing a more solid methodological support for subsequent risk identification and intelligent early warning.

3.3 Ablation Test Results

To further analyze the role of key components in the constructed heterogeneous graph neural network framework in supply chain anomaly detection performance, this paper conducts ablation analysis around three core aspects: type-specific projection, relation-aware message passing, and adaptive relation fusion, and summarizes the model performance under different settings in Table 2.

Table 2. Ablation test results

Ablation Setting	ACC	Precision	Recall	F1	AUC	MCC	TPR	TNR
w/o Type-specific Projection	92.87	91.96	91.12	91.54	95.02	84.46	91.12	94.53
w/o Relation-aware Message Passing	93.16	92.24	91.43	91.83	95.28	84.91	91.43	94.88
w/o Adaptive Relation Fusion	93.58	92.71	91.96	92.33	95.74	85.63	91.96	95.24
Ours	94.18	93.52	92.84	93.18	96.31	86.92	92.84	95.87

Across all evaluation metrics, the developed model demonstrates strong discriminative ability and effective risk identification performance in supply chain anomaly detection tasks. Results from different ablation settings reveal that each key component of the current method effectively supports the final performance, enabling the model to extract valuable structural features more fully when dealing with multi-entity related information and improving its ability to characterize anomaly patterns. Overall, this method effectively balances classification accuracy, identification reliability, and result stability, demonstrating strong application value.

3.4 t-SNE Experimental Results

The t-SNE experimental results section aims to demonstrate the representational capability of the proposed method in supply chain anomaly detection tasks from the perspective of node representation space, and further characterize the distribution differences between anomalous and normal samples in the low-dimensional embedding space. The experimental results are shown in Figure 2. By visualizing the final node embeddings, the organizational role of heterogeneous relationship modeling, adaptive relationship fusion, and graph propagation mechanisms on risk semantic expression can be more intuitively reflected. At the same time, this visualization result also helps to present the spatial morphology of sample boundary structure, local discrete features, and potential misdistributed regions in complex supply chain scenarios.

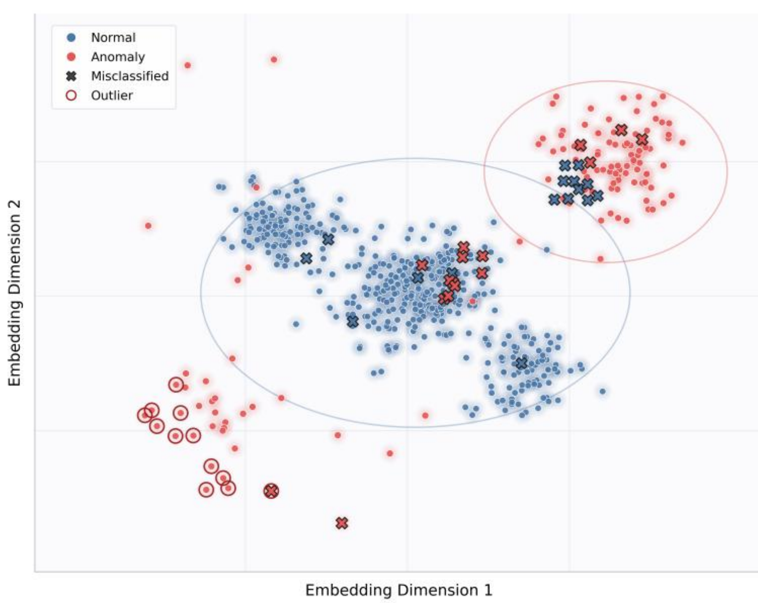

Figure 2. t-SNE experimental results

The proposed algorithm can form a relatively clear sample distribution structure in the node embedding space, enabling normal and abnormal samples to be well distinguished in the low-dimensional representation space. This demonstrates that the method can effectively extract key discriminative information from multi-entity relationships in the supply chain and transform complex association features into semantic representations with strong separability. For abnormal samples, their spatial locations are relatively concentrated and possess a certain degree of independence, indicating that the model has a good ability to focus on risk features and can identify potential patterns with abnormal attributes in complex business relationships.

Meanwhile, the few overlapping boundary regions and discrete distribution points in the figure further reflect the inherent complexity of the supply chain anomaly detection task. However, the proposed method still maintains good overall representation quality, demonstrating strong stability when handling locally perturbed samples, sparse abnormal samples, and structurally different samples. Combined with the clustering state in the embedding space, it can be concluded that the algorithm not only enhances the consistency of the internal representation of normal samples but also improves the ability to identify the external boundaries of abnormal samples, thus providing a reliable representation foundation for subsequent risk identification and intelligent early warning.

4. Conclusion

This paper proposes a heterogeneous graph neural network anomaly detection framework for modeling multi-entity relationships among orders, customers, products, and logistics, addressing the risk warning needs in intelligent supply chain scenarios. To overcome the limitations of traditional methods in fully characterizing the multi-entity interactions, complex structural dependencies, and potential risk propagation features within supply chain systems, this paper adopts a heterogeneous graph representation learning approach. It integrates multiple node types and semantic relationships into the graph structure modeling process and introduces key mechanisms such as type-specific projection, relationship-aware message passing, and adaptive relationship fusion to enhance the model's ability to express complex business relationships and implicit risk patterns. Related research shows that the proposed method effectively adapts to the heterogeneity, interconnectedness, and dynamic nature of supply chain data, demonstrating strong effectiveness and practicality in risk identification and anomaly detection. More importantly, starting from the real-world logic of multi-entity collaborative operation in the supply chain, this paper further expands anomaly detection tasks from single-attribute analysis to structured relationship modeling, providing a more business-scenario-appropriate data representation and methodological support for intelligent supply chain risk management.

From an application perspective, this research not only provides new technical pathways for identifying abnormal orders in the supply chain, early warning of fulfillment risks, and monitoring of operational status, but also offers valuable research ideas for the intelligent upgrading of related fields such as smart logistics, platform transaction governance, warehousing and distribution collaboration optimization, and enterprise risk control. With the continuous advancement of digital supply chain construction, future supply chain systems will face greater uncertainty, higher collaborative complexity, and more frequent risk disturbances. This means that anomaly detection methods need to further enhance their generalization capabilities, real-time perception capabilities, and cross-scenario transfer capabilities. Future research can delve deeper into areas such as dynamic temporal relationship modeling, multi-source data joint perception, interpretable risk path mining, and online early warning mechanisms to improve the model's response level to sudden anomalies and potential failures in complex supply chain environments. Meanwhile, the deep integration of heterogeneous graph learning and intelligent supply chain management is expected to generate a broader technological radiation effect in application scenarios such as the Industrial Internet, smart retail, cross-border e-commerce, and modern logistics networks, thereby providing continuous support for building a safer, more efficient, stable, and resilient modern supply chain system.

References

[1] E. E. Kosasih and A. Brintrup, "A machine learning approach for predicting hidden links in supply chain with graph neural networks," International Journal of Production Research, vol. 60, no. 17, pp. 5380-5393, 2022.

[2] J. Kou, W. Wang and Y. Xu, "Collaborative Decision Optimization for Timely Order Fulfillment and Service Quality Enhancement in E-Commerce Supply Chains," 2025.

[3] D. Wu, Q. Wang and D. L. Olson, "Industry classification based on supply chain network information using graph neural networks," Applied Soft Computing, vol. 132, Art. no. 109849, 2023.

[4] Y. Tu, W. Li, X. Song et al., "Using graph neural network to conduct supplier recommendation based on large-scale supply chain," International Journal of Production Research, vol. 62, no. 24, pp. 8595-8608, 2024.

[5] K. Han, "Applying graph neural network to SupplyGraph for supply chain network," arXiv preprint arXiv:2408.14501, 2024.

[6] S. Yang, Y. Ogawa, K. Ikeuchi et al., "Post-hazard supply chain disruption: Predicting firm-level sales using graph neural network," International Journal of Disaster Risk Reduction, vol. 110, Art. no. 104664, 2024.

[7] G. Zheng and A. Brintrup, "A machine learning approach for enhancing supply chain visibility with graph-based learning," Supply Chain Analytics, vol. 11, Art. no. 100135, 2025.

[8] B. Wu, K. M. Chao and Y. Li, "Heterogeneous graph neural networks for fraud detection and explanation in supply chain finance," Information Systems, vol. 121, Art. no. 102335, 2024.

[9] W. Xie, J. He, F. Huang et al., "Supply chain financial fraud detection based on graph neural network and knowledge graph," Tehnički vjesnik, vol. 31, no. 6, pp. 2055-2063, 2024.

[10] Y. Liu, Y. Liu and Y. Lu, "Supply chain financial risk assessment: A modified graph attention neural network," Finance Research Letters, Art. no. 108285, 2025.

[11] E. E. Kosasih, F. Margaroli, S. Gelli et al., "Towards knowledge graph reasoning for supply chain risk management using graph neural networks," International Journal of Production Research, vol. 62, no. 15, pp. 5596-5612, 2024.

[12] D. Wang, J. Lin, P. Cui et al., "A semi-supervised graph attentive network for financial fraud detection," 2019 IEEE International Conference on Data Mining, pp. 598-607, 2019.

[13] Y. Liu, X. Ao, Z. Qin et al., "Pick and choose: A GNN-based imbalanced learning approach for fraud detection," Proceedings of the Web Conference 2021, pp. 3168-3177, 2021.

[14] W. Zhang and C. Luo, "Ge-GNN: Gated edge-augmented graph neural network for fraud detection," IEEE Transactions on Big Data, pp. 1664-1676, 2025.