Hindawi
Complexity
Volume 2023, Article ID 8841451, 13 pages
https://doi.org/10.1155/2023/8841451

Research Article

Emergency Medical Resources Allocation of Periphery for Epidemic Areas: Based on Infectious Diseases Spatial-Temporal Transmission Path

Xuan Zhao, Benhong Peng, Chaoyu Zheng, and Anxia Wan

School of Management Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence should be addressed to Xuan Zhao; xuanzhao94@163.com and Benhong Peng; 002426@nuist.edu.cn

Received 25 March 2023; Revised 28 June 2023; Accepted 11 July 2023; Published 4 September 2023

Academic Editor: Chittaranjan Hens

People in the epicenter suffer from emergency medical supplies shortage in the early stage of a public health emergency because of imbalanced supply-demand in different regions or areas, which is a key issue in a major infectious disease. In response to the severe insufficiency of supplies in the epicenter, this study proposed a strategy of distributing supplies from peripheral areas to the epicenter and gave a supply-side selection model considering the epidemic influence and supplies condition in the candidate supply-side areas. First of all, the epidemic spatial-temporal transmission path (STTP) network describing the geographic spread of disease is obtained using a first-order conditional dependence approximation algorithm in a dynamic Bayesian network (DBN). Then, the structural information of the STTP network and the supplies condition characteristic information are combined using the Bipartite network embedding (BiNE) method. Finally, a graph convolutional neural network (GCN) is conducted to select the supply-side areas for peripheral-epicenter supplies distribution based on information achieved from the bipartite graph. The results show that the highest supplies allocation accuracy reaches 87%. Validation and supremacy of the proposed methodology are provided by applying it to the case in Hubei province. This study considers crossed-areas supplies distribution strategy and contributes to select suitable supply-side areas considering the epidemic and supplies condition in the peripheral areas, which is helpful to both epicenter and peripheral areas.

1. Introduction

Emergency medical supplies provide treatment and care for patients and protection for medical staff and the general public in large, resurgent epidemics. However, supplies are not evenly distributed in each region because of different demographics and structural functions. Some regions have fewer supplies for emergencies, and there is a possibility that the regions happen to be the epicenter. The imbalanced distribution and demand of emergency medical supplies may cause serious resource shortages, speeding up outbreak spread and increasing the severity of the pandemic [1]. Take Wuhan, China, during the COVID-19 outbreak for example, it suffered from a severe shortage of emergency medical supplies, and the production of supplies could not meet the supply in the short term. In addition, some certain emergency supplies have a short shelf life and have limited usage scenarios [2]. For example, the shelf life of medical protective clothing is about three years [3], and it is used less in normal times than the abnormal. If these kinds of supplies could not be used timely, they will expire and will lead to a lot of waste. At this point, it is particularly important to dispatch supplies to the worst-hit areas, which has many advantages including time-saving and higher resources utilization. For instance, dispatching supplies from the areas that are not affected by the disasters to the disaster-stricken areas allows the victims to achieve supplies earlier than waiting for the production. The sooner they get the supplies, the earlier the victims can survive the epidemic. More importantly, as the contagion characteristics of epidemic, the medical supplies can help slow the spread of the disease expanding to other areas, which is a win-win for both epicenter and other areas. Last but not least, sharing supplies is helpful to take advantage of resources and avoid resource waste.

The emergency supplies allocation methods can be categorized into two classes: model conduction and data-driven. The model-driven method develops the strategy and establishes the model of supplies allocation for specific scenarios and configuration goals [4]. Wright and Lim [5] proposed an interior point method to solve the resource allocation problem of nested constraints. A stochastic optimization model for allocating and sharing critical resources in the case of large epidemics is also proposed [6]. Different from model-building methods, data-driven methods achieve configuration objectives through regression, such as statistical approaches or machine learning algorithms, based on data. Data-driven methods are more effective for complex situations where mathematical models are difficult to build. Some scholars used sociodemographic data [7], socioeconomic factors [8], and land-use intensity [9] to utilize material supply allocation in specific regions. Based on Bayesian decision analysis, Wohlmann et al. [10] constructed a decision model of emergency supplies allocation for the major infectious diseases under five scenarios against seven dimensions, including time period, critical events, transmission dynamics, spatial distribution, infection scale, information characteristics, and medical supplies.

It is worth noting that the model construction method is easy to be practiced and understood. However, methods of this kind require predefined rules and the reasons behind abnormal results remain unexplained. Moreover, because there are certainly complexities related to the massive amount of real-world datasets, data-driven methods with high statistical order or too many nonlinearities might be not available. Besides, these above emergency supplies allocation studies, however, have mostly concentrated on how emergency medical resources are distributed within regions, neglecting the significance of the coordinated multiregional distribution of emergency resources in crossed-regional public emergencies.

The significance of multiregional collaborative materials allocation planning has steadily caught the attention of academics, and it is originated from the integrated emergency management decision support system proposed by Pérez-Rodríguez and Holguín-Veras [11] to deal with the hurricane disaster. Although the definition of the concept has not been unified, the understanding has gradually reached a consensus; that is, urban agglomerations are composed of several neighboring cities, which are closely connected with each other. For instance, Boin et al. [12] and Mehrtash et al. [13] discussed the importance of cross-regional coordination mechanisms through cost-benefit analysis under sudden cross-regional disasters. Pescaroli et al. [14] and Qiu et al. [15] put forward suggestions on how to construct emergency rescue synergistic processes. Urrutia et al. [16] used generalized network autoregressive to interdict potential higher infection areas. Similar to the general emergency materiel allocation, equity [17], risk acceptance and survivor satisfaction [18], minimal distribution cost, and delivery time [19] are all optimization objectives for multiregional collaborative emergency materials allocation. However, it is noteworthy that the majority of these researches concentrated on the victims' pleasure in the area of supply-demand, forgetting to consider the victims' contentment in the supply areas and rarely involve how to coordinate resource allocation among emergency response entities with differentiated characteristics based on regional coefficients when considering their decision preferences, which is exactly one key problem to be solved.

To improve multiregional collaborative emergency materials allocation, both peri-epidemic and epicenter areas needed to predict their respective emergency material demands. Defining the pandemic's spatial and temporal patterns, or the crucial issue of how the epidemic spreads in the immediate region, is one of the main elements in creating connections between the epidemic perimeter and the epidemic center. Pei et al. [20] predicted the geographical spread of influenza using human movement data and ensemble population models.

These techniques, however, need access to a sizable quantity of surveillance data, including information on human movement, traffic flow, and other more challenging-to-obtain data. It takes a lot of time and effort to gather these data in order to build forecasting and allocation methods for supplies needs. On the contrary, a dynamic Bayesian network (DBN) uses the time-lag connection between epidemic surveillance data from different places to indicate the probable transmission route of the epidemic. Therefore, we utilize DBN to employ the infectious disease transmission modeling to circumvent the challenges of data collecting issues.

For better prediction, other than the impact of pandemic on the supply-side areas, the supplies condition in the peri-epidemic areas is also rationale. Therefore, research for the selection of supplies allocation supply-side cities and regions for epidemic-centered areas and developing methods suitable for the scenario combining the advantages of model construction and data-driven methods are in tremendous need. To combine the advantages of both model conduction and data-driven methods, graph embeddings can combine the structural and parameter information together. Scholars have constructed many graph embeddings, such as DeepWalk [21] and Node2vec [22]. The DeepWalk algorithm works on homogeneous networks, whereas a bipartite graph is an inherently heterogeneous network, with edges connecting two different types of nodes, left and right. However, it fails to consider modeling the implicit relationship in the graph. BiNE [23] is capable of exploiting both structural and parameter characteristic information to reconstruct fusion and realize the graph embedding. Thus, we choose the BiNE algorithm to construct graph embeddings.

Prompted by the above two considerations, this study considers the selection of the supply-side area around the epicenter for emergency supplies distribution based on the epidemic transmission and supplies condition. First, we developed the infectious disease spatial-temporal transmission path (STTP) from the epidemic center to peripheral areas using DBN. Then, the structure of infectious diseases STTP network and parameters characteristic information data of supplies condition is represented as a bipartite graph for information fusion with the method of BiNE. Finally, a further step of bipartite graph information learning is performed based on the graph convolutional neural network (GNN) method to realize the supplies configuration supply-side areas selection. This method uncovers the hidden relationships between the STTP and supplies allocation supply-side candidate cities and provides excellent references for emergency medical supplies scheduling and donation decision-making. In order to verify the feasibility of the approach, this study applied the model to Wuhan, Hubei province, in the early outbreak of COVID-19 as a case study.

The main contributions of this study are summarized as follows: (1) this study considers supplies allocation from the perspective of the spatial-temporal transmission path of epidemic and provides insights for the effective supply-side area selection of supplies allocation. (2) The regional selection strategy of suppliers balances the resource delivery efficiency and the impact of the epidemic on the supply-side areas, which is a win-win selection. (3) In this study, an emergency medical supplies allocation model based on DBN-GCN methods is constructed, which combines the advantages of data-driven and model-building methods. Experimental results show that the proposed model achieves 87% accuracy in supplies allocation supply-side areas selection and outperforms in comparison with other methods, which can provide valuable suggestions for supplies allocation in emergency situations. This study can help areas around epidemic center to better allocate supplies, make supplies allocation decision reasonably, and improve the effective utilization rate of supplies. Additionally, the selection of epicenter peripheral areas is conducive to meet the explosive growth of demand in short term. It is of great value to urban public health emergency management and emergency risk reduction, and it is of great significance to improve the public's understanding of the importance of supplies allocation.

2. Related Work

2.1. Model Construction. In response to emergency medical supplies shortages and high demand in the initial outbreak of a major infectious disease, this study proposed a strategy that the surrounding areas distribute supplies to the epicenter and gave the methodology to select the supply-side areas. For the candidate supply-side areas, we analyze the influenced extent to peripherals by the epidemic center. The methodology considers both the impact of the pandemic on the peri-epidemic areas and supplies condition in the peri-epidemic areas.

In order to decide the influence factors from the epicenter to the peripherals, one of the most crucial steps is the description of the epidemic spatial-temporal transmission path (STTP) to mine relationships between the periphery and the epicenter of the outbreak. Dynamic Bayesian network is employed to learn epidemic STTP network and to avoid the obstacle of data collection. STTP network investigates the potential transmission direction of infectious diseases using epidemic surveillance data. Because DBN is able to characterize the network making good use of time-lag relationships of multisource data [24], it can be well applied to deal with infectious disease surveillance data. Therefore, DBN is used to learn the structure of infectious disease STTP network. The form of the DBN result is structural characteristic information.

The second step is to collect related data characterizing the impact factors of emergency medical supplies demand in peripheral areas, including socioeconomic conditions, hospital conditions, and emergency medical supplies industries conditions. The form of collected data is parameter characteristic information.

After obtaining both structural and parameter characteristic information on influence factors from the epicenter to the peripherals, the next step is to implement information infusion for predicting supplies demand in the peripherals. We choose BiNE as the method to implement the information infusion because BiNE is capable of exploiting both structural and parameter characteristic information [23] to reconstruct fusion and realize the graph embedding of the epidemic STTP network.

The third and final steps are selecting the suitable emergency supplies allocation areas. To investigate emergency medical supplies demand in the candidate supply-side areas, regressions are then taken driven by selected datasets. GCN is used to extract the spatial features of the topology and learn the multimodel characteristics data. That is, the emergency supplies allocation problem is such transformed into the edge label classification problem in the bipartite graph of the epidemic STTP.

Therefore, the emergency supplies allocation from peripherals to epidemic areas model building process consists of three parts: the epidemic spatial-temporal transmission path network, graph embedding, and supplies allocation supply-side areas selection. The flowchart of the model building is shown in Figure 1.

Therefore, the EMR allocation model building process consists of three parts: the epidemic spatial-temporal transmission path network, graph embedding, and EMR allocation supply-side areas selection. The flowchart of the model building is shown in Figure 1.

2.2. Research Methods. Based on Section 2.1, the details of three parts including epidemic spatial-temporal transmission path network, graph embedding, and EMR allocation supply-side areas selection modes are described in the section.

For readability, a brief overview of the parameters described in the following model is given in Table 1.

2.2.1. Epidemic Spatial-Temporal Transmission Path (STTP). To select the suitable supply-side areas for disturbing emergency medical supplies to the epidemic center, we first predict the impact of the pandemic from the epidemic center to the candidate supply-side areas so that the supplies demand of the peripheral areas can be estimated. Considering how to predict the effect of the pandemic on the peripheral areas, we choose epidemic spatial-temporal transmission

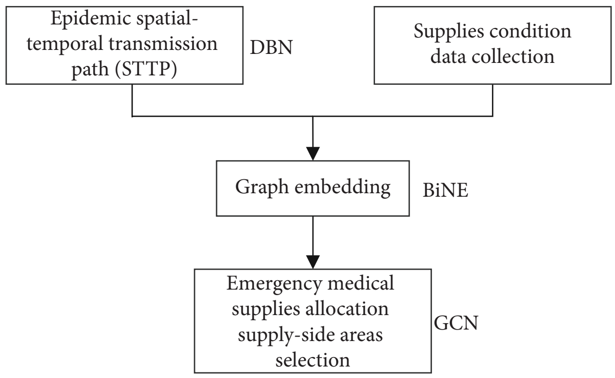


Figure 1: Model building flowchart.

Table 1: Parameters in the model.

Parameters	Description
G	Epidemic STTP network
$X = \{I(i,t)\}$	Infectious disease dataset
I	The number of new daily cases
i	The label of region or area
t	The time of new daily cases
V_t	The connection strength between two areas at time t
n	The number of statistical days
c	Epidemic center area node
a	Epidemic peripheral area node
$\mathbf{h}$	Embedding matrix
T	The number of convolution layer
e	Edge between the nodes
ω	Training weight vector
$\mathbf{G}$	Bipartite graph
$\mathbf{C}$	Nodes set of epicenter areas
$\mathbf{A}$	Nodes set of epidemic peripheral areas
m	The number of nodes in $\mathbf{C}$
n	The number of nodes in $\mathbf{A}$
$\mathbf{E} \subseteq \mathbf{C} \times \mathbf{A}$	The edge of nodes in sets $\mathbf{C}$ and $\mathbf{A}$
$e_{i,j} \in \mathbf{E}$	The edge connection between the node c_i and a_j
ω_{ij}	Edge weight
$\mathbf{f}$	Feature vectors
$\mathbf{F}$	Feature matrixes
l	Edge labels

path to form a structural network. The epidemic STTP refers to the time-lag correlation of the epidemic spread from place B (i.e., the epidemic center area) to place D (i.e., the peripheral of the epidemic) from a statistical point of view, which can be expressed by B $\longrightarrow$ D. The directional arc "$\longrightarrow$" between B and D indicates that node A has time-lag effects on the node B, and this kind of directional arc is defined as a spatial-temporal path in this study. All arcs will form a network, when two or more areas exist time-lag correlations, namely, epidemic spatial-temporal transmission path (STTP) network. Specifically, networks contain two types of information: (1) structural information, which is related to the directional arc; (2) parametric information, which measures the connection strengths between different nodes. Epidemic spatial-temporal transmission path network can simulate how infectious diseases are spread to the peripheral areas.

We define the epidemic STTP network as $G = (X, A)$, where X is the infectious disease dataset of the numbers of new daily cases $I(i,t)$ in different areas, e.g., $X = \{I(i,t)\}$, $(i = 1, 2, 3, \ldots\ldots)$, $(t = 1, 2, 3, \ldots\ldots)$, for convenience, X_t is used to display $X = \{I(i,t)\}$ as follows, and A is the arc set between any two areas in dataset X. $I(i,t)$ can be regarded as a time series. The primary goal of this part is to simulate the time-lag correlation between two time series $I(i,t)$ and $I(j,t)$ $(i \neq j)$ using dynamic Bayesian network (DBN) to realize epidemic STTP network structure learning. Bayesian network is a probabilistic graph model, which represents a group of random variables and their conditional dependencies through a directed acyclic graph (DAG) [25]. Formally, Bayesian networks use nodes to represent random variables (whose probability is Bayesian probability). Each node is associated with a node probability function that takes as input the values of all parent nodes and gives the probability of the node random variable. The process is illustrated in Figure 2.

The estimated value is the epidemic transmission probability of city i on city j, and it is obtained by combining the datasets with the Bayesian formula as shown in the following equation:

$$P(I(j) \mid I(i)) = \frac{P(I(j))P(I(i) \mid I(j))}{P(I(i))}. \tag{1}$$

Dynamic Bayesian network (DBN) is a dynamic directed acyclic graph, which uses nodes and arcs to represent the conditional probability dependence among a group of time series [26]. In the DBN model, an arc is drawn between two variables (in our cases, the two successive time points). For example, the arc from $I(i, t-1)$ to $I(j,t)$ represents the numbers of new daily cases in city j at time t which conditionally depends on that in city i at time $(t-1)$. In this study, we implemented our estimation using a first-order conditional-dependent approximation algorithm initialized by a DBN model. An overview of the procedure is presented in Figure 3.

In Figure 3, the probabilities for variance are time-dependent $(t = 1, 2, \ldots, n)$. The state of variable $I(i,t)$ is determined by $I(j,t)$ at any time, while the status of estimated values is determined by both $I(i,t)$ and $I(j,t)$. V_t is the connection strength (traffic flow from city i to city j) directional arc between city i and city j at time t. The joint probability distribution is obtained by the following equation:

$$P(X_t) = P(I(i,t), I(j,t), V_t) = P(I(i,t))P(I(j,t) \mid I(i,t))P(V_t \mid I(i,t), I(j,t)). \tag{2}$$

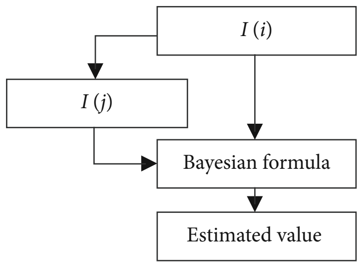


Figure 2: Bayesian network model.

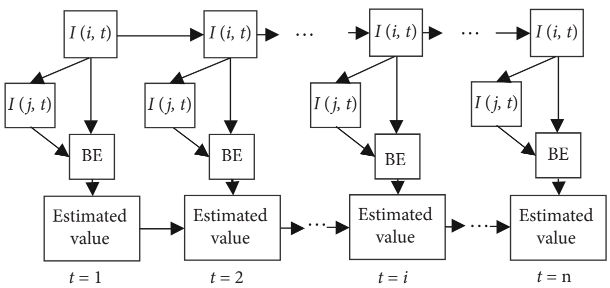


Figure 3: Dynamic Bayesian network model.

And the conditional probability between $I(i,t), I(j,t)$ and V_t is as follows:

$$P(V_t \mid I(i,t), I(j,t)) = \frac{P\big(I(i,t), I(j,t) \,\big|\, V_t\big)P(V_t)}{P(I(i,t))} = \frac{\sum_{t=1}^{n} P(V_t, I(i,t), I(j,t))}{P(I(i,t))}. \tag{3}$$

With the unsupervised learning algorithm for DBN learning, the training of an epidemic STTP network can be implemented in two steps. First, encode conditional dependencies in a DAG with the first-order algorithm. Then, infer the epidemic STTP network structure, so that we can get the result of the epidemic effect on the candidate supply-side areas in a structural formation. Therefore, the next step is to assess the effect of the epidemic spread on the peripheral cities.

2.2.2. Supplies Characteristics Parameters. For better preference of the prediction of the supplies demand in the candidate supply-side areas, the characteristics of supplies demand in supply-side nodes in the epidemic STTP network are needed. Considering the important factors of medical supplies in the epidemic STTP network, we select parameters including economy, market, and technology from the perspective of supplies production, consumption, and so on. The details of factors and selection reasons are given as follows:

(i) Sociodemographic factors

(1) Economic development level: Areas with an opposable high level of economic development usually display a relatively high density of health sector and higher demand for medical supplies in the pandemic. We used gross domestic product (GDP) to represent the level of economic development.

(2) Population: Typically, the larger the number of populations, the greater the supplies demand for emergency public health event response.

(ii) Market factors

(1) Industrial structure: In response to supplies shortages during the early age of public health emergencies, some manufacturing enterprises in China, such as car manufacturers, transfer their main business to supplies production temporarily. The industrial structure is expected to influence the demand substantially.

(2) Production quantity: Supplies production is adjusted for public health emergencies but with time lag. Current total production quantity of supplies is a significant consideration in allocating materials.

(3) Pricing index: The materials pricing index indicates the influence on the supplies price of public health emergencies. Market price depends on demand, and therefore, we consider the market pricing index a linear function coefficient.

(4) Transport efficiency: It could happen that public health emergencies have some impact on transport efficiency. For example, the material transmission pathways are blocked and subjected to natural hazards.

(iii) Technical factors

The improvement of material production technology will result in the impact on the production efficiency [27]. In this study, the number of material institutions represents the supplies production technological level.

Therefore, we select three types of supplies characteristic parameters to reflect the emergency medical supplies condition from both production and consumption.

2.2.3. Graph Embedding. The epidemic STTP network is a structural information, and the characteristic is the parameter information. To combine the two types of information together for better estimating the supplies situation under epidemic in the candidate supply-side areas, graph embedding is used in many fields to represent connection patterns among complex system components [28]. It can map nodes to low-dimensional vector space, combine subsequent graph problems such as node classification, and link prediction and node clustering with existing machine learning methods on the premise of node representation vector [29]. The process captures the topology of the graph, vertex-to-vertex relationships, and other relevant information about the graph, subgraph, and vertices, that is, information fusion. The obtained low-dimensional vector is used as input information for downstream machine learning. The graph embedding method has superior performance compared to traditional method such as planar graph [30, 31].

Bipartite graph is a universal data structure used to model the relationship between two types of entities, such as an event graph. Therefore, the STTP network obtained in Section 2.2.1 is a bipartite graph. We choose BiNE as the method to implement the information infusion because BiNE [23] is capable of exploiting to reconstruct fusion and realize the graph embedding of the epidemic STTP network, combining both structural and parameter characteristic information together. It can maintain the long-tail distribution of nodes in a bipartite graph by performing a biased random walk [32]. In this way, we can get both epidemic spreading information and supplies condition preparing for the prediction of supplies situation in peripheral areas.

2.2.4. GCN-Based Emergency Medical Supplies Allocation. In this section, we will estimate the supply situation in the peripheral areas to select the suitable candidate supply-side areas for distributing supplies to the epicenter. Graph convolutional network (GCN) is a neural network architecture that can effectively handle embedding graph [33], mainly for convolution operation of graph structure data. The core concept of GCN is to iterate and aggregate the nodes' features in the embedding graph. As opposed to other methods targeted at the processing graph, GCN is more suitable for processing the bipartite graph with topological structure and selected for the demand allocation. Based on these two reasons, GCN is selected to implement the prediction step in our study. We hypothesized that supplies demand in the periphery of the outbreak could be simulated by using epidemic information and supplied condition between regions. GCN-based supplies allocation decision model is constructed based on the structural and parameter information obtained from Sections 2.2.1 and 2.2.2 separately. The GCN model construction process is shown in Figure 4.

The concrete steps are as follows:

Step 1. Aggregate the adjacent nodes features.

Step 2. Update the initial nodes features through filters.

Step 3. Concatenate the updated features.

Step 4. Edge-labelled classification. The edge embedding matrix is obtained through Hadamard product [34]:

$$\left[\mathbf{h}_{e(\mathrm{ij})}^{(T)}\right]_l = \left[\mathbf{h}_{c_i}^{(T)}\right]_l \cdot \left[\mathbf{h}_{a_j}^{(T)}\right]_l, \tag{4}$$

where $\mathbf{h}_{c_i}^{(T)}$ represents the embedding matrix of epidemic center area node c_i, $\mathbf{h}_{a_j}^{(T)}$ represents the embedding matrix of epidemic peripheral area node a_j, and $\mathbf{h}_{e(\mathrm{ij})}^{(T)}$ represents the embedding matrix of edge connecting the node c_i and node a_j.

Step 5. SVM classification.

Support vector machine (SVM) classifier is used to select epidemic peripheral area as the allocation supply-side cities and regions. Specifically, the label is expressed as follows:

$$\hat{\mathbf{y}} = \mathrm{sgn}\left(\mathbf{H}_e^{(T)}\boldsymbol{\omega}\right), \tag{5}$$

where $\mathrm{sgn}(\bullet)$ is the sigmoidal function and ω is the training weight vector.

3. Case Study

This section presents a case study of the model established in Section 2. First, data were collected from the epidemic diseases COVID-19 in China. After that, the method of graph embedding BiNE is used to construct a bipartite graph of the epidemic STTP network. Finally, the selection of epidemic peripheral areas as the emergency medical supplies allocation supply-side is based on a graph convolutional neural network.

3.1. Data Collection and Processing. The dataset analyzed in this study includes 16 epidemic-centered areas (from Hubei province) and 19 epidemic peripheral areas (Anhui, Chongqing, Shaanxi, Jiangxi, Hunan, Henan, Beijing, Tianjin, Guangdong, Sichuan, Yunnan, Shanghai, Shandong, Zhejiang, Hainan, Guizhou, Ningxia, Hebei, and Jiangsu) from January 10, 2020, to March 31, 2020, during the early outbreak of COVID-19. The epidemic peripheral areas are selected with a relatively high level of economic development or relatively recent geographical distance. The daily accumulative number of confirmed cases is obtained from the National Center for Disease Control and Prevention (CDC). The division of administrative areas follows the government's official zoning principle.

The purpose of this study is to select some epidemic peripheral areas as the supply-side based on their supplies demand, providing decision-making basis for supplies allocation. The statistical data were summarized as 304 ($16^{*}19=304$) couples of epidemic-centered areas-peripheral areas. We represent each label as a pair of area nodes, for the high and the low supply intention, respectively. Features of the epidemic-centered and peripheral areas displayed in Table 2 were collected from the National Bureau of Statistics of China, China Industry Information Research network. In addition, the geographic distance between the epidemic center and the peripheral area was measured according to [35]. Features of epidemic peripheral areas, taking Jiangsu Province as an example, are shown in Table 3.

The difference in dimensions and orders of magnitude among the features is prone to errors. The dataset was preprocessed using [36] as follows:

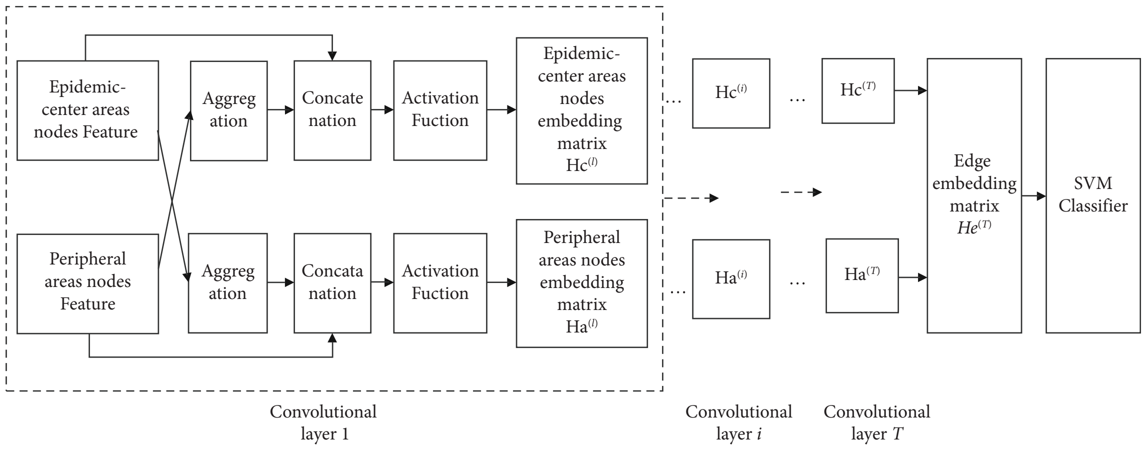


FIGURE 4: GCN-based emergency medical supplies allocation model.

TABLE 2: Features of epidemic-centered and peripheral areas.

Node type	Feature category	Feature
Epidemic-centered areas	Epidemic status	Proportion of the infected population
	Emergency medical supplies status	Supply/demand ratio
Epidemic peripheral areas	Economic status	GDP
		Industrial structure
	Demographic status	Population
		Production
	Emergency medical supplies status	Logistics efficiency
		The number of corresponding research institutions
		Market pricing index
		Emergency medical supplies demand
		STTP
	Epidemic status	Proportion of infected population

TABLE 3: Features of epidemic peripheral areas in Jiangsu Province.

Date	GDP (trillion RMB)	Industrial structure	Population (ten thousand)	Production (ten million RMB)	Market pricing index	Logistics efficiency (%)	The number of corresponding research institutions	EMR demand (ten thousand tons)
2019.9	9.073	12.1	7281	241.95	97.72	46.24	0	87.62
2019.10	8.329	12.12	7028	201.27	98.41	41.72	0	88.71
2019.11	6.870	13.08	6899	330.82	103.47	34.03	6	139.96
2019.12	5.533	13.86	7093	405.89	141.83	39.79	7	146.79
2020.1	4.137	14.67	7405	412.31	132.64	42.87	8	158.65
2020.2	4.543	15.79	7431	437.89	127.95	45.98	9	152.51
2020.3	4.677	16.81	7582	451.67	121.89	47.73	7	147.69

Data source: National Bureau of Statistics of China, China Industry Information Research network.

$$\overline{x} = \begin{cases} \dfrac{x - X_{\min}}{X_{\max} - X_{\min}}, & \text{if the attribute of x benefits for supplying,} \\[2ex] \dfrac{X_{\max} - x}{X_{\max} - X_{\min}}, & \text{otherwise,} \end{cases} \tag{6}$$

where x is the data before being processed and $\overline{x}$ is the processed data. And X is a list of data, $X_{\min}$ is the smallest data among X, and $X_{\max}$ is the largest data among X. After data preprocessing, we can change Table 2 into Table 4 so that it can be the parameter input of graph embedding. For readability, we label the dataset from Tables 3 in 4, and we can see the corresponding data after processing in Table 5.

3.2. Graph Embedding of Epidemic STTP Network. Epidemic STTP network is obtained based on DBN-based methods in Section 2.2.1, typically reflecting the impact of epidemic from epicenter to the peripheral areas. Using BiNE methods, graph embedding can combine information from both structural information of the epidemic STTP network and nodes parameter. In the graph embedding process, dataset of the epidemic-centered and epidemic peripheral areas can be presented in the form of bipartite graph $\mathbf{G} = (\mathbf{C}, \mathbf{A}, \mathbf{E})$ and tag its edge label classification, as shown in Figure 5.

In Figure 5, $\mathbf{C}$ and $\mathbf{A}$ indicate nodes set of epicenter areas and epidemic peripheral areas, respectively, and $c_i \in \mathbf{C}$, $a_j \in \mathbf{A}$, $m = |\mathbf{C}|$, and $n = |\mathbf{A}|$ are the numbers of nodes in $\mathbf{C}$ and $\mathbf{A}$, respectively. $\mathbf{E} \subseteq \mathbf{C} \times \mathbf{A}$ is the edge of nodes in sets $\mathbf{C}$ and $\mathbf{A}$, and $e_{i,j} \in \mathbf{E}$ indicates the edge connection between the node c_i and a_j. The edge weight ω_{ij} is calculated in the following way:

$$\omega_{\mathrm{ij}} = \exp\left(-\mathrm{distance}\left(c_i, a_j\right)\right), \tag{7}$$

where distance $(\bullet)$ represents the Euclidean distance between two nodes and exp $(\bullet)$ represents an exponential function.

The feature vectors of node c_i and a_j are symbolized by $\mathbf{f}_{c_i}$ and $\mathbf{f}_{a_i}$, respectively, and generate the matrixes $\mathbf{F}_\mathbf{c} = [\mathbf{f}_{c_1}, \mathbf{f}_{c_2}, \ldots, \mathbf{f}_{c_M}]^\mathbf{T}$ and $\mathbf{F}_\mathbf{a} = [\mathbf{f}_{a_1}, \mathbf{f}_{a_2}, \ldots, \mathbf{f}_{a_N}]^\mathbf{T}$. Let $l_{i,j} \in \{-1, 1\}$ be the label of edge $e_{i,j}$ (-1 = low epidemic impact in the epidemic peripheral area and 1 = high epidemic impact in the epidemic peripheral area). The edge labels compound set is denoted as $\mathbf{L} = \left\{l_{i,j}\right\}$ and is collected from the result of DBN in Section 2.2.1. Therefore, we can combine the structural and parameter information together in the epidemic STTP network. And then, the matrixes will be the input of the GCN-based emergency medical supplies allocation supply-side selection model.

3.3. GCN-Based Emergency Medical Supplies Allocation. In this study, to solve the issue of supplies allocation supply-sider selection, we can transform it to an epidemic-centered and peripheral areas pair label classification problem. Based on the obtained feature of edge embedding in Section 2.2.3, predict the selection of supplies allocation supply-sider in the epidemic peripheral areas using the given GCN model in Section 2.2.4.

The collected datasets are randomly divided into training, verification set, and test sets. Given the features of the epidemic-centered and peripheral areas $\mathbf{F}_\mathbf{c}$ and $\mathbf{F}_\mathbf{a}$, the distance between epicentral area and peripheral areas are given in $\mathbf{A}$ seperately. The labeled $\mathbf{Y}_{\mathbf{train}}$ and $\mathbf{Y}_{\mathbf{val}}$ account for 10%–80%, and the remaining edge labels $\mathbf{Y}_{\mathbf{test}}$ are predicted based on GCN. In the GCN model, the weight parameters are initialized using the parameter rectified linear units (PReLU) as the activation function, and the momentum of all convolution layers is set to 0.9. Adam optimizer [37] is selected as the training model, with the learning rate of 0.01 and the maximum number of training iterations of 500.

Table 4: Label data in Table 3.

No.	I	II	III	IV	V	VI	VII	VIII
1	9.073	12.1	7281	241.95	97.72	46.24	0	87.62
2	8.329	12.12	7028	201.27	98.41	41.72	0	88.71
3	6.870	13.08	6899	330.82	103.47	34.03	6	139.96
4	5.533	13.86	7093	405.89	141.83	39.79	7	146.79
5	4.137	14.67	7405	412.31	132.64	42.87	8	158.65
6	4.543	15.79	7431	437.89	127.95	45.98	9	152.51
7	4.677	16.81	7582	451.67	121.89	47.73	7	147.69

4. Results

4.1. Epidemic STTP Network. The epidemic STTP network concludes in two parts. The first part is the structure of influenza STTP achieved from the DBN learning, and the second part is estimating the impact strength of traffic distance in the influenza transmission. Considering the officially given incubation period of the epidemic COVID-19 varying from 1 week to 3 weeks, we set up three types of lag, respectively (that is, 1 week, 2 weeks, and 3 weeks). For readability, we transform the type of result from matrix to graph and draw the results directly in the map. The results of the DBN-based epidemic STTP network are presented in Figure 6 with daily accumulative number of confirmed cases data 1 week, 2 weeks, and 3 weeks lag, respectively, considering the latency of COVID-19. We can see the STTP of COVID-19 from the epidemic center Hubei to the surrounding areas such as Henan, Hunan, and Jiangxi. Also, Chongqing is properly controlled in the 1-week lag because of the traffic influence. But in the 2 weeks and 3 weeks lag, Chongqing is influenced with no doubt.

In the case of large infectious diseases, some neighboring cities indicate that there may be some potential stable paths between these cities, which can help the National Center for Disease Surveillance locate the key areas for infectious disease surveillance and provide a reference for the decision of allocating emergency medical supplies to epidemic centers.

4.2. GCN Model Training Results. In the GCN model, the input comprises the integrated bipartite graph data, encompassing both the topological structure and parameter features within the epidemic STTP network. The training results show that the GCN model could effectively select the suitable supplies allocation supply-sider and provide a reliable decision-making basis for supplies allocation. The performance of accuracy rate and $F1$ values under different test set proportions is provided in Table 6.

It can be observed in Table 4 that the model achieved the highest accuracy of 87% and a relatively high $F1$ value with a test set proportion of 20%. In addition, in order to verify the reliability of the GCN model, this study selected some

Table 5: Data in the table after processing.

No.	I	II	III	IV	V	VI	VII	VIII
1	0.99880335	0.998404115	0.039699288	0.968088895	0.98711158	0.993901345	1	0.988443682
2	0.998901477	0.998401477	0.073067792	0.973454234	0.987020575	0.994497494	1	0.988299921
3	0.999093907	0.998274862	0.090081773	0.956367713	0.986353205	0.995511738	0.999208652	0.981540491
4	0.999270245	0.998171986	0.064494856	0.946466631	0.981293854	0.994752044	0.999076761	0.980639673
5	0.999454366	0.998065154	0.023344764	0.945619889	0.982505935	0.994345819	0.998944869	0.979075442
6	0.99880335	0.998404115	0.039699288	0.968088895	0.98711158	0.993901345	1	0.988443682
7	0.998901477	0.998401477	0.073067792	0.973454234	0.987020575	0.994497494	1	0.988299921

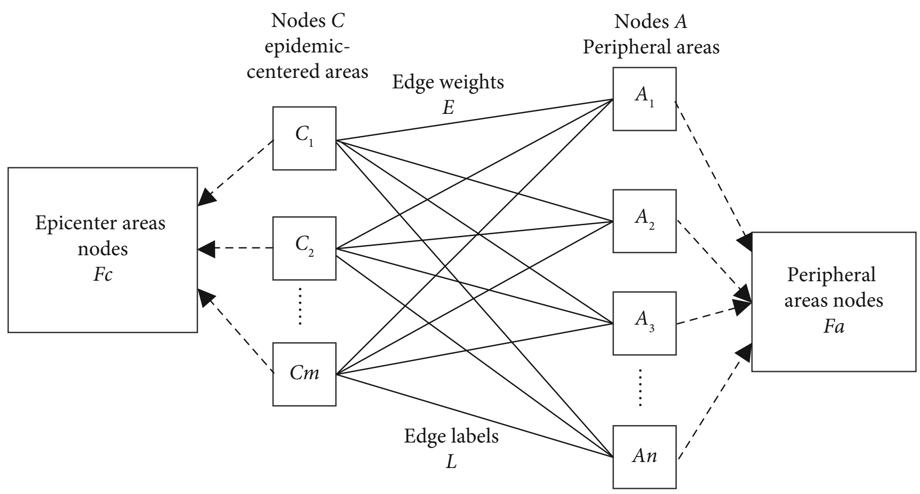


Figure 5: Bipartite graph of the epidemic-centered and peripheral areas.

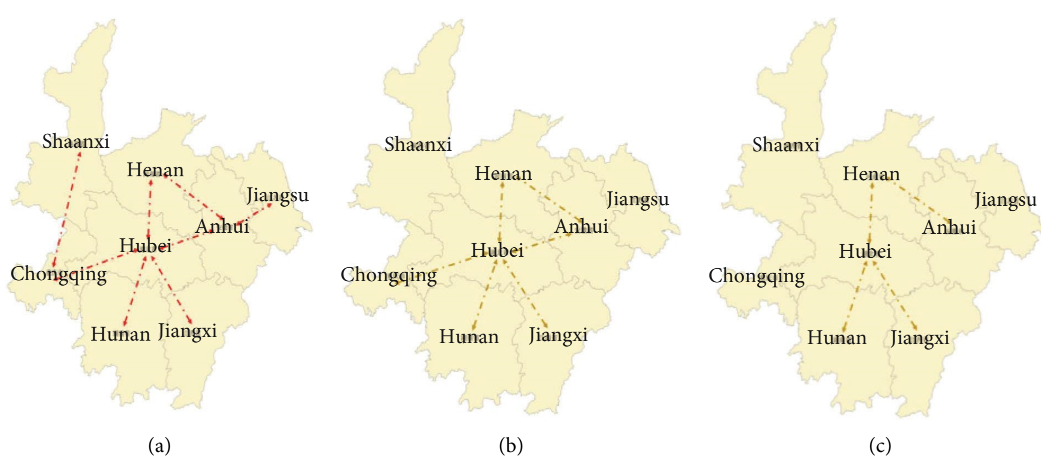


Figure 6: Epidemic STTP network of COVID-19: (a) 1-week lag, (b) 2-week lag, and (c) 3-week lag.

Table 6: Emergency medical supplies allocation in the epidemic center and peripheral areas.

Test set ratio	10%	20%	40%	60%	80%
Accuracy rate	0.85	0.87	0.84	0.82	0.7
$F1$ value	0.89	0.88	0.87	0.85	0.84

traditional machine learning classifiers: SVM [38] and GMM [39], statistical models: HAN [40], and the graph-based methods: Node2Vec [41] and GraphSAGE [42], as references to contrast with GCN method. In Tables 7 and 8, the emergency medical supplies allocation comparison performances are shown in the form of accuracy rate and $F1$ value.

To better interpret the results, we display the comparison results in the form of a diagram in Figure 7.

4.3. Emergency Medical Supplies Allocation Supply-Side Selection. Finally, the trained model was applied to the adjusted dataset in the early outbreak of COVID-19 in Hubei province. Data of daily new infections in the epicenter is provided as adjusted through incubation period length and subsequent relevant reports to fit with the designed emergency medical supplies allocation model. In the simulation, epidemic peripheral areas were not only considered as areas adjacent to the epidemic center but also areas with heavily interconnected in the context of transportation or developed economies. Based on the trained GCN-based emergency medical supplies allocation model, the selection of emergency medical supplies allocation supply-sider was given with a "high" willingness label to supply the epidemic-centered areas after the COVID-19 outbreak. The result is Beijing, Shaanxi, Henan, Jiangsu, Anhui, Sichuan, Chongqing, Hunan, Guangdong, Jiangxi, Zhejiang, and Shanghai and is marked in yellow in Figure 8. Besides, we also mark the unwanted candidate supply-side cities as blue. So that we can distinguish emergency medical supplies allocation supply-sider with high willingness from unsuitable ones.

4.4. Discussion. As is shown in Figure 7, GCN has the best overall performance. In terms of accuracy, the GCN model gained a training accuracy of 2.4–8.0% over other contrasting methods with 20–60% of the dataset for the test set. However, the performance of the GCN model is greatly degraded due to the influence of minimally overtraining when the test ratio reaches 80%. Generally speaking, GCN has certain advantages over other comparison approaches. There are two main reasons as follows: first, GCN takes the features and network structure of the bipartite graph into account. In contrast, Graph SAGE is designed for single-partite graphs and is not suitable for bipartite graphs. Second, the GCN model designed an end-to-end model optimization based on the cost function. While Node2Vec and SWM are unsupervised learning methods, they cannot make full use of the topology of the original graph in the training set to optimize the model. Therefore, the validity of emergency medical supplies allocation based on the GNN model is further verified.

The cross marks in Figure 8 indicate the epidemic peripheral areas, and the central point indicates the central area of the epidemic. As can be seen from Figure 8, GCN model gives 13 high willingness to supply the epidemic-centered areas (Anhui, Chongqing, Shaanxi, Jiangxi, Hunan, Henan, Beijing, Shanghai, Zhejiang, Guangdong, Sichuan, Hebei, and Jiangsu) and 6 low supply willingness to supply the

Table 7: Emergency medical supplies allocation accuracy of different methods.

Method	Test set ratio				
	10%	20%	40%	60%	80%
SVM	0.73	0.71	0.71	0.69	0.66
GMM	0.81	0.85	0.80	0.74	0.78
Node2Vec	0.67	0.57	0.66	0.65	0.61
Graph SAGE	0.66	0.72	0.70	0.70	0.67
HAN	0.64	0.72	0.68	0.66	0.65
GCN	0.85	0.87	0.84	0.82	0.7

Table 8: Emergency medical supplies allocation $F1$ values of different methods.

Method	Test set ratio				
	10%	20%	40%	60%	80%
SVM	0.76	0.79	0.75	0.74	0.73
GMM	0.84	0.82	0.84	0.83	0.83
Node2Vec	0.52	0.53	0.48	0.56	0.44
Graph SAGE	0.19	0.40	0.47	0.45	0.38
HAN	0.65	0.72	0.73	0.74	0.80
GCN	0.89	0.88	0.87	0.85	0.84

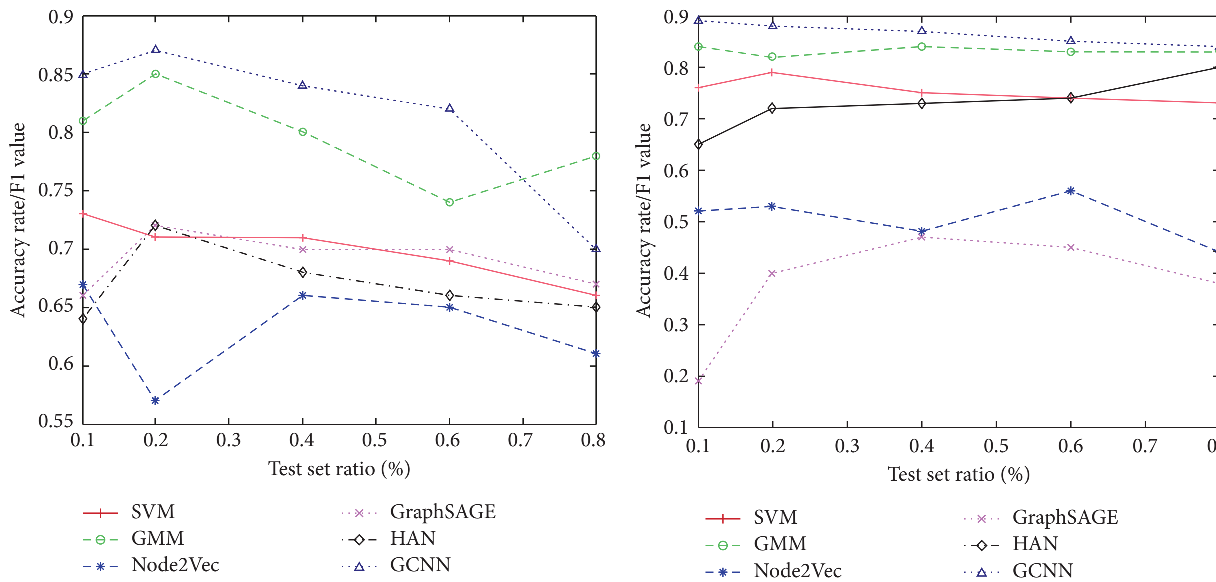
outbreak of the peripheral areas (Yunnan, Hainan, Guizhou, Ningxia, Shandong, and Tianjin). This result makes sense as epidemic peripheral areas with high supply intentions have both economic advantages and limited exposure to the spread of the epidemic. This result has important reference value for the decision-making of supplying from the epidemic peripheral areas to the center, which not only contributes to fast response to emergencies in the future but also makes preparation for public health emergency management.

The implications can be given as follows:

(1) Improving the capacity of emergency management departments or enterprises in different regions should allocate supplies with special emphasis on the basis of risk assessment, historical lessons, and experience summary and in accordance with comprehensive factors such as the population, social and economic situation, and traffic flow of the region.

(2) It is important for the government to evaluate the needs of the amount and types of supplies in the epidemic peripheral areas. The premise of regional supplies mobilization is to meet their own supply, while the emergency capacity of different regions varies greatly, and it is difficult to achieve a unified standard for the type and quantity of supplies. In the early stage of large-scale public health emergencies, the allocation of supplies in the peripheral areas can solve the urgent needs of the disaster areas.

(3) It is better to establish the strategic peripheral areas allocation mechanism of supplies. The reserve of supplies occupies a large number of places, and the corresponding costs would be high. Some certain types of supplies only apply to the rare large-scale infectious diseases and cannot be consumed under



(a) (b)

Figure 7: Results of GCN and comparison methods: (a) accuracy rate and (b) $F1$ value.

Figure 8: Emergency medical supplies allocation supply-side selection results based on the GCN model.

normal conditions. Blind reserve may cause a huge backlog and waste. Therefore, in the event of an outbreak, the emergency medical supplies allocation between epidemic centers and peripheral areas can improve the supplies use efficiency.

5. Conclusions

In the present study, an emergency medical supplies allocation method based on graph convolutional neural network (GCN) and bipartite graph of epidemic spatial-temporal transmission path (STTP) network in major infectious diseases. The first epidemic STTP network was obtained based on DBN, and then, the method of BiNE was used to transfer the structure and parameter features information into a bipartite graph. The last step of the GCN model is used to learn the epidemic-centered-peripheral areas pair, giving the decision-making basis for selecting supplies allocation supply-siders. The experimental results show that the supplies allocation accuracy is the highest with the loss balance hyperparameter α 0.1, the number of convolution layers 3, and the test set 20%. Compared with other classical methods and applied to a practical case, the reliability of the model is verified.

The main conclusions are presented as follows:

(1) The DBN-based epidemic STTP network makes full use of the traditional concept of transmission route in epidemiology, so as to better solve the problem of constructing a potential infectious disease spatial transmission network in the case with limited data. Most of the latest studies on infectious disease transmission networks are based on theoretical physics or Internet disciplines, but few of them are involved in the construction of infectious disease spatial transmission networks from the perspective of spatial and temporal distribution. Therefore, the STTP not only provides structural information parameters for learning supplies allocation in this model but also provides a theoretical basis for further studying the transmission and prevalence of infectious diseases from the spatial-temporal dimension.

(2) The supplies allocation method based on GCN deals with epidemic STTP network in the form of a bipartite graph fully solves the problems of complex data structures such as high dimension, high noise,

and nonlinear. The combination of DBN and GCN models can be well applied to complex infectious disease transmission networks in reality.

The contribution points of this study are as follows. (1) The emergency medical supplies allocation method proposed in this paper considers the epidemic peripheral areas as a supply-sider. The strategic assessment involves the selection of supply-side areas, considering both the impact of infectious diseases on potential suppliers and the distribution distances between epidemic center. (2) This study provides a reference for the supplies allocation for public health emergencies such as large-scale infectious diseases. The resource utilization efficiency of supplies is improved, and the possible waste of supplies is avoided. In extreme situations, such as the shortage of emergency supplies, the supplies of different cities can be fully deployed for emergencies. (3) This study is of great significance to urban public health emergency management and emergency risk reduction and has practical value in assisting local governments in making supplies allocation and scheduling decisions.

There are still some deficiencies in this study. The study gives emergency medical supplies allocation supply-sider without a specific list such as allocation pairs and quantity. In the future, this study is expected to further investigate the supplies allocation mechanisms.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are grateful to the case company for permitting and supporting this research. This work was financially supported by Graduate Research and Innovation Projects of Jiangsu Province (KYCX21_1035).

References

[1] Y. Wang, V. M. Bier, and B. Sun, "Measuring and achieving equity in multiperiod emergency material allocation," *Risk Analysis*, vol. 39, no. 11, pp. 2408–2426, 2019.

[2] M. R. Broccoli, "Emergency medicine training programs in low- and middle-income countries: a systematic review," *Annals of Global Health*, vol. 86, no. 1, pp. 1–18, 2020.

[3] Q. S. Zhou and T. L. Olsen, "Inventory rotation of medical supplies for emergency response," *European Journal of Operational Research*, vol. 257, no. 3, pp. 810–821, 2017.

[4] Z. S. Hou and Z. Wang, "From model-based control to data-driven control: survey, classification and perspective," *Information Sciences*, vol. 235, pp. 3–35, 2013.

[5] S. E. Wright and S. Lim, "Solving nested-constraint resource allocation problems with an interior point method," *Operations Research Letters*, vol. 48, no. 3, pp. 297–303, 2020.

[6] S. Mehrotra, H. Rahimian, M. Barah, F. Luo, and K. Schantz, "A model of supply-chain decisions for resource sharing with an application to ventilator allocation to combat COVID-19," *Naval Research Logistics*, vol. 67, no. 5, pp. 303–320, 2020.

[7] R. Aringhieri, G. Carello, and D. Morale, "Supporting decision making to improve the performance of an Italian Emergency Medical Service," *Annals of Operations Research*, vol. 236, no. 1, pp. 131–148, 2016.

[8] M. Amorim, S. Ferreira, and A. Couto, "Road safety and the urban emergency medical service (uEMS): strategy station location," *Journal of Transport & Health*, vol. 6, pp. 60–72, 2017.

[9] G. Grekousis and Y. Liu, "Where will the next emergency event occur? Predicting ambulance demand in emergency medical services using artificial intelligence," *Computers, Environment and Urban Systems*, vol. 76, no. JUL, pp. 110–122, 2019.

[10] J. Wohlmann, C. Juhra, L. Warnken et al., "Medical emergency datasets can improve the comprehensiveness of handover information: a simulation study," *Telemedicine and e-Health*, vol. 27, no. 9, pp. 1046–1053, 2021.

[11] N. Pérez-Rodríguez and J. Holguín-Veras, "Inventory-allocation distribution models for postdisaster humanitarian logistics with explicit consideration of deprivation costs," *Transportation Science*, vol. 50, no. 4, pp. 1261–1285, 2016.

[12] A. Boin, E. Stern, and B. Sundelius, *The Politics of Crisis Management: Public Leadership under Pressure*, Cambridge University Press, Cambridge, UK, 2016.

[13] M. Mehrtash, A. Kargarian, and A. Mohammadi, "Distributed optimisation-based collaborative security-constrained transmission expansion planning for multi-regional systems," *IET Generation, Transmission and Distribution*, vol. 13, no. 13, pp. 2819–2827, 2019.

[14] G. Pescaroli, R. Wicks, G. Giacomello, and D. Alexander, "Increasing resilience to cascading events: the M. OR. D. OR. scenario," *Safety Science*, vol. 110, no. 12, pp. 131–140, 2018.

[15] Y. Qiu, M. Shi, X. Zhao, and Y. Jing, "System dynamics mechanism of cross-regional collaborative dispatch of emergency supplies based on multi-agent game," *Complex & Intelligent Systems*, vol. 9, no. 3, pp. 2321–2332, 2021.

[16] P. Urrutia, D. Wren, C. Vogiatzis, and R. Yoshida, "SARS-CoV-2 dissemination using a network of the US counties," *Operations Research Forum*, vol. 3, no. 2, p. 29, 2022.

[17] J. Liu, C. Dong, and S. An, "Integration and modularization: research on urban cross-regional emergency cooperation based on the network approach," *International Journal of Disaster Risk Reduction*, vol. 82, pp. 103375–103400, 2022.

[18] Y. Wang, "Multiperiod optimal allocation of emergency resources in support of cross-regional disaster sustainable rescue," *International Journal of Disaster Risk Science*, vol. 12, no. 3, pp. 394–409, 2021.

[19] K. Liu, "Research on emergency personnel scheduling considering the psychological perception of both disaster victims and emergency personnel," *IAENG International Journal of Applied Mathematics*, vol. 52, no. 3, pp. 9–22, 2022.

[20] S. Pei, S. Kandula, W. Yang, and J. Shaman, "Forecasting the spatial transmission of influenza in the United States," *Proceedings of the National Academy of Sciences*, vol. 115, no. 11, pp. 2752–2757, 2018.

[21] B. Perozzi, "Don't walk, skip!: online learning of multi-scale network embeddings," *ACM*, vol. 7, pp. 1–8, 2017.

[22] A. Grover and J. Leskovec, "node2vec: scalable feature learning for networks," in *Proceedings of the 22nd ACM*

SIGKDD International Conference, San Francisco, CA, USA, December 2016.

[23] H. Chen, "Social boosted recommendation with folded bipartite network embedding," *IEEE Transactions on Knowledge and Data Engineering*, vol. 9, p. 1, 2020.

[24] Q. Chen and G. Pan, "A structure-self-organizing DBN for image recognition," *Neural Computing & Applications*, vol. 33, no. 3, pp. 877–886, 2021.

[25] M. Liu, J. Ma, Y. Duo, and T. Sun, "Reliability analysis of gasifier lock bucket valve system based on DBN method," *Mathematical Problems in Engineering*, vol. 2021, Article ID 8840545, 10 pages, 2021.

[26] R. Levy, "Dynamic bayesian network modeling of game-based diagnostic assessments," *Multivariate Behavioral Research*, vol. 54, no. 6, pp. 771–794, 2019.

[27] S. Zhang, M. Xu, Y. Yang, and Z. Song, "Technological innovation, production efficiency, and sustainable development: a case study from shenzhen in China," *Sustainability*, vol. 13, no. 19, pp. 10827–10912, 2021.

[28] C. Peng, B. Qiu, M. Li et al., "Gaussian diffusion sinogram inpainting for X-ray CT metal artifact reduction," *BioMedical Engineering Online*, vol. 16, no. 1, p. 1, 2017.

[29] H. Ca I, V. W. Zheng, and C. C. Chang, "A comprehensive survey of graph embedding: problems, techniques and applications," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, 2017.

[30] S. A. Fadaee and M. Amir Haeri, "Classification using link prediction," *Neurocomputing*, vol. 359, pp. 395–407, 2019.

[31] A. Aleta et al., *Link Prediction in Multiplex Networks via Triadic Closure*, American Physical Society, College Park, MD, USA, 2020.

[32] Y. Yoon, J. Hong, and W. Kim, "Item recommendation by predicting bipartite network embedding of user preference," *Expert Systems with Applications*, vol. 151, no. 3, Article ID 113339, 2020.

[33] M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, https://arxiv.org/abs/1609.02907.

[34] A. Grover and J. Leskovec, "node2vec: scalable feature learning for networks," *KDD: Proceedings. International Conference on Knowledge Discovery & Data Mining*, vol. 2016, pp. 855–864, 2016.

[35] E. M. Kamel, "The MENA region's need for more democracy and less bureaucracy: a gravity model controlling for aspects of governance and trade freedom in MENA," *The World Economy*, vol. 44, no. 6, pp. 1885–1912, 2021.

[36] O. Koch, "Scalable representation learning and retrieval for display advertising," 2021, https://arxiv.org/abs/2101.00870.

[37] D. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, https://arxiv.org/abs/1412.6980.

[38] D. Chikmurge and R. Shriram, "Marathi handwritten character recognition using SVM and KNN classifier," *Hybrid Intelligent Systems*, vol. 32, pp. 319–327, 2021.

[39] C. Baum, M. E. Schaffer, and S. Stillman, "IVREG2: stata module for extended instrumental variables/2SLS and GMM estimation," *Statistical Software Components*, vol. 8, 2019.

[40] Y. Xiong, Y. Zhu, and P. S. Yu, "Top-k similarity join in heterogeneous information networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 6, pp. 1710–1723, 2015.

[41] J. Qiu, "Network embedding as matrix factorization: unifying DeepWalk, LINE, PTE, and node2vec," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, Como, Italy, August 2017.

[42] J. Oh, K. Cho, and J. Bruna, *Advancing GraphSAGE with A Data-Driven Node Sampling*, Springer, Berlin, Germany, 2019.