Original Research

Multi-Criteria Attention-Based Graph Neural Network: A Heterogeneous Representation Learning Framework for Logistics System Optimization

Mohammad Shahbazi[1]. Hamid Tohidi[1*]. Majid Nojavan[1]

Received: 04 Feb 2025/ Accepted: 13 March 2025/ Published online: 17 March 2025

*Corresponding Author, h_tohidi@azad.ac.ir

1-Department of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

Abstract
Modeling the intricate relationships within complex logistics systems is essential for optimizing various operations—such as routing, scheduling, and distribution—in modern supply chains. These systems often exhibit significant diversity in their facilities, transportation modes, and capacity constraints, introducing a phenomenon known as “heterogeneity,” which complicates the modeling process. To simplify calculations, some researchers assume homogeneous systems, overlooking critical variability in nodes (e.g., warehouses, distribution centers) and edges (e.g., transportation routes, capacities). However, ignoring this heterogeneity can lead to a marked decrease in model accuracy. In this paper, a representation learning method specifically tailored for heterogeneous logistics systems is proposed, in which the multifaceted relationships among components are preserved and model performance in real-world scenarios is enhanced. Two novel extensions refine the underlying graph-based deep learning architecture by incorporating techniques from deep learning, graph probability models, and machine learning. The approach is evaluated on two popular Vehicle Routing Problem with Time Windows (VRPTW) datasets, using precision, F1 score, and recall as performance metrics. Experimental results indicate that this method outperforms existing approaches by providing higher precision and F1 scores, enabling more accurate classification of system components and better extraction of relationships within complex logistics networks.

Keywords- Machine learning; Deep learning; Representation learning; Heterogeneous systems; Logistics optimization

INTRODUCTION

Modeling is a critical step in analyzing big data generated by logistics and supply chain operations. Incomplete or imprecise modeling can omit essential relationships within raw data, ultimately reducing the accuracy of subsequent analytical models. Data is frequently categorized into structured (e.g., tabular demand records) and unstructured (e.g., route networks, vehicle tracking logs). Graph modeling is particularly valuable for representing unstructured data in logistics, as many distribution and transportation networks include diverse nodes (facilities, vehicles, customers) and edges (shipment routes, capacity constraints) [1]. In such heterogeneous logistics systems, disregarding the diversity of components introduces a fundamental challenge in modeling. Various supply chain entities and transport modes often require flexible forms of representation. Consequently, capturing the complexity of real-world relationships remains a priority, especially in large-scale networks [2, 3]. In parallel, researchers have noted that analyzing these multifaceted connections can be aided by data mining techniques on link structures, social interactions, and hyperlinked documents [4, 5, 6]. Representation learning provides a framework that allows feature representations to be learned automatically from data [7]. This approach is particularly powerful in logistics scenarios when the variety of nodes (e.g., hubs, warehouses, trucks) and edges (e.g., dynamic route constraints) must be captured without oversimplifications. However, many modeling efforts still assume a homogeneous structure [8], despite recent work suggesting that heterogeneity can reveal important patterns and improve predictive accuracy [9].

Contributions of this study

- **Novel Multi-Criteria Attention-Based GNN:** A GNN framework is proposed that captures heterogeneity in logistics systems by considering multiple types of nodes and edges simultaneously.
- **Two Architectural Extensions:** Two extensions are developed: (i) a feature refinement strategy and (ii) multiple aggregation functions, which enhance embedding quality for more accurate classification in complex VRPTW datasets.
- **Comprehensive Evaluation:** Evaluation is performed on two well-known VRPTW benchmarks (Solomon, Gehring & Homberger) and faster convergence and consistently higher F1 scores than previous methods are demonstrated.

LITERATURE REVIEW

Over the past two decades, a range of algorithms and models have been proposed for learning low-dimensional embeddings of graph-structured data. Early studies primarily leveraged traditional graph-mining techniques to identify patterns in social, information, or web networks [2, 3, 4]. Later, researchers recognized that these ideas could be extended and refined through automatic representation learning methods.

I. Unsupervised Representation Learning

One line of research focuses on extracting meaningful features from unlabeled graph data. [10] introduced DeepWalk, which uses random walks to capture local neighborhood information. Similarly, node2vec [11] generalizes this idea by applying biased walks to better sample diverse neighborhoods. Both methods excel at micro-level tasks such as node classification. Building on these, approaches like SDNE [12] and DNGR [13] adopt deep architectures (autoencoders) to preserve non-linear proximity. Although these unsupervised methods can uncover important structural patterns, they frequently do not incorporate additional node attributes (e.g., capacity or scheduling constraints) that are critical for logistics operations.

II. Semi-Supervised Representation Learning

When partial labels are available, semi-supervised approaches can blend labeled and unlabeled data to learn more robust embeddings. Early frameworks like graph convolutional networks introduced refined message-passing schemes for node classification tasks [14]. Further improvements were made using attention mechanisms, such as GAT [15], to better handle noisy graph neighborhoods. For heterogeneous networks, algorithms like HAN [16] and HetSANN [17] used meta-paths or hierarchical attention to explicitly model multiple node and edge types. This approach is more aligned with heterogeneous logistics networks that may involve depots, distribution centers, and diverse vehicles or shipping routes.

III. Summary and Motivation

Despite these advances, the majority of existing works primarily address micro-level predictions (e.g., node classification, link prediction) without explicitly focusing on complex operational constraints present in logistics [8]. Many ignore the richness of node attributes by assuming uniform embeddings. In contrast, our proposed method leverages multi-criteria attention and meta-path expansions to capture subtle relationships essential for optimizing routes or schedules in large-scale supply chains. Table 1 illustrates some representative methods. We group them by methodology (random walk, deep learning, hybrid) and list their main advantages and disadvantages for potential logistics applications.

TABLE 1
RELATED STUDIES IN NETWORK REPRESENTATION LEARNING.

Method Analysis	ALGORITHM [REFERENCE]	Advantages / Disadvantages
Random walk	DeepWalk [10], Node2Vec [11]	+ Preserves local neighborhood structure – Neglects node attributes, typically micro-level
Deep learning	SDNE [12], DNGR [13]	+ Learns complex, non-linear embeddings – Higher risk of overfitting, ignores heterogeneity
Hybrid	GAT [15], R-GCN [14], HAN [16], HetSANN [17]	+ Uses partial labels, meta-paths for diverse types – Often focused on node-level tasks, needs heavy tuning

PROPOSED MODEL

A novel graph-based deep learning framework is presented that captures heterogeneity in logistics systems. The method leverages a graph neural network (GNN) architecture, enriched by meta-paths and multi-criteria attention mechanisms, allowing multiple node and edge types to be modeled effectively.

I. Meta-Path Integration

Meta-paths are used to reveal hidden indirect relationships, especially when logistics networks have specialized transport modes or multi-hop connections between hubs. By selecting meta-paths that encode domain-specific constraints (e.g., capacity limits, preferred transport modes), relevant relationships are captured in the embedding.

II. Neighborhood Expansion via Graph Edges

To reinforce each node's representation, its neighborhood is expanded through both direct edges and meta-paths. A ring-like structure is employed so that nodes can retain their own features (such as capacity, cost, or time-window constraints) in the final embedding. An illustration of neighbor contributions to the embedding of a central node is provided in Figure 1.

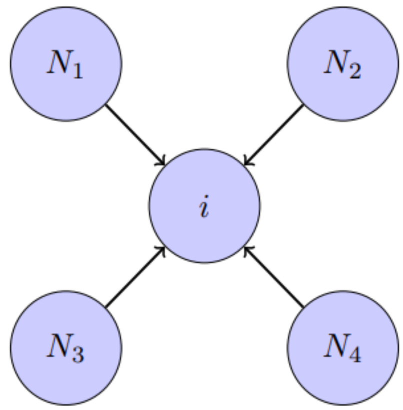

FIGURE 1
ILLUSTRATION OF NEIGHBOR CONTRIBUTIONS TO THE EMBEDDING OF A CENTRAL NODE I.

III. Unified Embedding and Attention

All features are projected into a shared embedding dimension. An attention layer is then applied to weight each neighboring node, with distinct coefficients for different edge or meta-path types. This ensures that the model is focused on the most relevant neighbors in highly heterogeneous logistics systems. The overall data flow of the proposed model is depicted in Figure 2.

FIGURE 2
DATA FLOW IN THE PROPOSED MODEL

IV. Representation Vector Generation

After attention-based weighting, node features are aggregated to yield a robust embedding that captures both local and indirect relationships, which is crucial for tasks such as routing, scheduling, and facility placement.

V. Representation Learning Block: Mathematical Formulation

Within each block (or GNN layer), computations are performed either in parallel or sequentially:

• Let h_{j}^{l} be denoted as the features of node j at layer l. A linear transformation is performed as:

$$h'^{(l+1)}_{j} = W^{(l+1)}_{\varphi(\mathrm{j})} h^{l}_{\mathrm{j}}, \tag{1}$$

where $W^{(l+1)}_{\varphi(\mathrm{j})}$ is a type-specific weight matrix, and $\varphi(\mathrm{j})$ is used to indicate the type of node j.

• For an edge $e = (i, j, r)$ of type r, the attention coefficient is computed as:

$$\eta^{(l+1)}_{\mathrm{e}} = \sigma([h'^{(l+1)}_{i} || h'^{(l+1)}_{j}] w^{(l+1)}_{\mathrm{r}}), \tag{2}$$

and softmax normalization is then applied:

$$\eta'^{(l+1)}_{\mathrm{e}} = \frac{e^{\eta^{(l+1)}_{\mathrm{e}}}}{\sum_{k\in E_i}^{K} e^{\eta^{(l+1)}_{\mathrm{k}}}}, \tag{3}$$

where E_i denotes the set of edges incident on node i.

• The new embedding for node i is then computed as:

$$h_i^{(l+1)} = \sigma(\sum_{e=(i,j,r)\in E_i} \eta'^{(l+1)}_{\mathrm{e}} h_j^{(l+1)}). \tag{4}$$

- ***First Extension (Feature Refinement)***

Before the linear transformation in (1), a feature-level attention is applied:

$$h'^{(l+1)}_j = att^{(l+1)}_{\varphi(\mathrm{j})} h^l_{\mathrm{j}}, \tag{5}$$

and then

$$h''^{(l+1)}_j = W^{(l+1)}_{\varphi(\mathrm{j})} h'^l_{\mathrm{j}}. \tag{6}$$

- ***Second Extension (Multiple Aggregation Functions)***

Instead of a single aggregator, multiple functions (e.g., mean, max, variance) are combined. For the t -th embedding dimension:

$$a_t^{(l+1)}[k] = A_k\left(\eta'^{(l+1)}_{e_1} h^{(l+1)}_{j_1}[t], \dots, \eta'^{(l+1)}_{e_n} h^{(l+1)}_{j_n}[t]\right) \dots \tag{7}$$

where A_k is an aggregation operator. Then,

$$h_i^{(l+1)}[t] = att_{h(i)}\left(a_t^{(l+1)}[1], \dots, a_t^{(l+1)}[m]\right). \tag{8}$$

- **Notation Explanation**

It is noted that the following indices and symbols have been used throughout the mathematical formulation:

• i is used to denote a node in the graph.

• j is used to denote a neighboring node of i.

• e is used to represent an edge in the graph and is often denoted as (i, j, r).

• r is used to indicate the type of edge or relationship (e.g., transport mode, capacity constraint).

• l is used to denote the layer in the graph neural network (GNN).

• φ(j) is used to indicate the type of node j (e.g., warehouse, hub, vehicle).

• t is used to denote each dimension of the embedding or feature vector.

• k is used as an index when multiple aggregation functions are employed.

• E_i is used to denote the set of edges incident to node i.

RESULTS

The method is evaluated on two widely used VRPTW datasets: the Solomon dataset (problems C1, C2, R1, R2) and the Gehring & Homberger dataset. Both contain multi-depot, capacity-constrained, and time-window components, making them highly representative for heterogeneous logistics networks.

I. Model Evaluation and Discussion

Precision, recall, and F1 scores (macro and micro) are reported. The F1 measure is defined as:

$$\text{F1} = 2\,\frac{\text{precision.recall}}{\text{precision+recall}} \tag{9}$$

Tables 1 and 2 compare the performance of the proposed model with various baselines.

- **Comparison on Solomon VRPTW**

Table 2 shows that the proposed model outperforms prior methods by roughly 2–4% in both macro and micro F1 scores.

TABLE 2
COMPARISON ON THE SOLOMON VRPTW DATASET (C1, C2, R1, R2).

Method	Macro F1	Micro F1
DeepWalk [23]	78.32	79.12
Metapath2Vec [41]	74.22	74.69
HERec [37]	82.93	83.32
GCN [42]	84.66	84.99
HAN [36]	85.20	85.66
R-GCN [42]	85.31	85.54
GAT [35]	86.59	86.78
HetSANN [37]	87.12	87.46
HetSANN.M [37]	88.33	88.59
HetSANN.M.R [37]	88.78	89.18
HetSANN.M.R.V [37]	89.12	89.46
Proposed Model	91.48	91.95

- **Comparison on Gehring & Homberger**

Table 3 shows that the proposed model also achieves superior performance on the Gehring & Homberger dataset.

TABLE 3
COMPARISON ON THE SOLOMON VRPTW DATASET (C1, C2, R1, R2).

Method	Macro F1	Micro F1
DeepWalk [23]	77.95	78.48
Metapath2Vec [41]	74.67	75.22
HERec [37]	83.57	84.16
GCN [42]	85.63	86.02
HAN [36]	86.15	86.71
R-GCN [42]	86.41	86.85
GAT [35]	87.78	88.19
HetSANN [37]	88.46	88.67
HetSANN.M [37]	89.11	89.45
HetSANN.M.R [37]	89.67	90.03
HetSANN.M.R.V [37]	90.08	90.46
Proposed Model	92.72	93.01

- **Discussion and Convergence Analysis**

- A 2–4% improvement in F1 is observed, highlighting the importance of modeling heterogeneous edges.
- Meta-path-based expansions capture indirect relationships that purely local methods tend to miss.
- The approach converges quickly despite additional architectural complexity.

Figure 3 plots the F1 score (%) versus training epochs for the proposed model and a strong baseline, demonstrating faster convergence and higher final accuracy.

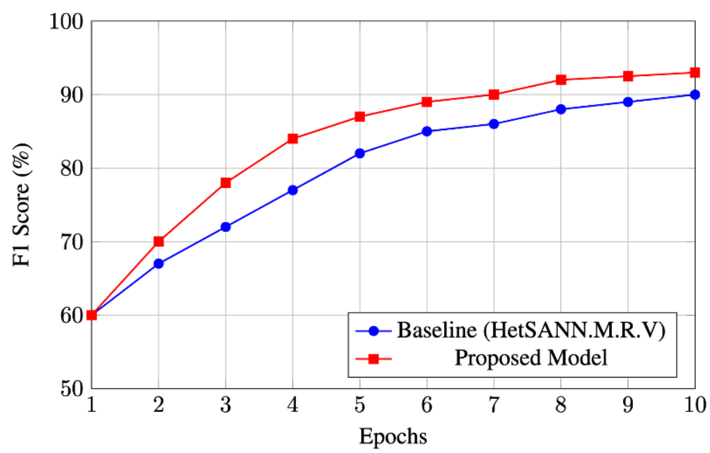

FIGURE 3
CONVERGENCE ANALYSIS ON VRPTW DATA.

CONCLUSION

In this study, a novel graph-based deep learning framework was proposed that employs multi-criteria attention mechanisms along with two key architectural extensions—feature refinement and multiple aggregation—to improve the representation of heterogeneous components within logistics systems. The model was specifically designed to address the complexity of real-world transportation and distribution networks, in which diverse nodes (e.g., facilities, hubs, vehicles) and edges (e.g., routes, capacity constraints) can significantly affect routing, scheduling, and other optimization tasks.

Initially, meta-paths were utilized to uncover hidden, indirect relationships among nodes, ensuring that relevant domain-specific interactions—such as capacity limits or multi-modal connections—were captured in the representation space. Subsequently, a multi-criteria attention module was implemented to selectively weight nodes and edges based on their importance, thereby facilitating a more focused and accurate modeling of complex logistics data. This attention-based weighting allowed the network to concentrate on the most critical elements, preventing dilution of valuable information in node and edge features.The evaluation was conducted using two well-known Vehicle Routing Problem with Time Windows (VRPTW) datasets: Solomon and Gehring & Homberger. These datasets incorporate various operational constraints, such as multiple depots, time windows, and vehicle capacity limitations, making them highly representative testbeds for heterogeneous logistics networks. The experimental results demonstrated that the proposed framework achieved a 2–4% improvement in F1 score compared to existing baseline methods, indicating its enhanced capability to classify and predict system components accurately under challenging, real-world conditions.

The feature-refinement extension played a key role by filtering and strengthening the input features before embedding, thereby preserving high-level information essential for downstream tasks. Furthermore, the multi-aggregation mechanism effectively captured a wide range of statistical characteristics within local neighborhoods—from means to variances and maxima—alleviating potential inaccuracies caused by the diversity and imbalance commonly observed in large-scale logistics networks. This multi-aggregation approach enabled the model to handle fluctuations and structural changes in the network more robustly, contributing to the observed performance gains. In addition to consistently outperforming baselines in precision and recall, the proposed method exhibited faster convergence, suggesting its suitability for scenarios requiring rapid or online inference. This higher convergence speed underscores the architecture's potential for practical deployment in real-time decision-making processes related to routing and scheduling. By leveraging both local and indirect relationships through meta-paths, the model succeeded in uncovering nuanced interactions in logistics data—interactions that traditional homogeneous approaches often overlook.

In summary, the findings confirmed that a graph neural network enhanced by multi-criteria attention, feature refinement, and multiple aggregations can effectively address the challenges posed by heterogeneous nodes and edges in logistics optimization. Through this architecture, a robust and flexible tool was introduced for analyzing and predicting hidden patterns in complex transportation networks. By incorporating these advanced representation-learning methods, the model not only achieved superior accuracy and comprehensiveness but also demonstrated heightened resilience and efficiency compared to standard methods. Consequently, it is anticipated that future work focusing on real-time learning, uncertainty management, and large-scale distributed training would further broaden the model's applicability, paving the way for advanced, data-driven decision-making in large, dynamic logistics systems.

- **Limitations and Future Work:**

- **Sparse Adjacency Challenge:** Highly sparse networks may degrade performance; advanced sampling or hierarchical modeling techniques could be explored.
- **Hyperparameter Tuning:** More extensive cross-validation or Bayesian optimization may further improve performance.
- **Online Learning:** Real-time adaptation is needed when dynamic factors (e.g., traffic, weather) change edge weights.
- **Scalability:** Distributed or GPU-accelerated strategies could tackle ultra-large networks more efficiently.

References

[1] Cook, D. J., & Holder, L. B. (2000). Graph-based data mining. *IEEE Intelligent Systems, 15*(2), 32–41. https://doi.org/10.1109/5254.848965

[2] Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of Information Science, 28*(6), 441–453. https://doi.org/10.1177/016555150202800601

[3] Chakrabarti, S. (2002). *Mining the web: Analysis of hypertext and semi-structured data*. Morgan Kaufmann. https://doi.org/10.1016/B978-012457530-7/50007-9

[4] Getoor, L., & Diehl, C. P. (2005). Link mining: A survey. *SIGKDD Explorations, 7*(2), 3–12. https://doi.org/10.1145/1081870.1081872

[5] Han, J. (2009). Mining heterogeneous information networks by exploring the power of links. In *Discovery Science* (pp. 13–30). Springer. https://doi.org/10.1007/978-3-642-00582-52

[6] Lewis, T. (2013). *Network Science: Theory and Applications*. Wiley.

[7] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

[8] Sun, Y., & Han, J. (2013). Mining heterogeneous information networks: A structural analysis approach. *SIGKDD Explorations, 14*(2), 20–28. https://doi.org/10.1145/2491436.2491441

[9] Pham, P., Nguyen, L. T. T., Vo, B., & Yun, U. (2022). Bot2Vec: A general approach of intra-community oriented representation learning for bot detection in different types of social networks. *IEEE Access, 10*, 12162–12174. https://doi.org/10.1109/ACCESS.2022.3148327

[10] Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). DeepWalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD* (pp. 701–710). https://doi.org/10.1145/2623330.2623732

[11] Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD* (pp. 855–864). https://doi.org/10.1145/2939672.2939754

[12] Wang, D., Cui, P., & Zhu, W. (2016). Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD* (pp. 1225–1234). https://doi.org/10.1145/2939672.2939755

[13] Cao, S., Lu, W., & Xu, Q. (2016). Deep neural networks for learning graph representations. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (pp. 1145–1151). https://doi.org/10.1609/aaai.v30i1.1145

[14] Schlichtkrull, M., Kipf, T. N., Bloem, P., van den Berg, R., Titov, I., & Welling, M. (2018). Modeling relational data with graph convolutional networks. In *Proceedings of the 15th European Semantic Web Conference (ESWC)* (pp. 593–607). https://doi.org/10.1007/978-3-319-93846-2_38

[15] Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). Graph attention networks. In *International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.1710.10903

[16] Wang, X., He, H., Cao, Y., Liu, M., & Chua, T.-S. (2019). Heterogeneous graph neural network. In *Proceedings of the 25th ACM SIGKDD* (pp. 793–803). https://doi.org/10.1145/3292500.3330924

[17] Shi, C., et al. (2017). Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering, 29*(2), 242–255. https://doi.org/10.1109/TKDE.2016.2569684

[18] Dong, Y., Chawla, N. V., & Swami, A. (2017). Metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd ACM SIGKDD* (pp. 135–144). https://doi.org/10.1145/3097983.3098036