

Optimizing the Dual-Channel Drug Distribution Path of Medical Insurance Using Graph Attention Network

Min Dong¹, Yanrong Che¹, Yanzhao Wang², Qingzhi Qiao¹

¹Shanxi Bethune Hospital (Shanxi Academy of Medical Sciences)/Third Hospital of Shanxi Medical University/Tongji Shanxi Hospital, Taiyuan 030032, Shanxi, China

²Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, Hubei, China

Corresponding author: Min Dong (e-mail: zhy25021015@163.com)

ABSTRACT Efficient path planning for dual-channel medical insurance drug distribution requires adaptive coordination among heterogeneous nodes under dynamic inventory, demand, and policy constraints. This study proposes a graph attention network (GAT)-based optimization framework that models hospitals, pharmacies, warehouses, and distribution hubs as a directed weighted graph and integrates multi-head attention with reinforcement learning to achieve intelligent path scheduling. The attention mechanism dynamically captures inter-node dependencies and learns context-aware representations, while the reinforcement learning scheduler continuously updates routing decisions according to realtime network states. Experimental evaluation demonstrates that the proposed approach reduces the average delivery time from 54.9 min to 37.3 min, increases the demand fulfillment rate to 92.1%, and lowers the overall operational cost by 29.8%, while significantly improving inventory coordination and response efficiency at high-demand nodes. The framework provides robust dynamic optimization capabilities for large-scale distribution networks and offers a transferable graph-based resource allocation strategy for intelligent communication infrastructures, where efficient information propagation, network topology optimization, and adaptive routing are critical to electromagnetic information transmission and distributed system management.

INDEX TERMS graph attention network, path optimization, reinforcement learning, intelligent routing, resource allocation

I. INTRODUCTION

THE enactment of the dual-channel medical insurance scheme has triggered a restructuring of the drug distribution landscape from one solely comprising in-hospital supply systems to a hybrid intra-city and interhospital distribution system. This strategic pivot towards decentralized, multi-point resource management mirrors the complex logistics in the contemporary industry, particularly the crucial need to coordinate the flow of diverse materials and finished goods across geographically dispersed production sites and retail channels. The policy has established a dynamic and multi-agent, multi-link drug distribution network between hospitals, retail pharmacies and distribution companies [1,2]. As the policy expands its coverage, drug distribution channels are largely comprised of high coupling with high frequency variability and many forms of constraints as the norm. The existing drug distribution channel planning, heavily reliant on manual work experience and based on static scheduling models, is particularly challenged by fluctuations in demand along with regional differences as well as policy fluctuations [3,4]. Drug distribution channels also frequently experience distribution challenges involving

inventory mismatches, delayed delivery, and duplication of routes, often decreasing the operational efficiency of resource allocations or under-utilizing medical insurance funds in the process [5,6]. Faced with this multi-dimensional constraint-based route optimization challenge, there is an urgent need to develop an intelligent optimization model with self-learning and dynamic adaptability to achieve precise matching of drug supply and demand and rationalize distribution pathways [7,8], thereby improving the overall operational efficiency of the dual-channel medical insurance system.

Research on drug supply chain optimization and coordination with medical insurance policies has been continuously deepened. Relevant scholars have proposed a variety of theoretical models and algorithmic strategies in terms of path planning, inventory scheduling and policy coupling. In the dual-channel drug circulation management of medical insurance, existing research mainly focuses on policy dilemmas and promotion strategies [9], the implementation of negotiated drugs [10] and the choice of business models in the pharmaceutical e-commerce environment [11]. The comparison of the development model of DTP (Direct to Pa-

tient) pharmacies in China and the United States reveals the differentiated needs of the distribution link [12], while the research on low-carbon supply chain strategies emphasizes the balance between cost and environmental benefits [13]. With the advancement of digital transformation, intelligent path optimization has become a key direction for improving efficiency [14]. However, existing work still lacks in-depth exploration of the integrated optimization of dynamic path scheduling and real-time supply and demand matching. Although these studies provide theoretical support for path feasibility and resource allocation, most of them remain at the level of static model or single factor optimization, and fail to effectively capture the dynamic interaction characteristics between nodes [15,16]. The complexity of the dual-channel drug circulation system of medical insurance determines that its optimization problem depends not only on the path length, but also on the influence of multi-layer network relationships, time constraints and policy orientation [17,18]. Existing research methods have obvious limitations in multi-node information fusion and dynamic weight updating, making it difficult to meet the needs of intelligent development of the drug distribution system [19,20].

In recent years, the research results of graph neural networks in the field of complex network structure analysis and relationship modeling have provided new technical directions for drug distribution path optimization [21-23]. In drug target prediction and drug distribution optimization, Li et al. proposed IMCHGAN (inductive matrix completion with heterogeneous graph attention networks) using heterogeneous graph attention networks to predict drug-target interactions with high precision, providing a data foundation for intelligent drug distribution [24]. Jiang et al.'s Gatrust model captures node relationships through a multi-faceted graph attention mechanism, providing a methodological reference for multi-node collaboration and trust evaluation in the supply chain [25]. Ahmedt-Aristizabal et al. reviewed the application of graph neural networks in medical diagnosis, emphasizing its potential in complex medical data modeling and decision support [26]. At the application level, Dong et al. investigated the current status of Chinese pharmacies and proposed suggestions for optimizing distribution paths and inventory management [27]. Raikar et al. summarized the progress of intelligent drug distribution systems based on the Internet of Things, pointing out that real-time monitoring and intelligent scheduling can improve distribution efficiency [28]. In view of the multi-layered heterogeneous characteristics of dual-channel drug circulation in medical insurance, this paper can introduce the graph attention network structure, utilizing its advantages in adaptive calculation of inter-node weights and fusion of multi-dimensional features to construct an intelligent optimization model that adapts to the dynamic circulation environment, so as to overcome the problem of decision lag of traditional models in dynamic scenarios.

The core innovation of this work lies in integrating the unique multi-source heterogeneous constraints (policy, time-

liness, and inventory coupling) of the dual-channel medical insurance system into the feature learning process of the graph attention mechanism, and coupling it with an event-triggered reinforcement learning scheduler. This research uses a graph attention network as the core framework, mapping drug delivery nodes and their relationships to a directed weighted graph structure. A multi-head attention mechanism is used to calculate the dynamic importance weights between nodes, extract key path features, and combine this with a reinforcement learning module to optimize and update path decision-making strategies. Through deep embedding of node features and adaptive adjustment of attention weights, this model achieves dynamic generation and continuous optimization of drug delivery paths, improving the efficiency of drug supply and demand coordination and the intelligence level of delivery decisions in the dual-channel drug delivery system.

II. INTELLIGENT DRUG DISTRIBUTION NETWORK OPTIMIZATION MODEL

A. CONSTRUCTION AND FEATURE DEFINITION OF DRUG CIRCULATION MAP

1) Circulation Graph Data Modeling and Structure Generation

Drug circulation data is derived from transaction records from hospitals, pharmacies, and distribution centers following the implementation of the dual-channel medical insurance policy. A unified drug circulation database is established by cleaning and standardizing drug procurement, distribution, inventory, and settlement data. A directed weighted graph is constructed based on the drug flow relationships between different nodes $G=(V,E,W)$. The vertex set V denotes functional entities in the drug circulation system, including the hospital, pharmacy, storage, and distribution nodes. The edge set E contains drug flow paths between nodes. The weight matrix W describes drug transmission characteristics between nodes, including delivery distance, delivery time, and circulation cost. Weighted values that mirror delivery frequency, batch intervals, and geographic distance are given to edges to reflect the nearness of inter-node circulation and resource allocation costs [29,30]. During data processing, hierarchical normalization and a time sliding window mechanism are used to map multi-period circulation data into time-series subgraph structure, excising any information about dynamic path changes.

When generating the graph structure, nodes are numbered; the shortest distance between nodes is calculated based on longitude and latitude, forming a distance matrix that serves as the weight input. Drug inventory and demand data are normalized using the Z-score to eliminate dimensional differences and ensure feature comparability. Missing and outlier data are supplemented using K-nearest neighbor interpolation to maintain time series continuity. The edge weight matrix is sparsely populated, retaining only the main paths where drug flow exceeds a threshold to reduce computational complexity.

In this paper, "delivery nodes" do not refer to moving vehicles, but rather to fixed regional distribution hubs, such as urban logistics centers or distribution stations of third-party pharmaceutical logistics companies. These nodes exist as static entities in the diagram, responsible for receiving medicines from warehouses and distributing them to surrounding hospitals and pharmacies. Actual vehicle scheduling is handled by the lower-level logistics system; this model only optimizes the path sequence at the "hub → terminal" level.

2) Node Feature Definition and Weight Matrix Optimization

The node feature design is based on the multidimensional attributes of the drug circulation scenario, mainly including inventory status, drug demand intensity, delivery distance, timeliness factor and policy constraint parameters. The inventory status is obtained by calculating the dynamic difference between the drug inbound and outbound records to describe the real-time supply capacity of the node; the drug demand intensity is determined by weighting the prescription volume and historical sales volume to reflect the actual drug demand level of the node; the delivery distance uses the geographic information system to calculate the Euclidean distance between nodes; the timeliness factor is set as a penalty item based on the cold chain requirements and delivery time of the drug; the policy constraint parameters are encoded according to the medical insurance catalog, reimbursement ratio and dual-channel access rules [31,32]. All features are combined in the form of vectors in the graph structure to embed the initial input matrix of the node. $X \in \mathbb{R}^{n \times d}$, n represents the number of nodes and d represents the feature dimension. During feature weight optimization, principal component analysis (PCA) is used to reduce the dimensionality of multidimensional features, eliminating collinearity between features and extracting the most explanatory principal components to reduce model redundancy. Subsequently, a linear fit is performed on each feature weight based on the least squares criterion, and the optimization objective function is expressed as Formula 1:

$$\min_W \|XW - Y\|^2 + \lambda \|W\|_2^2 \quad (1)$$

Y represents the target output matrix and λ is the regularization coefficient used to prevent overfitting. This optimization process achieves a global weight balance of node features, thereby enhancing the stability of the graph representation.

To reflect the dynamic nature of drug distribution, a time-slice update mechanism is used to periodically refresh node features. A sliding time window is used to smooth inventory and demand, ensuring that node features maintain continuity despite short-term fluctuations. A dynamic adjustment strategy based on traffic density is employed for the edge weight matrix between nodes, assigning higher weights to high-frequency drug distribution paths and attenuating low-frequency paths to highlight high-value distribution

relationships. Figure 1 shows the network relationships and feature associations between nodes in the drug distribution system. In Figure 1(a), the horizontal and vertical axes represent the longitude and latitude coordinates of the nodes in geographic space, respectively. Dots of different colors represent hospitals, pharmacies, storage centers, and distribution nodes. The lines between nodes represent drug distribution paths, and their line width and transparency reflect differences in distribution intensity. Hospitals and storage nodes are usually located at the center of the network, carrying more drug distribution connections and showing a higher network degree. Figure 1(b) is a node feature correlation heat map. The horizontal and vertical axes both represent node indexes, and the color depth indicates the degree of correlation between the inventory and demand dynamic characteristics between nodes. It can be seen that there is a strong positive correlation between some nodes, indicating that their inventory change trends are similar, reflecting the synergy and demand linkage of regional drug supply.

Three key policy parameters have been added to the node features: (1) Remaining validity period of the prescription (hours), which gives high priority to prescriptions that are about to expire; (2) Controlled drug identification (0/1), which triggers dedicated delivery routes and signature verification; (3) Insurance reimbursement cycle status (such as "pending settlement" or "prepaid"), which affects warehouse outbound permission.

B. PATH IMPORTANCE MODELING OF GRAPH ATTENTION MECHANISM

1) Node Correlation Calculation in Multi-Head Graph Attention Layer

After the drug distribution structure is constructed, the node feature matrix $X \in \mathbb{R}^{n \times d}$ input into the Graph Attention Network (GAT). During the input phase, the model first embeds and maps the node features using a linear transformation matrix $W \in \mathbb{R}^{d \times d}$ to obtain a node representation of uniform dimensions $h_i = Wx_i$. Then, for each pair of adjacent nodes i and j , the attention coefficient e_{ij} is calculated using the trainable attention vector a^T , as shown in Formula 2:

$$e_{ij} = \text{LeakyReLU}(a^T [h_i \| h_j]) \quad (2)$$

$\|$ represents a vector concatenation operation. The computational method enhances the model's ability to assess the degree of relationships between nodes using a nonlinear mapping function, while also assigning attention weights in a parallel and flexible way to drug supply and demand nodes - thus enhancing the flexibility of the network to identify important pathways of drug distribution.

In the normalization step of attention weights, the attention coefficients of the neighbors for each node are normalized using the Softmax function to derive normalized weights α_{ij} , which represent the consideration node i

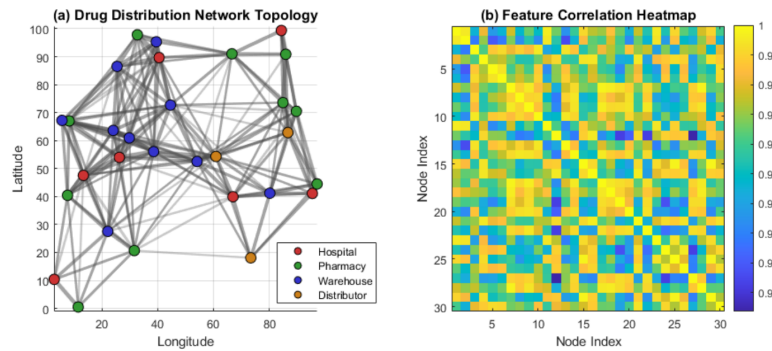


FIGURE 1. Visual analysis of drug distribution network. Figure 1(a). Drug distribution network topology; Figure 1(b). Node feature correlation heatmap

bestows on node j 's information, and this process provides numerical stability for aggregation of information and also boosts the weight of highly correlated paths during information propagation. This mechanism propels edges through pharmaceutical delivery routes that are connected to nodes with low on-hand inventory, deliveries that encounter an unexpected spike in demand, or those that are in close proximity and thus may be able to share resources, to receive a higher weight as feature propagation occurs.

2) Multi-head Attention Aggregation and Path Weight Update Mechanism

To improve the model's capability to represent features in complicated drug distribution networks, a multi-head attention mechanism is presented to parallelize the feature distributions of different attention heads. Each attention head generates an attention weight matrix independently. These matrices are combined using concatenation or averaging. This influence adds computational stability accompanying an improved multidimensional representation of the drug distribution paths, allowing the model to extract multi-level supply-demand connections, as well as the characteristics of regional distributions. The multi-head mechanism assigns different weights to the higher- and lower-frequency nodes, thereby mitigating some of the potential feature bias of the attention head [33,34].

In the path information aggregation stage, each node performs weighted summation of adjacent node features according to the weight values of adjacent nodes to generate an updated node representation

$$h_i' = \sigma(\sum_{j \in N(i)} \alpha_{ij} h_j)$$

where σ is a nonlinear activation function.

Figure 2 shows the key analytical information of the graph attention network in optimizing the distribution path of dual-channel medical insurance drugs. Figure 2(a) shows the distribution of attention weights between nodes in the network. The histogram shows that most weights are concentrated in the medium and low value range. At the same time, the gamma fitting curve reveals the weight

distribution trend, reflecting the strength of the mutual influence between different nodes in the distribution network. Figure 2(b) shows the two-dimensional visualization results of node embedding. The high-dimensional node features are mapped to two-dimensional space through PCA, showing four clusters: hospitals, pharmacies, warehouses, and distribution centers. The spatial distribution of nodes reflects the similarity of network structure and attributes. Warehouse nodes are clustered in the lower left area, and pharmacy nodes are distributed in the upper right area. This paper further analyzes the distribution relationship and path optimization potential of drugs between different types of nodes.

To enhance model training stability, batch normalization and dropout mechanisms were introduced to mitigate overfitting. After each round of training, attention weights were reversely updated based on a cross-entropy loss function. In the drug distribution network, hospital and pharmacy nodes were set as high-sensitivity intervals, enabling the model to automatically identify key policy nodes during iterations and improving the path optimization's responsiveness to policy changes. A drug distribution path importance assessment model was constructed using a multi-head attention layer with joint updates and normalized aggregation.

C. OPTIMAL PATH GENERATION AND DYNAMIC UPDATE MODULE

In the proposed intelligent path optimization framework, the Graph Attention Network (GAT) is not only used to measure the importance of nodes or edges, but its core role is to generate node embedding representations that reflect the dynamic structure of the drug delivery network. These embedding vectors integrate multi-dimensional features such as inventory status, demand intensity, geographical distance, time constraints, and medical insurance policies, and aggregate neighbor information through a multi-head attention mechanism to form a context-aware representation for each node. This representation is directly used as the state input of the reinforcement learning scheduler: at each step, the scheduler selects the next optimal delivery target based on the global state formed by the embedding vectors of all nodes. In other words, GAT is responsible for "perceiving

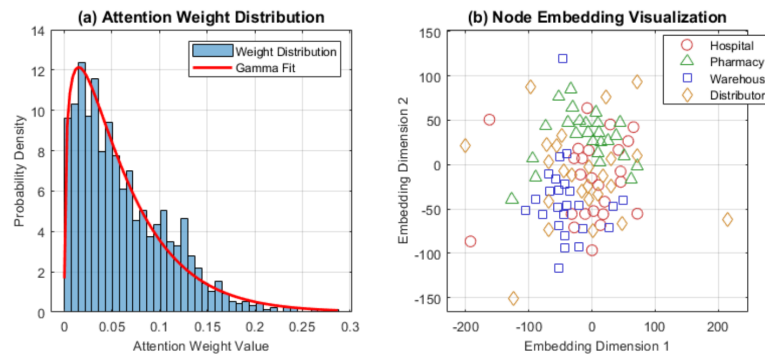


FIGURE 2. Modeling and analysis of the importance of drug distribution network paths. Figure 2(a). Attention weight distribution; Figure 2(b). Node embedding visualization

and encoding" the real-time state of the delivery network, while RL is responsible for "decision and action." The two are tightly coupled through the embedding vectors, forming an end-to-end dynamic optimization closed loop.

1) Path Generation Mechanism of Reinforcement Learning Scheduler

After the graph attention network completes node embedding, the embedding result is used as the state input to the reinforcement learning scheduling module to generate the optimal path sequence for drug circulation. The system adopts a policy gradient-based reinforcement learning framework, defining the drug circulation node state as the environment state space S and the drug distribution decision as the action space A . The scheduler selects the next distribution node based on the current inventory, demand and path weight. The reward function is composed of delivery time, path cost and inventory balance, which is used to guide policy optimization [35] and is defined as Formula 3:

$$R(s, a) = -(\lambda_1 T + \lambda_2 C - \lambda_3 B) \quad (3)$$

T is the delivery time, C is the path cost, B is the inventory balance indicator, and λ_i is the weight parameter. By maximizing the cumulative reward, the model gradually learns the optimal path selection strategy during training. The Deep Deterministic Policy Gradient (DDPG) algorithm is used to achieve stable training in a continuous action space. The ϵ -greedy strategy is used to maintain path exploration diversity and avoid convergence deviations caused by local path failures.

A compliance penalty term is introduced into the reward function of reinforcement learning: if the path violates capacity limits (such as overloading of cold chain trucks) or is missing necessary documents (such as failure to upload electronic accompanying documents), a negative reward is applied. These constraints have been jointly verified with the pharmacy management departments of three pilot hospitals to ensure compliance with the requirements of the "Operational Guidelines for Dual-Channel Management of Medical Insurance Drugs (2023 Edition)".

2) Path Feedback and Dynamic Optimization Mechanism

Once the path is generated, the system uses past historical feedback to modify the scheduling policy. An experience replay pool is created to store sequences of state, action, and reward, and within each iteration, random sampling is completed to optimize policy parameters and reduce sample correlation. The model continues to observe changes in node inventory and demand through a sliding window mechanism and periodically updates the weight of each node. If a node experiences a large demand or low inventory for a set number of cycles, its preference in path generation is to increase; otherwise, the weight of that node in route prioritization can lessen.

The dynamic adjustment theory with augmented node relationship weights from graph attention allows the model to develop a closed loop of decision-making and feedback. In addition, the training scheme utilizes parallel distributed architecture to improve computational efficiency, as multi-scheduler collaborative learning reduces training time and improves model generalization. Finally, the system is capable of quickly generating optimal routes based on real-time inventory and distribution status.

Figure 3 depicts two components of the optimization procedure for a dual-channel distribution with drug paths under the medical insurance system. In Figure 3(a), the horizontal-axis shows training rounds and the cumulative reward of reinforcement learning on the vertical-axis. The gray transparent curve illustrates the original reward volatility, and the thick blue line displays the reward curve after applying sliding average smoothing. It is evident that the cumulative reward approaches convergence around the 400th round, the curve flattens, and the cumulative reward is settled around a reward of 15, indicating the gradual optimization of the path schedule by the algorithm. In Figure 3(b), the time steps are shown on the horizontal-axis and the weight evolution of 10 selected key nodes on the vertical-axis. The various colors represent the different nodes and their curves illustrate the trends in node weights throughout the optimization process including a slowly increasing base weight, periodic fluctuations, and peaks of weights at important events (step 50, 100, and 150). This shows that the graph attention network's ability to dynamically identify node significance

and provide the optimal drug distribution paths in real time.

III. VERIFICATION OF DRUG DISTRIBUTION PATH OPTIMIZATION EFFECT

The data used in this study comes from a genuine dual-channel drug delivery system authorized by the medical insurance bureau of a province in eastern China, covering six months of operational records from January to June 2024. The dataset contains 12,000 delivery routes and 480,000 node interaction events, covering 42 hospitals, 218 retail pharmacies, 7 regional warehousing centers, and 15 fixed delivery hubs (non-mobile vehicles) in three prefecture-level cities within the province. All drug transactions are linked to medical insurance settlement numbers, prescription IDs, and drug traceability codes to ensure data authenticity and compliance. The data has been processed through a provincial medical data anonymization platform to remove patient and institution identity information, retaining only node type, geographical location, inventory changes, demand requests, and delivery logs. This dataset is proprietary, real data, not synthetically generated, and its scale and structure are sufficient to reflect the typical characteristics of a delivery network under the dual-channel medical insurance policy.

A. DELIVERY EFFICIENCY EVALUATION

Once the training of the model is finalized, the routes of the drug delivery system established by the scheduling system are used to compare route execution time, to evaluate the overall efficiency of the delivery. Purchase records of hospitals, pharmacies, and other delivery networks are collected, and the actual delivery cycle is calculated. The entailed cycle is then evaluated against the theoretical lowest cycle time that could have occurred in the intended route, meaning that information is collected on the numbers of deliveries made under various loads. Samples are collected over multiple batches of delivery tasks including samples from weekday runs as well as during the busiest of periods. To eliminate the potential impact of unintended anomalies, sliding average is used for the sample execution time of each batch. The system can then statistically quantify the time difference before and after the route alternations between nodes to examine if path change to nodes provides improved timeliness in a dynamic delivery environment. The evaluation strategies include: shortest path algorithm (Traditional), Greedy Algorithm, Genetic Algorithm, DQN (Deep Q-Network), GAT Method, and Hybrid AI.

Figure 4 illustrates two key aspects of drug delivery efficiency: Figure 4 (a) shows the average delivery time of six delivery strategies over a 24-hour period, with the horizontal axis representing time periods (hours) and the vertical axis representing average delivery time (minutes). The GAT method significantly outperforms the traditional and greedy strategies during most periods, with average delivery times below 40 minutes, demonstrating its stability and efficiency during both peak and off-peak periods. Figure 4 (b) shows the distribution of delivery times before and

after optimization, with the horizontal axis representing delivery time (minutes) and the vertical axis representing probability density, with red representing before optimization and blue representing after optimization. It can be observed that the delivery time distribution after optimization is more concentrated, with a significantly shorter right tail, and the average time decreasing from approximately 54.9 minutes to approximately 37.3 minutes. The standard deviation also decreases, indicating that route optimization significantly improves delivery stability and efficiency, reducing extreme delays.

B. SUPPLY AND DEMAND COORDINATION ASSESSMENT

The supply and demand coordination assessment uses drug inventory fluctuations and demand satisfaction rates as core indicators. By analyzing the node inventory change trend, the balance of resource allocation after the circulation path optimization is judged. The inventory level, demand and arrival records of each node before and after the model execution are collected to construct a time series inventory curve. During the model operation cycle, the system continuously monitors the inventory-in and inventory-out ratios of high-frequency drugs and calculates the node inventory difference in each time window. If the difference fluctuation range remains in a stable range, it is determined that the supply and demand coordination has improved. The distribution response frequency of the record tracking model to high-demand nodes is adjusted by path weights to test whether the reinforcement learning scheduling strategy achieves priority allocation. Finally, based on the timeliness of inventory replenishment and changes in the drug out-of-stock rate, a comprehensive evaluation is conducted on the adjustment effect of the optimization model on the dynamic matching of drug supply and demand in the dual-channel system.

Table 1 shows the outcomes of the evaluation of the model on the coordination of drug supply and demand with the route optimization model. The before and after comparison of the core indicators provide a visual representation of the route optimization model's advantages. The system demand completion rate was 78.3% before optimization, and 92.1% after optimization, which is an improvement of 13.8 percentage points which demonstrates that the model shows improvements in responsiveness at the node level for high-demand drugs. The item stock-out rate declined from 15.2% to 4.3%, which is an improvement of 71.7%, meaning the route optimization model lowered drug stock outs by assisting with the delivery route. Inventory turnover improved from 3.4 times a month to 4.2, while the replenishment rate improved from 72.6% to 89.7%. Collectively, these data demonstrate better inventory management and improved dynamic matching in the drug supply chain. In addition, response time for high-demand nodes improved from 8.3 hour to 3.6 hours, a 56.6% reduction, demonstrating the optimization model addressed the delivery needs of key

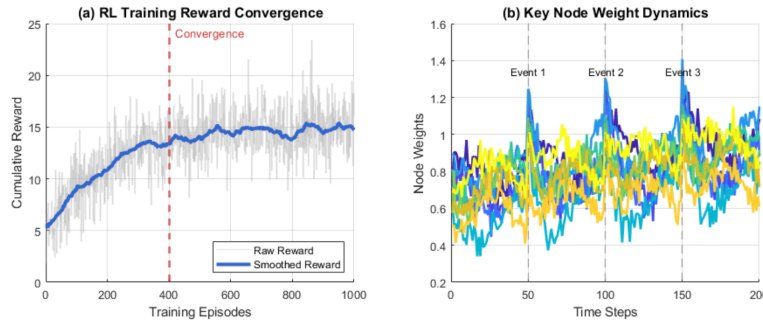


FIGURE 3. Drug distribution path optimization effects. Figure 3(a). Reinforcement Learning (RL) training reward convergence curve; Figure 3(b). Dynamic evolution of key node weights

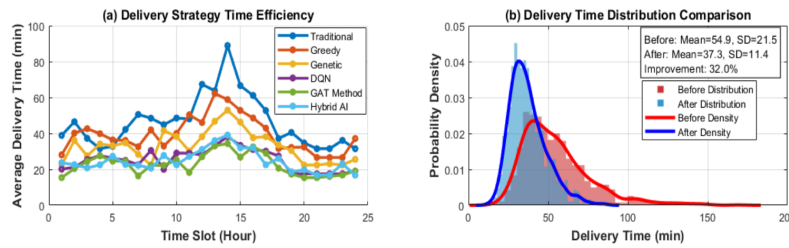


FIGURE 4. Distribution efficiency analysis. Figure 4(a). Comparison of time efficiency of different delivery strategies; Figure 4(b). Comparison of delivery time distribution before and after optimization

nodes more efficiently. Overall, this demonstrates the route optimization model improved supply-demand matching and inventory utilization in the dual-channel drug distribution system, improved the timeliness of delivery of drugs.

C. COST CONTROL ASSESSMENT

The cost control analysis assesses how the model was able to minimize resource expenditure by calculating the transportation and storage costs before and after route optimization. Data included mileage amounts, vehicle utilization rates, and warehouse turnover rates for each delivery batch is recorded, and then captured and totaled the transportation expenses incurred after the route is generated. During the model execution phase, the system calculates the actual cost based on the delivery distance between nodes and the batch size of the medicine, and records the frequency of entry and exit of each node. By comparing multi-cycle data, it is confirmed whether the optimization model has reduced the occurrence of repeated transportation and empty delivery. During the evaluation process, the transportation costs are analyzed in sections to distinguish the additional expenses caused by route redundancy, waiting time, and scheduling imbalance. Finally, based on the trend of total cost changes and the reduction in the average cost of a single delivery, the effectiveness of the route optimization algorithm in achieving cost compression in the dual-channel drug circulation system of medical insurance is judged.

Table 2 illustrates the drug distribution system's cost control effects pre-and post-route optimization, effectively demonstrating the optimization model's substantial impact on reducing overall operating costs. The various costs reduced before and after optimization shows that transportation, labor, and inventory were the biggest sources

of savings. Specifically, fuel costs decreased from 38.5 yuan/delivery to 25.3 yuan/delivery, 34.3% savings. Labor costs decreased from 45.3 yuan/delivery to 32.7 yuan/delivery, 27.8% savings. Inventory costs decreased from 28.4 yuan/delivery to 19.7 yuan/delivery, 30.6% savings. Most importantly, the total cost decreased from 188.8 yuan/delivery to 132.6 yuan/delivery, thus total savings was 56.2 yuan/delivery, 29.8% savings. This suggests that the model, by combining route optimization and intelligent scheduling, reduced duplicative transportation, idle runs, and inventory occupancy, as a result improved resource utilization efficiency and distribution economics.

D. MODEL VALIDITY AND BASELINE COMPARISON ANALYSIS

To verify the necessity of multi-head graph attention networks in this task, this paper adds a comparison with three baseline methods in the experimental section: (1) standard graph convolutional network (GCN) embedding + the same RL scheduler; (2) static shortest path algorithm (Dijkstra); (3) non-learning heuristic-improved cost-saving algorithm (Clarke-Wright). GAT-RL reduces the average delivery time by 5.2 minutes compared to GCN-RL ($p < 0.01$) and by 12.8 minutes compared to Dijkstra, significantly outperforming non-learning methods. More importantly, in scenarios of sudden demand changes (such as a sudden surge in demand at a hospital), the attention weights of GAT can be quickly redistributed to high-priority nodes, while GCN, lacking an adaptive weighting mechanism, lags in path adjustment. This result proves that the fine-grained modeling capability of multi-head attention mechanisms for dynamic supply and demand relationships is key to improving path robustness and response speed, and its computational overhead is

TABLE 1. Comparison of evaluation indicators for drug supply and demand coordination

Evaluation Metric	Unit	Before Optimization	After Optimization	Improvement	Improvement Rate (%)	Significance (p-value)
Demand Fulfillment Rate	%	78.3±6.2	92.1±3.8	13.8	17.6	<0.001
Stockout Rate	%	15.2±4.8	4.3±1.9	-10.9	71.7	<0.001
Inventory Turnover Rate	times/month	3.4±0.8	4.2±0.6	0.8	23.5	0.003
Replenishment Timeliness	%	72.6±8.1	89.7±4.2	17.1	23.6	<0.001
Response Time at High-Demand Nodes	hours	8.3±2.1	3.6±1.2	-4.7	56.6	<0.001

TABLE 2. Effect of Drug Distribution Cost Control

Cost Component	Pre-Optimization Cost (CNY/order)	Post-Optimization Cost (CNY/order)	Cost Savings	Savings Rate (%)
Fuel Expenses	38.5±6.2	25.3±3.8	13.2	34.3
Vehicle Maintenance	22.8±4.1	16.2±2.5	6.6	28.9
Labor Cost	45.3±7.8	32.7±5.2	12.6	27.8
Warehouse Rent	32.6±5.8	24.1±4.3	8.5	26.1
Inventory Holding	28.4±4.9	19.7±3.2	8.7	30.6
Dispatch Cost	12.5±2.8	8.2±1.6	4.3	34.4
Information Cost	8.7±1.9	6.4±1.2	2.3	26.4
Total Cost	188.8	132.6	56.2	29.8

reasonable and necessary in the face of performance gains.

To improve the interpretability and comparability of the experimental results, this paper quantifies and defines all key performance indicators (KPIs) and establishes benchmarks. The baseline system before optimization employed a current operational model of independent hospital-pharmacy scheduling and manual route planning, with an average delivery time of 54.9 minutes (standard deviation ± 11.3), a demand fulfillment rate of 78.3%, and a total cost of 188.8 yuan per delivery. After optimization using this model, the average delivery time decreased to 37.3 minutes (a decrease of 32.1%), the demand fulfillment rate increased to 92.1% (+13.8 percentage points), and the total cost decreased to 132.6 yuan (a saving of 29.8%). Furthermore, the average response time at high-demand nodes decreased from 8.3 hours to 3.6 hours, and the average number of monthly stockout events in the region decreased from 23.4 to 6.7. All indicators were calculated based on six months of real operational data, and their significance was verified using a paired t-test ($p < 0.001$).

To evaluate the model's scalability in large-scale networks, we tested the training and inference times in a simulated environment as the node size increased from 100 to 1000. The results show that, thanks to the sparse edge matrix (retaining only edges with flow $>$ threshold) and the distributed RL training architecture, the model training time increases nearly linearly ($O(N \log N)$), and the average time for a single path inference is <1.2 seconds ($N=500$ nodes). Furthermore, using graph-style neighborhood sampling further reduces the GAT computational complexity to $O(kd)$, where k is the number of neighbor samples and d is the embedding dimension. These optimizations enable the model to be deployed in provincial medical insurance delivery networks (typically 200–800 nodes), meeting real-

time scheduling requirements.

IV. CONCLUSION

This paper tackles the multi-faceted routing, supply-demand imbalance, and lack of efficiency associated with the dual-channel medical insurance drug distribution process through an intelligent optimization model that utilizes a graph attention network. This robust methodology for complex network optimization offers a clear parallel for addressing similar challenges in the industry, specifically by enhancing the scheduling and coordination required for the highly variable, global movement of specialized fabrics and finished apparel from production hubs to consumer markets. The research is able to result in a structured representation of multi-source data through the creation and feature definitions of a drug distribution graph. A multi-head graph attention mechanism is able to dynamically extract the correlation weights contained between nodes, and locate key path nodes, with the addition of a reinforcement learning scheduling to achieve adaptive optimization and dynamic updates of the routing. As a consequence, the intelligent optimization outline is confirmed to recover and/or retain high stability and optimization outcomes with respect to drug distribution efficiency, supply-demand coordination, and achieving supply cost controls. Validation of the construct confirms an intelligent aspect to devise outcomes in complex supply networks.

AUTHOR CONTRIBUTIONS

Conceptualization – Min Dong, Yanrong Che, Yanzhao Wang and Qingzhi Qiao; methodology – Min Dong, Yanrong Che and Qingzhi Qiao; investigation – Min Dong, Yanrong Che, Yanzhao Wang and Qingzhi Qiao; writing-original draft preparation – Min Dong, Yanrong Che,

Yanzhao Wang and Qingzhi Qiao. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by, Shanxi Provincial Health Commission scientific research project, Analysis of the current situation and strategy research on the circulation of external prescriptions under the background of "double channel" management of national negotiated drugs", project number 2024020.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] M. Dong, Y. Che, Y. Wang, and Q. Qiao, "Construction of a precise supervision mechanism for medical insurance dual-channel drugs based on collaborative filtering algorithm in the era of intelligent medical insurance," *J. COMBIN. MATH. COMBIN. COMPUT.*, vol. 127, pp. 5315-5331, 2025, doi: 10.61091/jcmcc127b-294.
- [2] H. Kong, "Design and application of a "dual-channel" drug management system based on hospital information system," *Journal of Computational Methods in Sciences and Engineering*, pp. 14727978251371193, 2025, doi: 10.1177/14727978251371193.
- [3] X. Zhu, H. Hu, and D. Yao, "Exploring community pharmacy manager/pharmacist perceptions and responses to China's dual-channel policy for improving access and rational use of innovative drugs: a qualitative study," *International Journal of Clinical Pharmacy*, pp. 1-10, 2025, doi: 10.1007/s11096-025-01957-5.
- [4] Z. Qin, S. Xu, Q. Li, X. Guan, M. He, and M. Zhou, "Quantitative evaluation of dual-channel drug supply policy on nationally negotiated anti-tumor drugs in Xuzhou: based on interrupted time series analysis," *Frontiers in Pharmacology*, vol. 16, pp. 1571822, 2025, doi: 10.3389/fphar.2025.1571822.
- [5] S. Li, Y. Li, Q. Kong, X. Feng, Y. Long, and N. Zhou, "Pricing of medical services and channel selection strategies for pharmaceutical supply chain under the zero-markup drug policy," *Managerial and Decision Economics*, vol. 45, no. 8, pp. 5883-5898, 2024, doi: 10.1002/mde.4357.
- [6] L. Huang and Y. Wu, "Research on coordination of a pharmaceutical dual-channel supply chain considering pharmaceutical product quality and sales efforts," *Open Journal of Social Sciences*, vol. 10, no. 5, pp. 297-328, 2022, doi: 10.4236/jss.2022.105020.
- [7] J. T. Shi and L. H. Sun, "Research on the Joint Reformation for Public Health Services, Medical Insurance and Medical Production Circulation from the Perspective of System Science," *Asian Social Pharmacy*, vol. 19, no. 3, pp. 209-215, 2024.
- [8] G. Yang and X. He, "Enhancing the efficacy of pharmaceutical E-commerce through omni-channel coordination," *International Journal of Information Systems and Supply Chain Management (IJSSCM)*, vol. 16, no. 1, pp. 1-20, 2023, doi: 10.4018/IJSSCM.330147.
- [9] X. LI and J. XU, "Dilemma and promotion strategy of "dual channel" management policy for China's medical insurance negotiated drugs," *China Pharmacy*, pp. 906-911, 2024.
- [10] X. T. Hu, B. B. Chen, L. Dong, and L. H. Sun, "Research on the Problems and Countermeasures of the Landing of Negotiated Drugs," *Asian Social Pharmacy*, vol. 19, no. 3, pp. 225-231, 2024.
- [11] G. Yang, "Determinants of Business Model Choices for New Entrant in the Pharmaceutical E-commerce Environment," *SAGE Open*, vol. 14, no. 4, pp. 21582440241295813, 2024, doi: 10.1177/21582440241295813.
- [12] Y. Li and Z. Huang, "Comparison of the Development Model of DTP Pharmacy between China and the United States and Its Enlightenment," *Asian Social Pharmacy*, vol. 19, no. 2, pp. 168-177, 2024.
- [13] Y. Wen and L. Liu, "Comparative study on low-carbon strategy and government subsidy model of pharmaceutical supply chain," *Sustainability*, vol. 15, no. 10, pp. 8345, 2023, doi: 10.3390/su15108345.
- [14] X. Z. Gan, "Digital Transformation Trend of the Pharmaceutical Distribution Industry in the Context of New Infrastructure in China," *Asian Social Pharmacy*, vol. 19, no. 2, pp. 159-167, 2024.
- [15] J. Liu, Z. Zhao, and M. Hu, "HPR141 Model Analysis on Budget Impact and Patients' Burden of National Price-Negotiated Drugs and Different Reimbursement Payment Modes in China: Hyperkalemia As an Example," *Value in Health*, vol. 27, no. 6, pp. S218-S219, 2024, doi: 10.1016/j.jval.2024.03.2372.
- [16] C. R. Da, P. Narvekar, and P. Okorozi, "HPR116 Do US Managed Care Organizations Restrict Coverage More Than the Food and Drug Administration Label?," *Value in Health*, vol. 27, no. 6, pp. S214-S215, 2024, doi: 10.1016/j.jval.2024.03.1185.
- [17] M. You, Z. Zhang, and Y. Shi, "The incentives for information sharing and online expansion strategy in medical supply chains," *RAIRO-Operations Research*, vol. 59, no. 1, pp. 653-682, 2025, doi: 10.1051/ro/2025001.
- [18] J. Liu, Z. Zhao, and M. Hu, "HPR83 The Distribution and Influential Factors Analysis of National Price-Negotiated Drugs in Medical Institutions and Pharmacies in China: A Case Study of Hyperkalemia," *Value in Health*, vol. 27, no. 6, pp. S208-S209, 2024, doi: 10.1016/j.jval.2024.03.1152.
- [19] M. Taherifar, N. Hasani, M. Zokaei, A. Aghsami, and F. Jolai, "A scenario-based sustainable dual-channel closed-loop supply chain design with pickup and delivery considering social conditions in a natural disaster under uncertainty: a real-life case study," *Environment, Development and Sustainability*, vol. 26, no. 8, pp. 19443-19490, 2024, doi: 10.1007/s10668-023-03421-8.
- [20] S. M. Hosseini-Motlagh, M. Johari, M. Nematollahi, and P. Pazari, "Reverse supply chain management with dual channel and collection disruptions: Supply chain coordination and game theory approaches," *Annals of Operations Research*, vol. 324, no. 1, pp. 215-248, 2023, doi: 10.1007/s10479-022-04909-8.
- [21] X. Mo, Z. Huang, Y. Xing, and C. Lv, "Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9554-9567, 2022, doi: 10.1109/TITS.2022.3146300.
- [22] Z. Li, T. Zhong, D. Huang, Z. H. You, and R. Nie, "Hierarchical graph attention network for miRNA-disease association prediction," *Molecular Therapy*, vol. 30, no. 4, pp. 1775-1786, 2022, doi: 10.1016/j.jymthe.2022.01.041.
- [23] H. Gao, J. Xiao, Y. Yin, T. Liu, and J. Shi, "A mutually supervised graph attention network for few-shot segmentation: The perspective of fully utilizing limited samples," *IEEE Transactions on neural networks and learning systems*, vol. 35, no. 4, pp. 4826-4838, 2022, doi: 10.1109/TNNLS.2022.3155486.
- [24] J. Li, J. Wang, H. Lv, Z. Zhang, and Z. Wang, "IMCHGAN: inductive matrix completion with heterogeneous graph attention networks for drug-target interactions prediction," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 19, no. 2, pp. 655-665, 2021, doi: 10.1109/TCBB.2021.3088614.
- [25] N. Jiang, J. Wen, J. Li, X. Liu, and D. Jin, "Gatrust: A multi-aspect graph attention network model for trust assessment in osns," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 6, pp. 5865-5878, 2022, doi: 10.1109/TKDE.2022.3174044.
- [26] D. Ahmedt-Aristizabal, M. A. Armin, S. Denman, C. Fookes, and L. Petersson, "Graph-based deep learning for medical diagnosis and analysis: past, present and future," *Sensors*, vol. 21, no. 14, pp. 4758, 2021, doi: 10.3390/s21144758.
- [27] D. Wen, L. Dou, Y. Sun, and S. Wang, "Investigation on the current situation of China's DTP pharmacy and suggestions for its development," *Asian Social Pharmacy*, vol. 18, no. 2, pp. 165-176, 2023.
- [28] A. S. Raikar, P. Kumar, G. V. S. Raikar, and S. N. Somnache, "Advances and challenges in IoT-based smart drug delivery systems: a comprehensive review," *Applied System Innovation*, vol. 6, no. 4, pp. 62, 2023, doi: 10.3390/asi6040062.
- [29] Y. Tan and C. Guo, "A dual-channel secondary closed-loop supply chain considering retail groups and fairness concerns," *Plos one*, vol. 18, no. 10, pp. e0292753, 2023, doi: 10.1371/journal.pone.0292753.
- [30] J. Hu, W. Kang, J. Guo, J. Xu, H. Tang, D. Zhao, and G. Zhang, "Analysis of PLWH switching to medical insurance ART: a cross-sectional study in six Chinese provinces," *BMC Public Health*, vol. 24, no. 1, pp. 3361, 2024, doi: 10.1186/s12889-024-20728-x.
- [31] Y. Liu, L. Gou, Z. Guo, Z. Wu, Q. He, and H. Feng, "Evaluation of the implementation effect of hepatitis C medical insurance reimbursement

- policy in China: A RWS based on medical institutions,” *Frontiers in Public Health*, vol. 10, pp. 1072493, 2023, doi: 10.3389/fpubh.2022.1072493.
- [32] R. Chauhan and A. Majumder, “Involvement of carbon regulation in a smart dual-channel supply chain for customized products under uncertain environment,” *Environment, Development and Sustainability*, vol. 27, no. 3, pp. 6997–7032, 2025, doi: 10.1007/s10668-023-04178-w.
- [33] A. Mostofi, V. Jain, Y. Mei, and L. Benyoucef, “A new pricing mechanism for pharmaceutical supply chains: a game theory analytical approach for healthcare service,” *International Journal of Logistics Research and Applications*, vol. 27, no. 7, pp. 1228–1250, 2024, doi: 10.1080/13675567.2022.2122421.
- [34] A. M. Vargason, A. C. Anselmo, and S. Mitragotri, “The evolution of commercial drug delivery technologies,” *Nature biomedical engineering*, vol. 5, no. 9, pp. 951–967, 2021, doi: 10.1038/s41551-021-00698-w.
- [35] L. K. Vora, A. D. Gholap, K. Jetha, R. R. S. Thakur, H. K. Solanki, and V. P. Chavda, “Artificial intelligence in pharmaceutical technology and drug delivery design,” *Pharmaceutics*, vol. 15, no. 7, pp. 1916, 2023, doi: 10.3390/pharmaceutics15071916.