International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Fast optimization for large scale logistics in complex urban systems using the hybrid sparrow search algorithm

Yao Yao, Siqi Lei, Zijin Guo, Yuanyuan Li, Shuliang Ren, Zhihang Liu, Qingfeng Guan & Peng Luo

To cite this article: Yao Yao, Siqi Lei, Zijin Guo, Yuanyuan Li, Shuliang Ren, Zhihang Liu, Qingfeng Guan & Peng Luo (2023): Fast optimization for large scale logistics in complex urban systems using the hybrid sparrow search algorithm, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2023.2190371

To link to this article: https://doi.org/10.1080/13658816.2023.2190371

Published online: 24 Mar 2023.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2023.2190371

RESEARCH ARTICLE

Fast optimization for large scale logistics in complex urban systems using the hybrid sparrow search algorithm

Yao Yao[a,b], Siqi Lei[a], Zijin Guo[a], Yuanyuan Li[a], Shuliang Ren[a], Zhihang Liu[c], Qingfeng Guan[a] and Peng Luo[d]

[a]School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei province, P. R. China; [b]Center for Spatial Information Science, The University of Tokyo, Chiba, Japan; [c]School of Urban Planning & Design, Peking University Shenzhen Graduate School, Shenzhen, Guangdong province, P. R. China; [d]Cartography and Visual Analytics, Technical University of Munich, Munich, Germany

ABSTRACT

Urban logistics is vital to the development and operation of cities, and its optimization is highly beneficial to economic growth. The increasing customer needs and the complexity of urban systems are two challenges for current logistics optimization. However, little research considers both, failing to balance efficiency and cost. In this study, we propose a hybrid sparrow search algorithm (SA-SSA) by combining the sparrow search algorithm with fast computational speed and the simulated annealing algorithm with the ability to get the global optimum solution. Wuhan city was selected for logistics optimization experiments. The results show that the SA-SSA can optimize large-scale urban logistics with guaranteed efficiency and solution quality. Compared with simulated annealing, sparrow search, and genetic algorithm, the cost of SA-SSA was reduced by 17.12, 18.62, and 14.72%, respectively. Although the cost of SS-SSA was 11.50% higher than the ant colony algorithm, its computation time was reduced by 99.06%. In addition, the simulation experiments were conducted to explore the impact of spatial elements on the algorithm performance. The SA-SSA can provide high-quality solutions with high efficiency, considering the constraints of many customers and complex road networks. It can support realizing the scientific scheduling of distribution vehicles by logistics enterprises.

ARTICLE HISTORY
Received 4 March 2022
Accepted 8 March 2023

KEYWORDS
Large-scale urban logistics; sparrow search algorithm; multiple depot vehicle routing problem; complex road network

1. Introduction

Large-scale urban logistics and distribution, such as garbage collection (Gemma *et al*. 2016), and courier services (Martin *et al*. 2021), refer to the delivery of large volumes of goods that serve the large metropolitan area. It relies on the city's complex road network and is essential for people's lives (Kim *et al*. 2015, Groß *et al*. 2020). With the rapid expansion of cities and population, especially in developing countries such as China, the role of the logistics industry in economic development is increasingly

CONTACT Peng Luo peng.luo@tum.de

essential (Cattaruzza *et al.* 2017, Tu *et al.* 2017). However, the fast-growing urban population is increasing the demand for logistics services in recent years, logistics efficiency has been challenged, which also burdens urban transport networks more (Bergmann *et al.* 2020). Optimizing large-scale logistics distribution routes and reducing costs under the existing urban transportation network has become an essential task.

The growing demand for distribution and the highly complex road network in large cities are the main challenges of city logistics. First, large-scale urban logistics tasks require service to thousands of customers with heavy distribution tasks and high time pressure (Laporte 2009). Therefore, efficient heuristic algorithms are needed to obtain high-quality vehicle path solutions (Wang *et al.* 2021). Second, the urban traffic network is a complex spatial network that contains topological and spatial information about the relevant road sections (Yao *et al.* 2018). However, current research rarely considers both the demand for large-scale urban distribution and highly complex road networks, leading to a poor balance between optimization performance and computational efficiency (Hill and Benton 1992, Janic 2007).

From the perspective of Geographical Information Science (GIS), logistics optimization is a pathfinding problem under complex spatial constraints. The carrier of logistics distribution is a traffic network constrained by node complexity, road features, network efficiency and other spatial elements (Nasiri 2014, Shen 2020, Hina *et al.* 2020). From the individual scale of customer points, in a complex urban network, the accessibility of customers relative to the warehouse will affect the efficiency of logistics distribution. From the local scale, the complexity and connectivity of the road network in a specific area will affect the logistics distribution. In addition, the urban structure of mega-cities often has considerable spatial heterogeneity, which also inevitably affects the performance of logistics distribution tasks in regions. These spatial characteristics present challenges to logistics optimization algorithms.

Scholars in the GIS domain have developed a series of pathfinding algorithms for logistics optimization based on the shortest path problem (Li *et al.* 2015, Liu *et al.* 2018, Shirabe 2014, Vanhove and Fack 2012, Zeng and Church 2009). Currently, most logistics optimization studies consider 20–100 customer points, and some scholars define the logistics task with 200–500 customers as a large-scale optimization problem (Zachariadis and Kiranoudis 2010, Zhou *et al.* 2018). With the development of the city, the logistics needs of the city far exceed the size of 500 customers, and some studies indicate that urban logistics optimization is more practical with more than 1,000 customers (Dong *et al.* 2021). In addition, most existing research is based on simulated data through Euclidean distances to evaluate logistics optimization algorithms. In practice, the urban traffic network is a carrier for logistics transport. The traffic network's spatial structure transformation and speed limits can significantly impact logistics transport. Therefore, to meet urban logistics needs, conducting effective and efficient optimization to consider both factors is essential in current urban logistics.

Large-scale urban logistics usually transfer and distribute goods from multiple warehouses (Kim *et al.* 2015) to meet customers' needs at different locations. Therefore, logistics distribution is usually modeled as a multi- depot vehicle routing problem (MDVRP) (Cattaruzza *et al.* 2017). In this study, we choose the most common logistics distribution optimization scenario in cities: multi- depot, delivery vehicles with capacity constraints, and delivery vehicles return to the warehouse after the distribution tasks. The delivery vehicles are required to traverse all customer points, and the optimization objective is the minimum driving distance. Such scenarios include express delivery, garbage collection, etc. In the multi-warehouse problem, there are more complex space constraints. First, a reasonable distribution range for each warehouse has to be considered, and second, the distribution path from the warehouse has to be optimized. We proposed a hybrid sparrow search algorithm to solve large-scale urban logistics optimization problems by combining a sparrow search algorithm (SSA) and a simulated annealing algorithm (SA). The SSA aimed to obtain the solution efficiently. To avoid fall into local optimum, the SA was used to improve the quality of the solution. The proposed algorithm was applied to the logistics optimizaiton in Wuhan city and was compared with the other classical heuristic algorithms.

The remainder of this paper is structured as follows: Section 2 introduces the related works of logistics optimization. Section 3 describes the study area and data used in this study. Section 4 covers the definition and derivation of the proposed SA-SSA. Section 5 presents the results of logistics optimization in the study area. Findings and contributions are discussed in Section 6, and conclusions are summarized in Section 7.

2. Related work

2.1. Logistics optimization problem definition and *topological representation*

Logistics optimization, as a classical pathfinding task under spatial constraints, is one of the essential applications of GIS. It refers to the supply of goods from one or more warehouses to serve multiple spatially discrete customers with different needs. It requires quickly choosing the optimal distribution path for many clients, which can boost logistics companies' economic efficiency and customer satisfaction. Spatial elements are the main components of the optimization objective and problem construction. The optimization problem's basic elements, such as warehouses, vehicles, and customers, can be modelled as spatial objects, and the distribution routes are based on the urban road network.

According to the number of warehouses in the vehicle path problem, the logistics optimization problem can be divided into single-depot vehicle routing problems and multi-depot vehicle routing problems (Han and Wang 2018, Gayialis *et al*. 2019, Wang *et al*. 2020). The single-depot vehicle routing problem includes only one warehouse, and all vehicles start from the same warehouse to serve the customer point (Chan *et al*. 2002, Rahim *et al*. 2016). Urban logistics path optimization requires the consideration of multiple warehouses and many customers. The multi-depot vehicle routing problem is more in line with application scenarios, such as logistics distribution and garbage collection (Shen *et al*. 2021, Nozari *et al*. 2022). Multiple warehouses serve the customer point at the same time. Therefore, the distribution vehicles need to choose the departure warehouse reasonably for efficient service of the customer point. However, since MDVRP is an NP-hard problem (Laporte 2009), it is challenging to obtain high-quality solutions, and it is challenging to construct an efficient algorithm.

In the field of GIS, the topological representation of road networks is an essential topic for logistics optimization (Liu *et al.* 2018). The most commonly used method describes the road network and customers as a graph consisting of nodes and edges. Some studies express streets as a node of the graph, and two streets are considered to have connected edges between them if they have proximity (Jiang and Claramunt 2004, Jiang and Liu 2009). Such street-street topologies take into account the complexity of the topology. However, this approach does not provide the most realistic description of the logistics distribution process in a traffic network, where drivers' wayfinding is usually based on road junctions or customer locations. Therefore, in GIS studies, the optimal road network representation for path planning and logistics optimization is expressing road junctions or customer points as nodes and edges as road segments between two nodes. It can maximize the topological information of the road network and thus can realistically simulate the logistics planning in Spatial constraints.

2.2. Logistics optimization algorithm

A range of urban logistics optimization algorithms has been developed to solve large-scale urban logistics optimization problems. Common methods applied in vehicle route planning can be classified into exact and approximate solution methods according to the strategy of solving. Exact algorithms were widely used in early logistics optimization studies (Eilon *et al.* 1974, Toth and Vigo 2002, Shimizu *et al.* 2020). For example, Laporte and Nobert proposed branch delimitation algorithms to serve 25 customer points for delivery (Laporte 1984). The exact algorithm continued to evolve, and the number of customers that could be satisfied was expanded to 80 (Bettinelli *et al.* 2011, Laporte *et al.* 1988). However, the computation time also increases exponentially with the increase in distribution size (Salman *et al.* 2020). In urban logistics tasks with thousands of customers (Crainic *et al.* 2009), the exact algorithm cannot find the optimal solution in a limited time. Hence, it failed to meet the increasing need for urban logistics (Toth and Vigo 2002, Tu *et al.* 2017). The exact methods can get the full optimal solution, but it is difficult to apply in practical scenarios (Dasdemir *et al.* 2022).

Approximate solution methods are developed to conduct the logistics optimization task in real-world problems. They include reinforcement learning and heuristic algorithms. In reinforcement learning-based logistics and distribution optimization, the model is trained with a reward mechanism that enables high-quality urban logistics and distribution solutions (Li *et al.* 2017, Lu and Gzara 2019). For example, the Q-learning algorithm (Watkins and Dayan 1992) and the DQN algorithm (Mnih *et al.* 2013) optimize urban logistics. Reinforcement learning has the ability of autonomous learning and decision-making and can handle the uncertainty of environmental changes in the face of complex, large scale vehicle path planning scenarios (Arulkumaran *et al.* 2017, Phiboonbanakit *et al.* 2021). However, reinforcement learning-based methods are time-consuming, and the model is challenging to be converged. For urban logistics optimization problems with complex road networks and many user requirements, reinforcement learning still lacks the capability for large-scale engineering applications. In addition, the exploration process of reinforcement learning has an unstable scenario, which can lead to a very time-consuming convergence and does not apply to the solution of real VRP problems.

Compared to exact algorithms and reinforcement learning, heuristic algorithms have a proven theoretical basis and can converge quickly to obtain high-quality solutions in an efficient time. The proposal of heuristic algorithms offers opportunities for the rapid implementation of large-scale logistics route optimization (Mester *et al.* 2007, Wang and Chen 2012, Mouthuy *et al.* 2015).

Heuristics models can be classified as local search methods and intelligent search methods (Tu *et al.* 2015). Local search heuristics use a spatial domain structure by traversing the domain of the current solution and transferring to a better domain solution until the termination condition is satisfied. The simulated annealing algorithm (SA) (Hwang 1988, Wang *et al.* 2022), a representative of local search heuristics, simulates the principle of annealing of solid combustibles and has a complete theoretical basis. As a result, the algorithm is widely used in urban logistics optimization. However, the SA strongly depends on the quality of the initial solution, and its urban logistics optimization results are difficult to converge. To sum up, more efficient optimization methods must be explored to improve the efficiency of heuristic algorithms in large-scale (e.g., 1000 customers) urban logistics distribution.

Intelligent search heuristic algorithms simulate the behavior of organisms in nature and design intelligent search rules to improve the quality of solutions using group search continuously. Among them, ant colony algorithms (ACO) (Ouyang and Zhou 2011, Qin *et al.* 2021) and genetic algorithms (GA) (Imani and Ghoreishi 2021, Mester and Bräysy 2005) are representatives of intelligent search heuristics. ACO are based on the idea of ants searching for food and determining the shortest path by staying on the pheromone on the path. For example, Lin et al. used an ACO to optimize the logistics distribution of a tram (electri vehicle) to ensure that the tram had sufficient power to complete its service to customers (Shi *et al.* 2022). GA mimic biological evolution in nature and optimize urban logistics through mechanisms such as selection and mutation. For example, Abolfazl et al. used genetic algorithms to optimize the distribution of medical items in the city of Tehran (Aliakbari *et al.* 2022). The study shows that swarm intelligence optimization algorithms can ensure the quality and efficiency of logistics optimization in urban logistics. However, as the size of urban customers increases, the optimization quality and efficiency of ACO and GA show a significant decline, making it difficult to meet the actual needs of existing cities for logistics optimization.

As a new intelligent search heuristic algorithm, the SSA provides opportunities for solving urban logistics distribution path optimization (Xue and Shen 2020). The algorithm simulates the foraging and warning behaviors of sparrows. It has the characteristics of fast convergence, strong optimization ability, and short operation time compared with the classical heuristic search algorithm (Ouyang *et al.* 2021). Due to its few control parameters and easy implementation, SSA has been applied to practical engineering applications such as power management of hybrid renewable energy sources (HRES), sustainable energy system optimization, and unmanned aerial vehicle route planning (Kumaravel and Ponnusamy 2020, Liu and Rodriguez 2021, Liu *et al.* 2021). However, the SSA is updated by approaching the optimal forward position and approaching the origin, which leads the algorithm to fall into local optimal solution.

Whether the SSA can be applied to large-scale urban logistics path optimization is unknown. Improving the SSA while considering route optimization and computational time to solve urban logistics path optimization problems is of great significance.

3. Study area and data description

Wuhan has about 8,500 square kilometers and more than 11 million population by the end of 2019. The built-up area of Wuhan city has a ring and radial road system, with the urban motorway, trunk road, arterial road, secondary road, and tertiary road as the main skeleton. According to the Wuhan City Open Data Development Platform (http://english.wuhan.gov.cn/), the total mileage of roads in Wuhan has reached 12,000 kilometers, with a road density of 149.5 kilometers per 100 square kilometers. Because of the unique geographical environment, Wuhan has many bridges across the Yangtze River, increasing the road network complexity. According to the Wuhan City Open Data Development Platform (http://english.wuhan.gov.cn/), Wuhan's Gross Domestic Product (GDP) exceeds RMB 150 billion. The logistics industry, an essential part of Wuhan's tertiary industry, completed 636 million tons of freight in 2020. Among them, the volume of road cargo transportation has reached 317 million tons, topping all modes of transportation. Due to the complex road network and massive cargo transportation demand, Wuhan city is an ideal place to study the optimization of urban logistics distribution paths.

Logistics data are the main datasets used in this study (Figure 1). We obtained a logistics company's customer location and warehouse data from Gaode Map

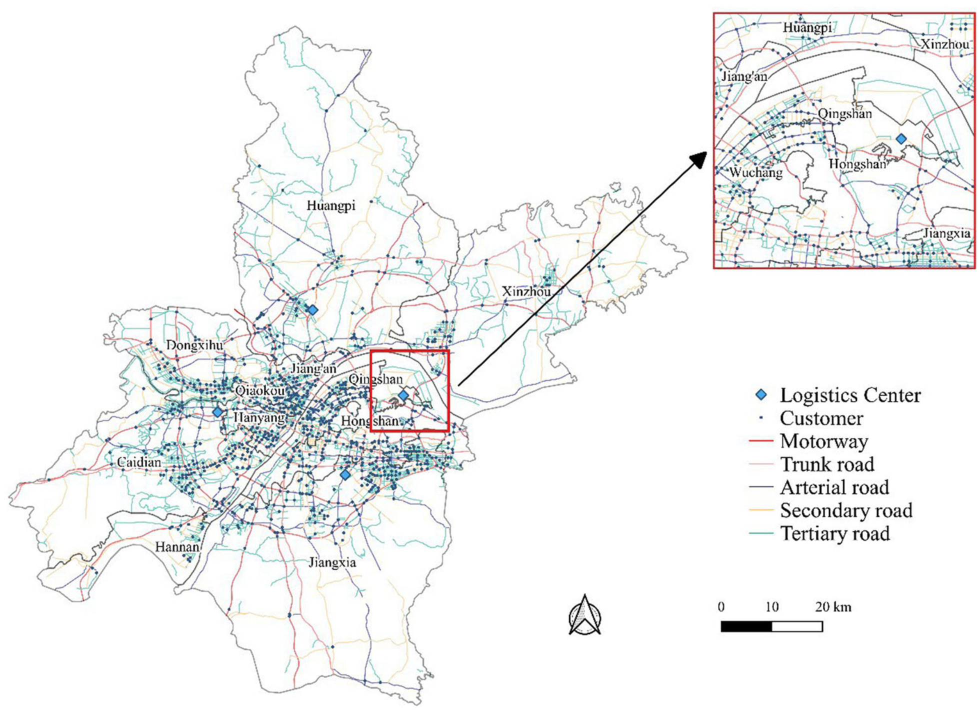

Figure 1. Logistics centres, customers and road network data in the study areas.

Table 1. Location coordinates of demand points and delivery quantity.

Sequence	Longitude	Latitude	Delivery quantity/kg
1	114.06231	30.40521	15
2	114.13948	30.48460	30
3	114.38444	30.46892	45
4	114.30151	30.46185	30
5	114.39554	30.38778	25
6	114.25430	30.58867	30
…	…	…	…

(https://lbs.amap.com/). One thousand customer points and four warehouses were randomly selected. In addition to the location information, the customer points also contain the delivery quantity data shown in Table 1, which is a random value in the 1–100kg interval. In addition, we obtained vehicle condition data from the logistics company, including the maximum load of the vehicles.

The road network data was obtained from OpenStreetMap (OSM) (http://www.openstreetmap.org). OSM is an open-source mapping website that provides free and easily accessible digital map data (Pourabdollah *et al.* 2013). The quality of data is crucial to logistics distribution which ensures the accuracy of the distribution route for the logistics model. The OSM data provides spatial data with high quality by evaluating its positional and attribute accuracy, completeness, and consistency (Haklay 2010). The accuracy and reliability of OSM data for characterizing Wuhan's urban transportation network have been verified (Wang *et al.* 2013). Wuhan city's road network data contains 81,711 road segments and related attributes, such as road latitude and longitude and road type descriptions. In this study, the roads were classified into four classes based on the type description information from OSM data: motorway, trunk road, arterial road, secondary road, and tertiary road (Figure 1).

4. Methodology

In this study, a hybrid heuristic algorithm was proposed, combing a SSA and a simulated annealing algorithm. We constructed an urban logistics optimization scenario and researched logistics route optimization using road network and logistics data. This study consists of four main parts (Figure 2): (1) Scene modeling for multi-depot logistics optimization. (2) A SA-SSA was used to conduct the path optimization to solve the multi-depot vehicle routing problem. (3) The optimization results were compared and analyzed with other classical heuristic algorithms to verify the effectiveness and reliability of the SA-SSA. (4) Three sets of simulation experiments at different spatial scales were conducted to explore the impact of spatial elements on logistics optimization algorithms.

4.1. Scene modeling for multi-depot logistics optimization

In the MDVRP, logistics centers, customers, and transportation networks can be represented by a weighted graph $G(N, E, W)$. The vertex $N = \{n_1, n_2, n_3 \ldots, n_L\}$ is defined as a set of logistics centers $\{n_1, n_2, n_3 \ldots, n_l\}$ and customers $\{n_{l+1}, n_{l+2}, n_{l+3} \ldots, n_L\}$. The $E = \{e_1, e_2, e_3 \ldots, e_r\}$ is the set of roads in the traffic network. The $W = \{w_{1\cdot 1}, w_{1\cdot 2}, w_{1\cdot 3} \ldots, w_{i\cdot j}\}$ is represent the shortest weight. The $w_{i\cdot j}$ is the set of roads in n_i

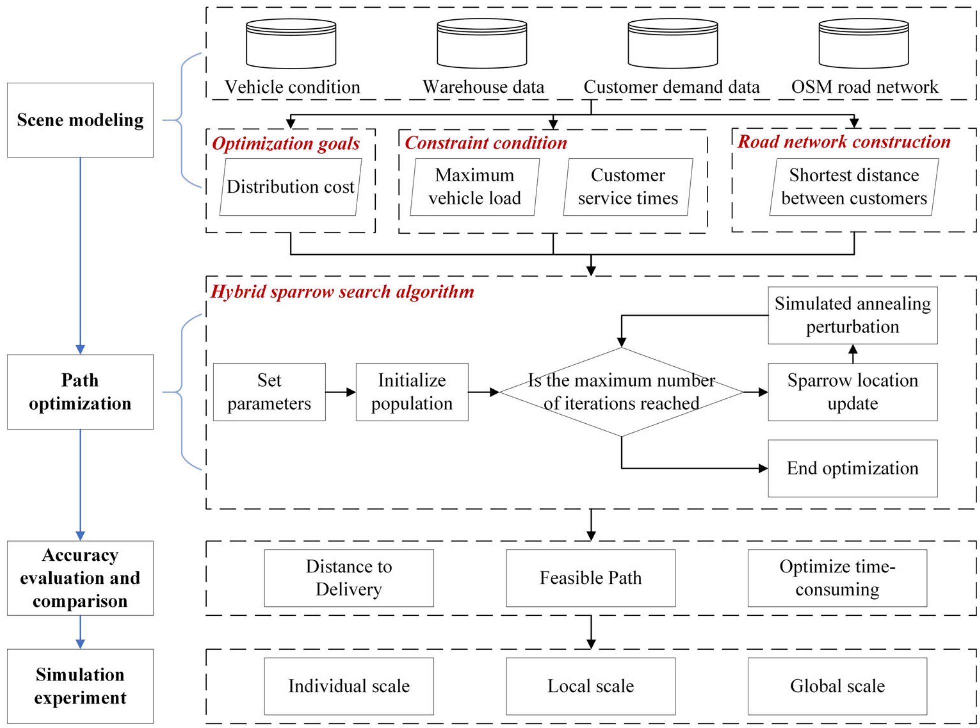

Figure 2. The workflow of multi-depot urban logistics optimization.

and n_j. In the paper, the roads in the Wuhan city are divided into five categories: motorway, trunk road, secondary road, tertiary road, and arterial road, referring to the description of data, and the Urban Road Engineering Design Specification (2016 version). The average speed S' of each road class was regarded as the road attribute. The travel time ($t_r = \frac{d_r}{s_r}$) of each road was calculated by the distance and speed of the road. The shortest path $w_{i \cdot j}$ is calculated by the travel time.

To clarify urban logistics distribution, we further explain the MDVRP model. The logistics vehicles are all the same type and have a maximum capacity C and travel distance D in the paper. The logistics vehicle loads goods and depart in the logistics warehouse, then the vehicle delivers the goods in the city road network. The Formula (1) specify the logistics optimization:

$$\text{Min } F(x) = \sum_{i \in N} \sum_{j \in N} \sum_{k \in K} w_{i \cdot j \cdot k} x_{i \cdot j \cdot k} \;\; i \neq j \tag{1}$$

Where the K reprensets the set of vehicle, the $w_{i \cdot j \cdot k}$ reprensets the cost from i *and* j by vehicle k. If vehicle k departs from point i and arrive to j, the $x_{i \cdot j \cdot k} = 1$. Otherwise $x_{i \cdot j \cdot k} = 0$.

The constraint conditions are shown in Formula (2)–Formula (6):

- A customer point can only be served once by one vehicle:

$$\sum_{j \in N} \sum_{k \in K} x_{i \cdot j \cdot k} = 1, \;\; i \neq j, \;\; j \epsilon \{n_{l+1}, n_{l+2}, n_{l+3} \ldots, n_L\} \tag{2}$$

- The vehicle departures and retures from the same logistics warehouse:

$$\sum_{i\in N}\sum_{a\in N} x_{i\cdot a\cdot k} - \sum_{j\in N}\sum_{b\in N} x_{i\cdot j\cdot k} = 0, \quad i \neq j, \ i \ j\epsilon\{n_1, \ldots, n_l\} \tag{3}$$

- The total demands (q) of customers does not exceed the maximum load capacity C :

$$\sum_{i\in N} q_i x_{i\cdot k} \leq C \tag{4}$$

  Where the q_i reprensets that quantity demanded for customer i.
- The total vehicle distribution path distance does not exceed the maximum vehicle travel distance D :

$$\sum_{i\in N}\sum_{j\in N} d_i x_{i\cdot k} \leq D, \ i \neq j \tag{5}$$

  Where the $d_{i\cdot j}$ reprensets that travel distance from customer i to customer j.
- Each vehicle can only complete one delivery service. In this study, the delivery service refers to the task that starts from one warehouse to complete a full distribution and then returns to the same warehouse:

$$\sum_{k\in K} x_{i\cdot k} = 1 \tag{6}$$

4.2. Hybrid sparrow search algorithm for large-scale logistics optimization

In order to solve the large-scale urban logistics distribution problem, this paper proposes a hybrid sparrow search algorithm by combining the sparrow search algorithm and the simulated annealing algorithm. By simulating the foraging behavior of a sparrow population, the SSA algorithm classifies sparrows into three types of intelligence: discoverers, followers, and scouters. In urban logistics optimization, the discoverer provides an optimization direction for urban logistics optimization, while the follower continues to explore the global optimum in the optimization direction. When the discoverer and the follower fall into a local optimum in that optimization direction, it is difficult to improve the urban logistics optimization solution. The early-warning scouter sends a timely message to guide the discoverer to open a new optimization direction to explore the global optimal solution.

Compared to other population intelligence optimization algorithms, the discoverer-follower-scouter mechanism of the SSA improves the local search capability and has significant advantages in terms of convergence speed, search accuracy and stability. However, the above features of the sparrow search algorithm also reduce its global exploration capability. Therefore, this paper introduces the SA into the SSA, expecting to borrow the probabilistic sudden jump idea of SA to accept poor solutions with a certain probability, thus improving the global exploration ability.

The process is divided into two steps in the hybrid sparrow search algorithm for logistics distribution paths (Table 2): First, the SSA optimizes the logistics paths and obtains the initial optimal solution. Second, the initial optimal solution obtained by the SSA is perturbed by the SA. This approach is expected to help the algorithm jump

Table 2. SA-SSA algorithm operations.

The hybrid sparrow search algorithm
Step 1: Initialize corresponding parameters: #Set the number of maximum iterations, the number of producers, the number of sparrows who perceive the danger, the alarm value, and the number of sparrows. #Set the Correlation coefficient of the simulated annealing algorithm #Rmax: the maximum iterations, PD: the number of producers, SD: the number of sparrows who perceive the danger, and R2: the alarm value. **Step 2:** Randomly generate the initial solution. for $i = 1$: Rmax **Step 3:** Update discover location by Formula (8); Update follower location by Formula (9) **Step 4:** Randomly select the scouter and update the location by Formula (10) **Step 5:** Calculate the objective function and select the best individual for perturbation operation. **Step 6:** Increase perturbation for the best individual. #Simulated annealing algorithm for $j = 1$:Gmax **Step 6:** Creating neighbourhood solution: Domain solutions are created based on exchange operator, reversal operator and insertion operator. **Step 7:** Accepting the critical domain solution by Formula (11) **Step 8:** Update the best individual. **Step 9:** Find the optimal solutions. End

out of the local optimum to get efficient and quality logistics distribution path planning.

4.2.1. Initial solution solving based on the sparrow search algorithm

The logistics optimization problem can be understood as a rational ordering of the access order of customer points. In the logistics optimization algorithm based on the sparrow search algorithm proposed in this study, the access weights of K sparrows to all distribution points form a matrix, and the distribution vehicles will serve the distribution points in the order of weight. Therefore, logistics optimization aims to obtain a reasonable access weight matrix for proper distribution efficiency and quality. The access weight matrix is as follows:

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,j} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,j} & \cdots & x_{2,d} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{i,1} & x_{i,2} & \cdots & x_{i,j} & \cdots & x_{i,d} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{K,1} & x_{K,2} & \cdots & x_{K,j} & \cdots & x_{K,d} \end{bmatrix} \tag{7}$$

where d denotes the dimensionality of the problem variable to be optimized, i.e. the logistics distribution points. x_{i*} denotes the solution of the i th intelligence, i.e. the city logistics distribution solution. $x_{i,j}$ denotes the weight of the i th intelligence at the j th customer point, the higher the weight, the higher the probability of access. The K intelligences are classified according to the optimization objective: discoverers, followers and early warners. The dimensional characteristics of each type of intelligent body are updated according to the following formula.

The weight update formula for the discoverer is:

$$x_{i,d}^{t+1} = \begin{cases} x_{i,d}^{t} * \exp\left(\frac{-i}{\alpha * iter_{\max}}\right), R_2 < ST \\ x_{i,d}^{t} + Q, R_2 \geq ST \end{cases} \tag{8}$$

where $x_{i,d}^{t+1}$ is the d-dimensional position of the i-th individual in the t-th iteration, α is a random number in the interval (0, 1], and Q is a random number that conforms to the standard normal distribution. R_2 is a random number in the interval [0,1], and ST is a warning value in the interval [0.5,1.0].

For followers, the position update formula is as follows.:

$$x_{i,d}^{t+1} = \begin{cases} Q * \exp\left(\frac{xw_{i,d}^{t} - x_{i,d}^{t}}{\alpha * iter_{\max}}\right), i \geq n/2 \\ xb_{i,d}^{t} + \frac{1}{D}\sum_{d=1}^{D}(rand\{-1,1\} * (|xb_{i,d}^{t} - x_{i,d}^{t}|)), R_2 < ST \end{cases} \tag{9}$$

where xw is the worst position of the sparrow in the current population and xb is the best position of the sparrow in the current population.

The equation for updating the location of the scouters is as follows.

$$x_{i,d}^{t+1} = \begin{cases} xb_{i,d}^{t} + \beta * (x_{i,d}^{t} - xb_{i,d}^{t}), f_i \neq f_g \\ x_{i,d}^{t} + K\left(\frac{x_{i,d}^{t} - xw_{i,d}^{t}}{|f_i - f_W| + \varepsilon}\right), f_i = f_g \end{cases} \tag{10}$$

where β is the random number conforming to the standard normal distribution, K is the random number in the interval [−1,1], ε is the smaller number, and f_W is the fitness value of the worst-position sparrow.

4.2.2. Optimal solution update based on simulated annealing algorithm

In order to avoid the sparrow search algorithm falling into a local optimum and improve the optimization quality for large-scale urban logistics distribution, this paper introduces a simulated annealing algorithm. The idea of this algorithm is inspired by the principle of simulated annealing of solid combustibles, which is essentially a probability-based stochastic algorithm. Therefore, the simulated annealing algorithm can enable the sparrow search algorithm to jump out of the local optimum with a certain probability.

In the simulated annealing algorithm, $F(x)$ is the optimization objective, x_{best} is the current optimal solution obtained by the sparrow search algorithm. During the iterative solution process, the algorithm perturbs the current solution x_{old} randomly, resulting in a new solution x_{new}. if $F_1(x_{old}) \geq F_1(x_{new})$, then the new solution x_new will overwrite the old solution x_{old}. Otherwise, the poor solution x_{new} is chosen to be accepted with a certain probability p, which is defined by the Metropolis criterion:

$$p = pexp\left(-\frac{F_1(x_{new}) - F_1(x_{old})}{pkT}\right) \tag{11}$$

where k is indicated as Boltzmann, in this study, indicates the number of iterations. T is the current temperature value, which will gradually decrease during the optimizaiton process. Its temperature difference represents the difference between the new and old solutions. In this study, T controls the rate at which the model converges.

4.3. Model evaluation

In order to demonstrate the effectiveness of the proposed SA-SSA, a series of model validations and comparisons were carried out. First, we carried out logistics optimization experiments in the study area using SA-SSA, calculated the total distance and time of the optimized paths, and compared them with each other and with classical algorithms. The four selected algorithms are SA, SSA, GA, and ACO. In addition, we selected several typical paths and analyzed the rationality of the optimization results of the four methods.

Second, we analyze the rationality of the spatial distribution of the distribution paths in two different scenarios. The spatial statistics of the logistics distribution paths obtained using the SA-SSA were carried out in two cases, considering road class and not considering road class. Third, a sensitivity analysis of the parameters of the SA-SSA is carried out to validate the model's robustness.

4.4. Simulation experiment

Logistics optimization is an optimization problem under complex spatial constraints. It is important to explore how spatial characteristics affect algorithm performance to understand the mechanism of optimization algorithms. It also helps us to use GIS for logistics optimization under more complex spatial constraints in further research. In this study, we compute spatial features at three different scales for customer points. By grouping the logistics customer points according to the spatial features and then optimizing them separately using the proposed SA-SSA, we explore how the spatial features affect the performance of the algorithm.

4.4.1. Individual-scale simulation

The individual-scale simulation experiments are conducted to analyze the performance of logistics optimization algorithms from the perspective of spatial characteristics of an individual customer point. The network efficiency (NE) from the warehouse to the customer point is important for logistics planning tasks. It is the reciprocal of the shortest path distance between nodes (Latora and Marchiori 2001), which characterizes how easy it is for a warehouse to serve a customer point. In this study, we calculate the average network efficiency of each customer point to four warehouses and analyze the performance of SA-SSA for serving customer groups with different levels of network efficiency. The average network efficiency for a customer point is calculated as follows:

$$\mathrm{NE} = \sum_{1}^{m} \frac{1}{SD_i} \tag{12}$$

Where SD_i is the shortest distance from customer point to warehouse i, and m is the number of warehouses.

We calculated the average network efficiency of 1000 customer points and then divided them equally into two subgroups (low network efficiency group and high network efficiency group) according to the value of network efficiency. After that, for each subgroup, low-density (100 points) and high-density (300 points) sampling was performed twice, and then logistics optimization was performed using SA, SSA, and SA-SSA.

4.4.2. Local-scale simulation experiment

The local-scale simulation experiment evaluates the optimization algorithm from the local spatial structure of the location of the customer points. The logistics optimization task is determining the order of access to each customer point. The ease with which a customer point can be served depends not only on its location from the warehouse but also on its ease of access to other customer points. If a customer point is relatively close to all other customer points, it is more convenient to be served by a logistics vehicle. In this way, the overall distance of logistics is shorter.

In this study, we choose closeness centrality to characterize the proximity between other customer points of a customer point in the urban road network (Okamoto *et al.* 2008). For a specific node, closeness centrality is calculated as the average length of the shortest path from this node to all other nodes. Similar to the simulation experiment of network efficiency, we calculated the closeness centrality of 1000 customer points and then divided them equally into two subgroups. After that, for each grouping, low-density (100 points) and high-density (300 points) sampling was performed twice, and then logistics optimization was performed using SA, SSA, and SA-SSA.

4.4.3. Global-scale simulation

The global scale simulation experiment evaluates the algorithm from the perspective of urban structure. The urban road network and customers have a spatially heterogeneous pattern. For example, customers in some areas show spatial aggregation, while others are sparse. In addition, the spatial structure of the urban road network affects the service capability and efficiency of logistics in different areas.

To explore the performance of SA-SSA in serving regions with different spatial characteristics, we conducted a network-based spectral clustering of the 1000 customer points in the study area to obtain clusters with different spatial patterns (Ng *et al.* 2001). Then logistics optimization experiments are conducted separately for customer points within each cluster to evaluate the algorithm's performance.

5. Result

5.1. Model evaluation result

To verify the effectiveness of the proposed hybrid algorithm, we constructed an urban logistics distribution scenario based on four logistics warehouses, 1000 customer points, and road network data in Wuhan city. The proposed SA-SSA, SA, SSA, GA, and ACO were used for logistics path optimization. We conducted a comparative analysis of the optimization results of each algorithm. Each algorithm was run 30 times, and the average total path length, standard deviation, and calculation time were recorded

Table 3. Comparison of results of different optimization algorithms: simulated annealing (SA), sparrow search algorithm (SSA), genetic algorithm (GA), ant colony algorithm (ACO), and the hybrid sparrow search algorithm (SA-SSA).

Algorithms	Optimal solution/km	Operation time/s
SA	5317.845 ± 108.51 (+17.12%)	936.14 ± 14.34 (+14.05%)
SSA	5415.927 ± 110.88 (+18.62%)	749.52 ± 11.06 (−7.35%)
GA	5168.396 ± 106.25 (+14.72%)	793.76 ± 15.56 (−1.37%)
ACO	3952.983 ± 97.36 (−11.50%)	85291.38 ± 469.65 (+99.06%)
Proposed SA-SSA	4407.619 ± 91.74	804.61 ± 12.73

Values in parentheses represent relative differences compared to SA-SSA.

for model comparison and sensitivity analysis. After several experiments and comparisons of the optimization result, the maximum number of iterations $iter_{max}$ of the SSA-SA algorithm was set to 1000, the sparrow population K was 50, and the warning value ST was 0.6. In addition, the proportion of discoverer PD was 0.7, the initial temperature T was 100, and the cooling speed *alpha* was 0.97.

Table 3 shows the optimization results for five algorithms. The proposed SSA-SA algorithm had the highest solution quality regarding the shortest path for logistics distribution. Compared with the SSA and GA, the minimum shipping distance of the SSA-SA algorithm was reduced by 17.12 and 14.72%, respectively. It is noteworthy that the minimum shipping distance of the SSA-SA algorithm was decreased by 18.62% compared to the original SSA. This indicates that the introduction of SA improves the global search capability of the SSA algorithm. In addition, the shortest shipping distance of the ACO outperformed the hybrid SSA by 11.50% compared to the SSA-SA algorithm due to its strong global search capability.

The proposed SSA-SA algorithm had high computational efficiency, optimizing logistics routes for 1000 customer instances in 804.61s on average. Its computation time was not significantly different from the SSA, SA and GA. However, it had a considerably shorter computation time (99.06%) than the ACO. In addition, the SSA-SA algorithm had the smallest standard deviation (91.74 km) of the average minimum shipping distance and the smaller standard deviation of the average optimization time (Table 3). This indicates that SSA-SA had better optimization performance and solving stability than other algorithms. In summary, the proposed algorithm can consider both shipping cost and computation time in the real logistics scenarios, better meeting the logistics need of large urban with complex systems.

To better demonstrate the advantages of SSA-SA, we have selected several typical distribution paths for case analysis. Each complete distribution route starts from a warehouse, serves some customer points, and returns to the same warehouse. The paths planned by SSA-SA tend to be high-speed roads for delivery to more dense customer points in the city centre. For example, the vehicle selects the Dai Huang motorway (expressway) from the distribution centre (Figure 3(A1)). The vehicles in the city center chose arterial roads such as Jiefang Avenue and Joy Avenue. For customer points at the city's edge area, the vehicles are affected by road speed during distribution. The vehicles chose the motorway (Wuhan Fourth Ring Road) and arterial road (Gaoxin Avenue) for distribution as much as possible (Figure 3(A2)). For the cross-district customer points, the vehicle mainly selected the Wuhan Beltway to deliver to the customer points far from the distribution center (Figure 3(A3)). In addition, SA-SSA

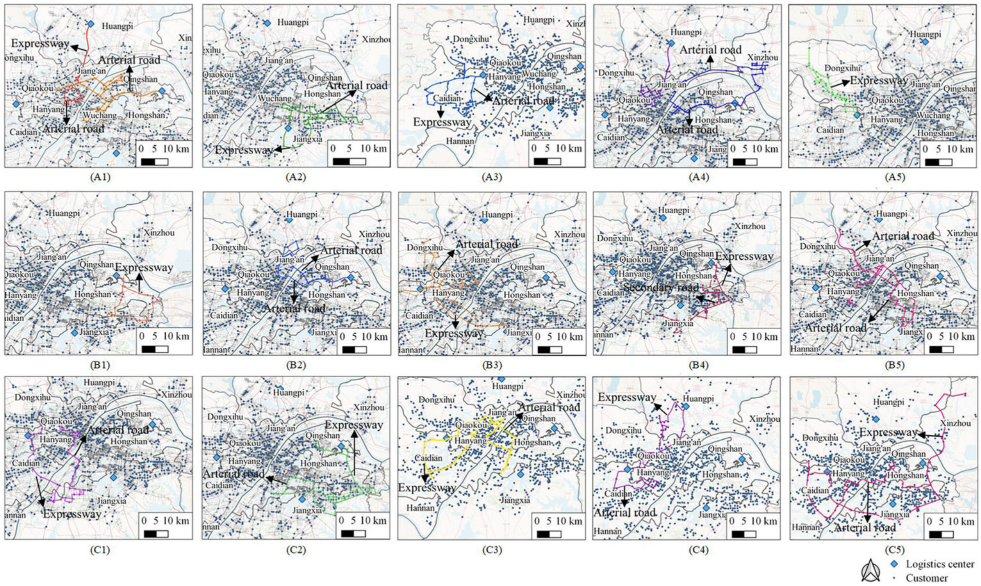

Figure 3. Three examples (A,B,C) of the multi-depot logistic optimization results for five models: (1) Proposed SSA-SA, (2) ACO, (3) GA, (4) SA, (5) SSA.

prefers to concentrate on serving customer points around one logistics centres. It is less likely that one distribution vehicle departs from one logistics centre but serves the customers around another logistics center. Therefore, customer points show spatial aggregation at the logistics centre. In summary, the SA-SSA reduces the vehicle distribution time by considering the influence of vehicle speed and guiding vehicles to choose highways, expressways and main roads for distribution.

Compared with other algorithms, the SSA-SA algorithm allowed the logistics centre to reasonably choose the demand point of the service. For the city with a large area and complex consumer demands, the proposed SSA-SA could divide the large area into small regions serviced by different logistic centres, efficiently carrying out logistics transportation in an extensive range. However, other algorithms for logistical distribution paths may result in unreasonable storage and distribution. For example, some vehicles may depart from one warehouse and travel to deliver to a customer point near another warehouse (Figure 3(B1,C3,D3,E2)). This can lead to increased distribution cost.

5.2. Reliability interpretation and analysis for two scenarios

To verify the performance of SSA-SA for path optimization in real urban scenarios, the set of optimized distribution paths using the SSA-SA algorithm considering road class and without considering road class is shown in Figure 4. The results show that the distribution of logistics routes optimized by the SSA-SA algorithm was reasonable. The shipping routes departing from the same distribution centres had significant spatial aggregation. As is shown in Figure 3, the routes of different colours were sent from different distribution centres. The distribution of the same colour paths was generally

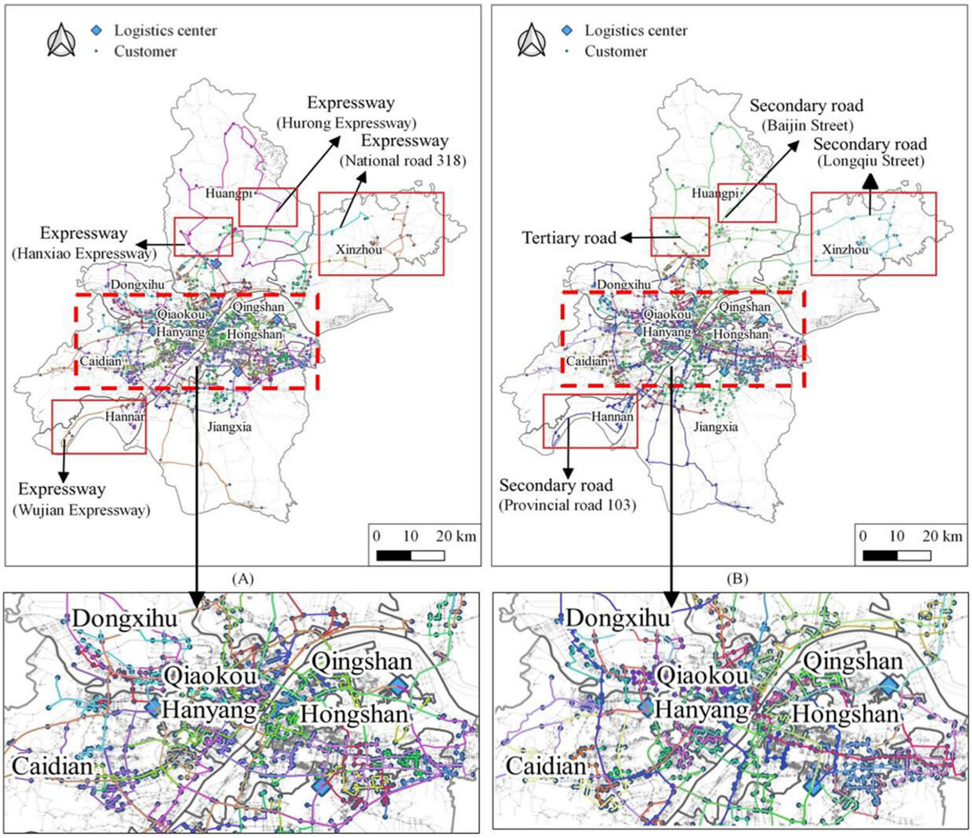

Figure 4. Overall path distribution map of logistics scheduling optimization in Wuhan under two demands (A) considering road class (B) without considering road class. Different colours represent different delivery routes.

more concentrated and distributed in a certain range around the corresponding distribution centres.

In addition, the path optimization results of the proposed algorithm consider the effect of vehicle speed on the optimization results. The optimization results varied for logistics scheduling under different demands. Considering the road level (Figure 4(A)), the vehicle distribution and return to the warehouse prioritized the high-speed road. Without considering the road level (Figure 4(B)), the vehicles were more likely to choose secondary roads, tertiary roads, and other paths with shorter distances for distribution.

The optimization results in two demands were further quantitatively analyzed. We calculate the distance traveled on the different classes of roads through which the logistics distribution path passes. The analysis was carried out from two perspectives: based on individual customer points and based on the delivery vehicles.

First, we analyzed the optimization results based on 1000 customer points (Figure 5(A)). Assuming there is a customer point *i*, the path from the previous customer point *i*-1 to *i* is considered the distribution path for customer point *i*. We calculate the road class of the distribution path for each customer point. Figure 5(B) shows that, on

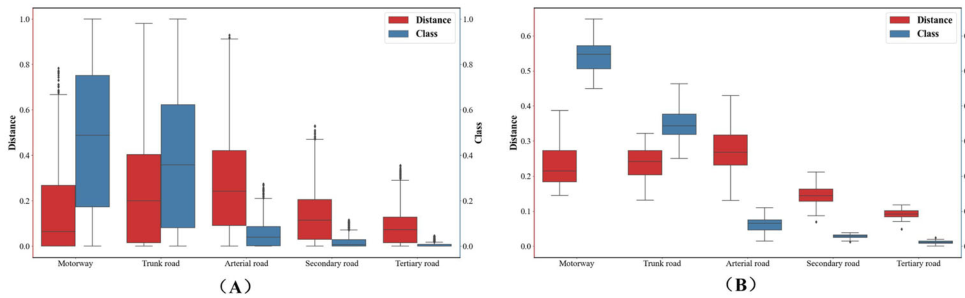

Figure 5. Proportion of logistics distribution routes passing through different road classes in two contexts: (A) based on each customer point; (B) based on each distribution route.

average, the customer point-based distribution path includes 46.85 and 38.60% of motorway and trunk when road class is considered. In comparison, the average percentage of roads passing through the two types of roads without considering road class is 19.30 and 24.83%, respectively. Furthermore, when road class is considered, only an average of 14.56% pass through the other three lower grades of roads. In contrast, 55.86% of customer points are delivered via low-grade roads when road classes are not taken into account, which is 3.84 times more than when road grades are taken into account.

Second, we counted the proportion of the total length of different road classes that each delivery vehicle passes through, as shown in Figure 5(B). The number of vehicles to be delivered is automatically optimized by the algorithm. With 55.43 and 33.64% of motorway roads and trunk roads, respectively, when road class is considered, the distribution path passes through an average of just 10.92% of low grade roads. In contrast, when road class is not considered, an average of 23.24 and 25.19% of motorway roads and trunk roads are passed, with an average of 51.57% passing through low-grade roads. Figure 6 shows the distances and proportions of road grades for all logistics distribution routes for both scenarios. The results show that all delivery vehicles are mainly via motorway (red line) when road class is considered. When road class is not considered, the delivery vehicle tends to go through the trunk and arterial roads.

5.3. Parameters sensitivity analysis

In the SSA-SA algorithm, when the initial temperature *T* and cooling speed were large enough, the maximum number of iterations $iter_{\max}$, the population *N*, and the finding ratio *PD* had some influence on the calculation results. We performed the sensitivity analysis for these three parameters. Thirty optimization experiments were performed at each parameter condition.

Tables 4–6 show the optimization result and the corresponding standard deviation, including the average shortest shipping distance and average computation time. It can be seen from Table 4 that the larger N is, the better the quality of logistics optimization. This is because a larger N can be to increase the search capability of the algorithm and help avoid getting trapped in a local optimum

As shown in Table 5, when *N* and *PD* took specific values, the average shortest distribution path decreased significantly as the $iter_{\max}$ value increased. However, the time

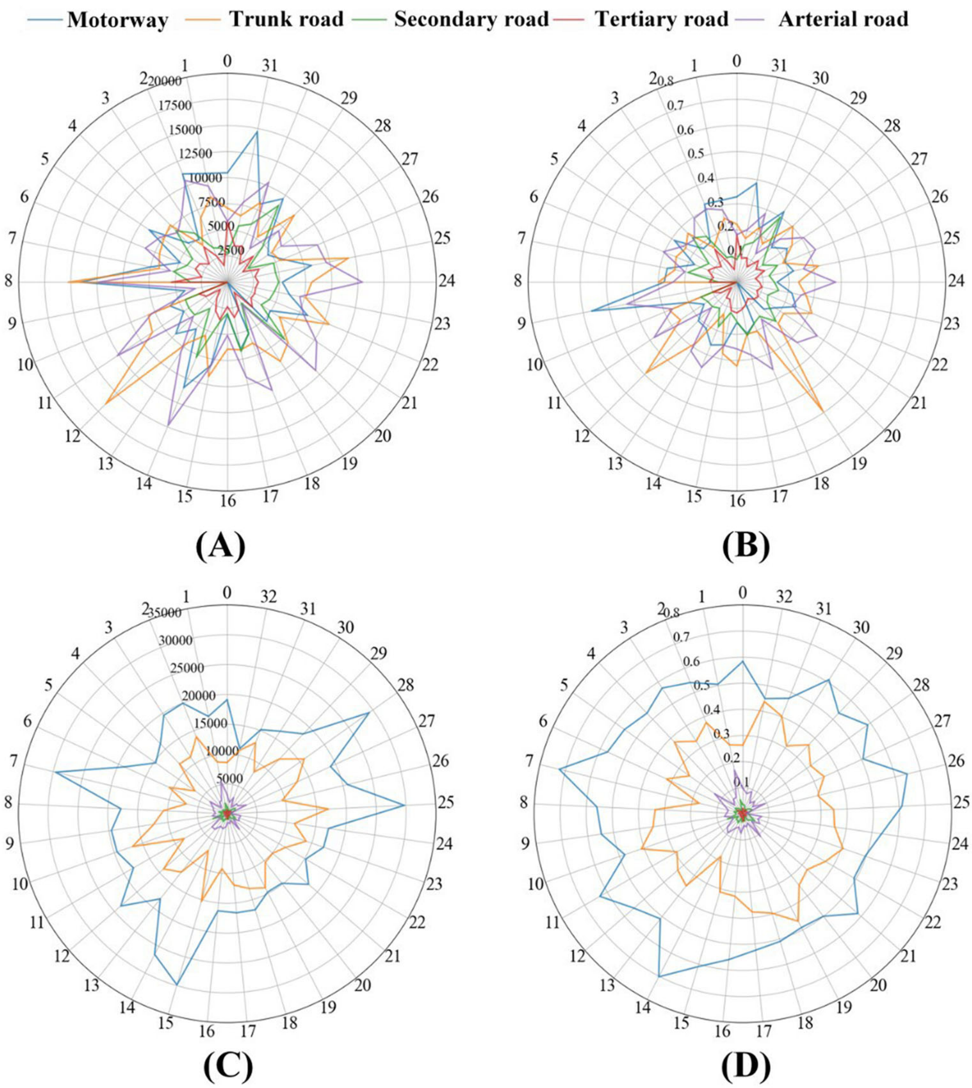

Figure 6. Distance (A,C) and proportion (B,D) of each distribution vehicle's distribution route through different classes of roads: (A,B) without consideration of road class; (C,D) with consideration of road class.

Table 4. Experimental results of different parameter *N*.

$iter_{max}$	*PD*	*N*	Optimal solution/km	Operation time/s
1000	0.7	20	4693.874 ± 90.98	671.68 ± 10.51
1000	0.7	50	4407.619 ± 91.74	804.61 ± 12.73
1000	0.7	80	4352.736 ± 88.14	986.26 ± 13.85

consumption also increased significantly, and the standard deviation value of the solution time became significantly larger. As can be seen from Table 6, the optimal solution and solution time did not change significantly as the PD value of the discoverer increased, but the optimization results were more desirable when the PD value was

Table 5. Experimental results of different parameter. $iter_{max}$.

$iter_{max}$	PD	N	Optimal solution/km	Operation time/s
500	0.7	50	4584.327 ± 98.27	483.15 ± 9.62
1000	0.7	50	4407.619 ± 91.74	804.61 ± 12.73
1500	0.7	50	4361.258 ± 92.80	1357.82 ± 22.54
2000	0.7	50	4218.493 ± 90.38	1546.76 ± 23.08

Table 6. Experimental results of different parameter PD.

$iter_{max}$	PD	N	Optimal solution/km	Operation time/s
1000	0.6	50	4486.239 ± 94.43	791.32 ± 10.66
1000	0.7	50	4407.619 ± 91.74	804.61 ± 12.73
1000	0.8	50	4389.394 ± 93.52	769.03 ± 11.85

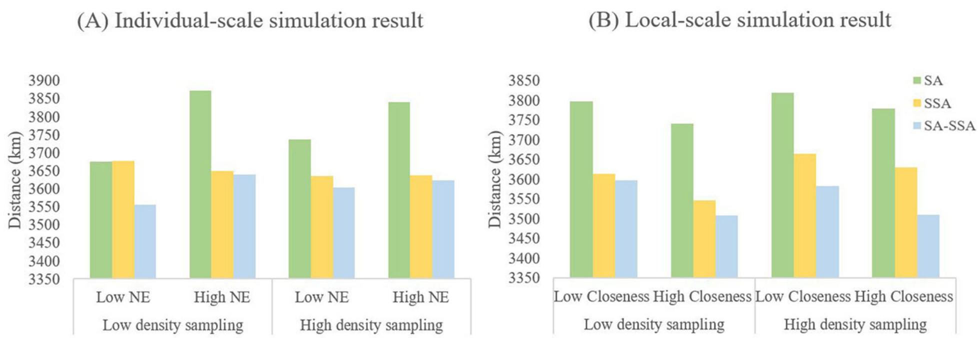

Figure 7. (A) Individual-scale simulation results. Logistics optimization results for customer groups in different levels of network efficiency. (B) Local-scale simulation results. Optimization results for customer groups in different levels of closeness centrality.

chosen reasonably. The results illustrate that the optimization results of the SSA-SA algorithm were most affected by the values of $iter_{max}$ and N. The proposed algorithm can have high efficiency and consider both distribution cost and computation time with the appropriate combination of parameters. In conclusion, the sensitivity analysis results indicate the necessity of choosing the proper parameters when using SA-SSA for logistics optimization.

5.4. Simulation experiment result

5.4.1. Individual-scale simulation result

Network efficiency characterizes how easy it is for the warehouse to serve a particular customer point. Results show that SA-SSA can guarantee the optimal solution when serving customer points of any network efficiency level (Figure 7(A)). When the network efficiency of the customer point is higher, i.e., closer to the warehouse, the distribution distance of SA is longer. In particular, in the case of high-density sampling, the distribution distance of SA improves by 2.66%, the path of SSA improves by 0.09%, and the path of SA-SSA improves by 0.57% when serving higher network efficiency customers. In the case of low-density sampling, the distribution path of SA improves by 5.05%, the path of SSA decreases by 0.75%, and the path of SA-SSA improves by 2.29% when serving customers with higher network efficiency.

The results reveal that it is challenging for the optimization algorithm to achieve a reasonable spatial division in multi-warehouse logistics distribution. Specifically, there is the situation that customer points near a specific warehouse are distributed by vehicles from another warehouse, leading to an increase in distribution cost. This problem also affects the performance of SA-SSA, increasing its distribution path distance. However, when the network efficiency increases, the distribution path of SSA decreases (0.75%, low-density sampling) or only slightly increases (0.09%, high-density sampling). This indicates that SSA can consider the spatial partitioning problem of the multi-warehouse logistics task. SA-SSA has a partial improvement in this problem due to the introduction of SSA. The algorithm of SA-SSA do not show a similar degree of distribution efficiency degradation as SA when serving customer points around the warehouse.

5.4.2. Local-scale simulation result

Closeness centrality can characterize the proximity between a certain customer point and all other customer points in the urban road network. Results show that in the case of high-density sampling (Figure 7(B)), SA-SSA reduced the distribution distance by 71.41 km for customer points serving high closeness centrality (3511.02 km) than those serving low closeness centrality (3582.43 km). It indicates that a customer point is more likely to be selected by SA-SSA when it is closer to other customer points, which leads to high-quality distribution routes.

In contrast, the effect of closeness centrality of customers on SA is small. In the same situation, the distribution distance optimized by SA and SSA in serving high-closeness centrality customer points shrinks by 39.63 and 35.40 km, almost half the size of SA-SSA. In summary, results show that when the customer point has high closeness centrality, it will be preferred by SA-SSA, resulting in a better solution and lower cost.

5.4.3. Global-scale simulation result

Figure 8(A) shows the spatial clustering results that all customer points were divided into four clusters. SA-SSA had the best optimization results among all four clusters in terms of the distribution distance (Figure 8(B)). Among them, the distance optimized

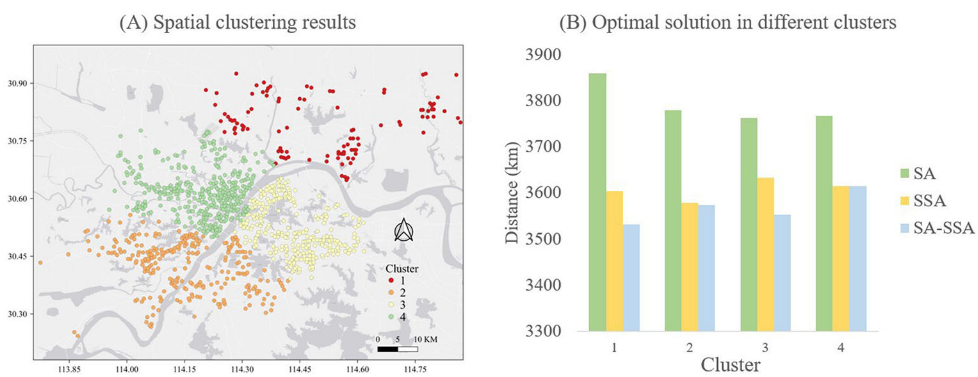

Figure 8. (A) Spatial clustering results of customer points based on spectral clustering and (B) logistics optimization results for customers in different spatial clusters.

by SA-SSA are 2.05 and 2.25% shorter than those of SSA in clusters 1 and 3, respectively.

In addition, results show that in a sparse urban road network, the logistics optimization algorithm is more likely to fall into a local optimum, especially for SA In the northeast suburb of Wuhan (cluster 1, red dots), the road network is sparse and the distribution of customer points is more dispersed, and the local optimum problem of logistics distribution algorithm is very significant. In addition, the logistics distribution distance of SA is 7.12% higher than that of SSA, and 9.31% higher than that of SA-SSA. In the southern suburb of Wuhan city (cluster 2), the logistics distribution distance of SA is 5.58 and 5.74% higher than SSA and SA-SSA, respectively. With such a sparse urban road network, it is difficult for SA to find a high-quality distribution path. At the same time, SSA can effectively search for high-quality paths and greatly reduce the distribution distance compared to SA. In the urban centre area (clusters 3 and 4), the urban road network is tighter, and the local optimum problem of the logistics optimization algorithm is relatively minor compared with the suburban area.

In conclusion, spatially sparse road networks challenge logistics optimization algorithms to find high-quality solutions and thus make them fall into the local optimum problem. By combining SSA with SA, the SA-SSA can effectively cope with logistics distribution tasks under sparse road networks (e.g., suburban areas), significantly improve distribution efficiency and reduce costs.

6. Discussion

6.1. Interpretation of findings

As a wayfinding problem, logistics optimization is a classic application of GIS. In a complex urban road network, various spatial features and spatial relationships between customer points and warehouse locations affect the performance of logistics optimization algorithms. In addition, the increasing customer demand also challenges the efficiency of logistics optimization algorithms (Laporte 2009, Yao *et al*. 2018). However, few studies are oriented to real urban logistics shipping scenarios, ignoring the increasingly complex road networks and growing customer demands in large cities. This study introduced the simulated annealing algorithm into SSA and constructed a hybrid heuristic algorithm (SA-SSA) to address this problem. The logistics optimization was performed for four logistical warehouses and 1,000 customer points in Wuhan. The results show that the proposed SSA-SA algorithm can efficiently solve the large-scale urban logistics optimization problem.

Compared with classical heuristics such as SA, ACO, and GA, the SSA-SA algorithm had the best performance in solving urban logistics problems. Compared with the SSA, the SSA-SA algorithm reduced the minimum delivery distance by 18.62% and completed the delivery in 13.40 minutes (804.61s) on average. It can consider complex constraints in the real urban logistics scenario, such as urban roads and multi-bin distribution. The theoretical innovation of SA-SSA is to combine the advantages of the SSA with the SA. The sparrow search algorithm has greater potential for large-scale logistics optimization, and can achieve acceptable optimization results in a short time. While the SA has a greater dependence on the initial solution, a good initial solution can improve the optimization efficiency. SA-SSA combines the complementary advantages of SSA's fast solving capability and avoids getting stuck at the local optimum of SA. In this way, the SSA-SA can balance the time and cost requirements of urban logistics optimization, thus having the greater potential to solve large-scale logistics optimization problems in cities.

The experimental results in Wuhan city prove the effectiveness of the proposed algorithm in solving logistical optimization problems. The interpretable analysis demonstrates that the service points of each logistic centre show spatial aggregation using the SSA-SA algorithm. Without using other optimization strategies, the SSA-SA algorithm can spontaneously perform partitioned distribution services for large-scale customer points, making the vehicles serve small-scale logistics distribution and more conducive to setting logistics service sites. In small-scale distribution, the proposed algorithm can conduct distribution with less cross-bridge, effectively avoiding cross-regional large-scale distribution and thus reducing the overall logistics cost.

The proposed SSA-SA algorithm can efficiently solve the large-scale multi-depot vehicle routing problems for one thousand customers within 20 minutes, which has excellent potential for solving practical optimization problems. In this study, it is found that the selection of parameters $iter_{\max}$ and N had a great influence on the performance of the SSA-SA algorithm. The results of parameter sensitivity analysis show that the parameter $iter_{\max}$ affected the global optimal solution of the algorithm. When the parameter $iter_{\max}$ increased, the optimal solution improved significantly, but the time overhead of the algorithm also increased. In future research, coupling our previous proposed high-performance spatially intelligent computing framework (Laure 2001) can be considered to achieve real-time mega-city logistics scheduling performance.

We conducted three simulation experiments at different scales, which were used to explore the influence of spatial elements on logistics optimization algorithms. The results show that it is still difficult for the logistics optimization algorithm to achieve good spatial partitioning in a multi-warehouse logistics distribution scenario. Specifically, there is an unreasonable problem that customer points near a particular warehouse are distributed by vehicles from another warehouse, leading to an increase in distribution cost. Our proposed SA-SSA can partially solve this problem, but further improvement is still needed. In addition, the proximity of a customer point to other customers in the road network affects the performance of the algorithm. When a customer point is more accessible in the road network, it will be preferred by our proposed SA-SSA, thus reducing the distance of the whole distribution path. Finally, the sparse road network and more dispersed distribution of customer points make it difficult for the logistics optimization algorithm to find high-quality solutions and thus fall into the local optimum problem. Our proposed SA-SSA can greatly improve distribution efficiency and reduce costs in areas with sparse road networks.

The computational efficiency of the optimization algorithm is essential for applications in large-scale logistics. The algorithm complexity affects the performance as well as the applicability of the logistics optimization algorithm in practical application scenarios. In solving logistics optimization problems, the time complexity of the exact algorithm is O (d!) (d is the number of customer points). In small-scale urban logistics optimization problems, exact algorithms are often used. In the case of large-scale logistics optimization, the actual complexity is astronomical, and it is impossible to obtain results in a limited time. Therefore, we choose a heuristic algorithm to solve the large-scale logistics optimization problem.

Our results show that the shortest path of ACO is lower than that of SA-SSA, but its computation time is too long and, thus, difficult to be applied in practical logistics distribution. The time complexity of the ACO is O (d^2). In contrast, the time complexity of its SSA is O (d + Klog(K)) (K is the number of sparrows). The time complexity of the SA is O (d + K). The time complexity of the SA-SSA is O (d + Klog(K)). The ACO, therefore, requires an iterative selection of clients compared to the sparrow hybrid algorithm, which also results in higher time consumption.

6.2. The application of SA-SSA in a complex urban system

The city can be considered a complex system with high uncertainty (Iturriza *et al.* 2020). In addition, some large-scale urban logistics tasks, such as postal delivery or garbage collection, often require servicing a large number of customers (e.g. over 1000). These two factors pose the challenge for logistics optimization algorithms.

This research is oriented to the application scenario of mega-cities and large-scale urban logistics, conducting large-scale pathfinding optimization in a real scenario using GIS technology. We developed an algorithm to solve the logistics route optimization problem efficiently. Some previous studies improve the heuristic algorithm to optimize logistics (Yang *et al*. 2015, Xue and Cao 2016, Tu *et al.* 2017). However, these theoretical studies usually cannot balance shipping cost and efficiency when facing highly complex road networks and huge customer demand in practical application scenarios. Dedicated to the engineering implementation of logistics optimization, this study has the following main contributions. First, to apply to the most realistic urban logistics distribution scenarios, the proposed algorithm considers complex constraints such as road class, multi-bin distribution and vehicle capacity limits (Cattaruzza *et al.* 2017).

Second, the real-time changing urban traffic can make the logistics optimization results uncertain, which is a tremendous challenge for logistics optimization. However, the related research is still in the gap. We build a customer inter-node dataset based on a transportation network graph, balancing transportation efficiency and cost. The edge weights W of the weighted graph can be set as traffic flow, such as real-time traffic data, or traffic flow predicted by taxi data. Therefore, by combining the dynamic traffic flow data, the SSA-SA model can dynamically select the path with a shorter delivery time for delivery service, thus optimizing real-time and dynamic mega-city logistics scheduling.

Third, we applied the heuristic algorithm for the first time to optimize large-scale urban logistics considering road class. The results show that the proposed algorithm has a strong application background and can provide a reference for practical urban logistics planning. When the SSA-SA algorithm provides distribution service to dense customer points in the city center, the planned path selects the main city roads and secondary roads for distribution.

6.3. Limitations and future works

There are still some shortcomings in this study. First, the factors affecting urban logistics scheduling are complex, and some are dynamic. However, only three urban constraints are considered in this study to solve the urban logistics optimization problem: customer point distribution volume, multiple warehouse centres, and complex road networks. In subsequent studies, constraints such as real-time traffic conditions, carbon emissions, and customer service time windows can be considered to solve urban logistics optimization problems in different scenarios. In addition, the primary objective of this work is to design a generalized optimization algorithm in a basic multi-warehouse logistics optimization scenario. Therefore more complex logistics optimization problems are not considered. For example, we do not consider the constraints on the number of vehicles and the subloop elimination problem. In subsequent studies, we will explore the optimization performance of SA-SSA under more complex constraints.

Second, as cities have heterogeneity in scale and structure, different types of cities need to be selected to test the validity of SA-SSA. The research context of this paper is that the issue of efficiency in logistics and distribution is being challenged with the accelerating urbanization process. Therefore, we have chosen a typical fast-growing mega-city for our experiment. We believe logistics and distribution problems are the most challenging and problematic in such a city. Based on this principle, we chose Wuhan as our study area. In future research, we will try to select several cities of different types and sizes to test the robustness of our proposed SA-SSA model.

Third, reinforcement learning has become one of the popular methods for solving combinatorial optimization problems. Since urban logistics path optimization is complex and usually large-scale, the exploration strategies commonly used in reinforcement learning often fail. In subsequent studies, we will introduce deep reinforcement learning for heuristic search operators to speed up the solution and improve the quality of the solution.

7. Conclusions

For large-scale urban logistics optimization tasks, traditional heuristics cannot provide high-quality vehicle path planning solutions quickly. This study proposes a hybrid heuristic algorithm based on the SSA for effectively solving the multi-depot vehicle routing problems considering complex road networks. Experiments on large-scale logistics optimization in Wuhan city show that the proposed SA-SSA algorithm can efficiently provide high-quality vehicle path solutions with stable performance, which is suitable for large-scale urban logistics optimization. The SA-SSA can provide vehicle path design services for relevant logistics enterprises, thus improving logistics efficiency, reducing logistics costs and realizing logistics intelligence. In addition, this study explored the impact of spatial characteristics of cities as well as customer points on logistics optimization algorithms. However, this study still has some limitations, such as the lack of consideration of the real-time traffic condition. In the future, we will apply the proposed method in different types of cities and try to introduce deep reinforcement learning to improve the efficiency of logistics route optimization.

Acknowledgements

We would like to thank the editors and anonymous reviewers for their constructive suggestions and comments for improving this manuscript. We also thank Dr. Yongze Song for his insightful suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Key Research and Development Program of China [Grant No. 2019YFB2102903], the National Natural Science Foundation of China [Grant No. 41801306 and 42171466], Alibaba Innovative Resarch Project [No. 20228670], the "CUG Scholar" Scientific Research Funds at China University of Geosciences (Wuhan) (Project No. 2022034) and a grant from State Key Laboratory of Resources and Environmental Information System.

Notes on contributors

Yao Yao is a professor at China University of Geosciences (Wuhan), a researcher from the Center for Spatial Information Science at the University of Tokyo, and a senior algorithm engineer at Alibaba Group. His research interests are geospatial big data mining, analysis, and computational urban science.

Siqi Lei is currently a master's student at China University of Geosciences (Wuhan). Her research interests include geospatial spatial analysis, logistics optimization, and applied artificial intelligence.

Zijin Guo is a master's student at China University of Geosciences (Wuhan). His research interests are trajectory data mining and complex network analysis.

Yuanyuan Li is currently a Ph.D. candidate at China University of Geosciences (Wuhan). His research interests include geospatial spatial analysis, logistics optimization, and applied artificial intelligence

Shuliang Ren is currently a Ph.D. candidate in GIScience at the Institute of Remote Sensing and Geographical Information Systems,Peking University, Beijing. His research interests are GeoAI, spatial analysis, and urban data mining.

Zhihang Liu is a master's student at Peking University. He is currently a visiting student at the Technical University of Munich, German. His research interests include Urban Computing, Computational Social Science and Complex Networks.

Qingfeng Guan is a professor at China University of Geosciences (Wuhan). His research interests are high-performance spatial intelligence computation and urban computing.

Peng Luo is a Ph.D. candidate at the Chair of Cartography and Visual Analytics at the Technical University of Munich, Germany. His research interests include spatial association modelling, social sensing, and applied artificial intelligence.

ORCID

Yao Yao http://orcid.org/0000-0002-2830-0377
Qingfeng Guan http://orcid.org/0000-0002-7392-3709
Peng Luo http://orcid.org/0000-0002-3680-8509

Data and codes availability statement

The data and codes that support the findings of the present study are available on Figshare at https://doi.org/10.6084/m9.figshare.19289150.

References

Aliakbari, A., *et al.*, 2022. A new robust optimization model for relief logistics planning under uncertainty: a real-case study. *Soft Computing*, 26 (8), 3883–3901.

Arulkumaran, K., *et al.*, 2017. Deep reinforcement learning: a brief survey. *IEEE Signal Processing Magazine*, 34 (6), 26–38.

Bergmann, F.M., Wagner, S.M., and Winkenbach, M., 2020. Integrating first-mile pickup and last-mile delivery on shared vehicle routes for efficient urban e-commerce distribution. *Transportation Research Part B: Methodological*, 131, 26–62.

Bettinelli, A., Ceselli, A., and Righini, G., 2011. A branch-and-cut-and-price algorithm for the multi-depot heterogeneous vehicle routing problem with time windows. *Transportation Research Part C: Emerging Technologies*, 19 (5), 723–740.

Cattaruzza, D., *et al.*, 2017. Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics*, 6 (1), 51–79.

Chakraborty, A., and Kar, A.K., 2017. Swarm intelligence: a review of algorithms. *Nature-Inspired Computing and Optimization*, 475–494.

Chan, L.M.A., *et al.*, 2002. Effective zero-inventory-ordering policies for the single-warehouse multi-retailer problem with piecewise linear cost structures. *Management Science*, 48 (11), 1446–1460.

Crainic, T.G., Ricciardi, N., and Storchi, G., 2009. Models for evaluating and planning city logistics systems. *Transportation Science*, 43 (4), 432–454.

Dasdemir, E., *et al.*, 2022. A multi-objective open vehicle routing problem with overbooking: Exact and heuristic solution approaches for an employee transportation problem. *Omega*, 108, 102587.

Dong, X., *et al.*, 2021. ITÖ algorithm with local search for large scale multiple balanced traveling salesmen problem. *Knowledge-Based Systems*, 229, 107330.

Eilon, S., *et al.*, 1974. Distribution management-mathematical modelling and practical analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, 21 (6), 589–589.

Gayialis, S.P., Konstantakopoulos, G.D., and Tatsiopoulos, I.P., 2019. Vehicle routing problem for urban freight transportation: a review of the recent literature. *In*: *Operational research in the digital Era–ICT challenges: 6th International symposium and 28th national conference on operational research*. June 2017. Thessaloniki, Greece: Springer International Publishing, 89–104.

Gemma, P.L., *et al.*, 2016. Cost efficiency in municipal solid waste service delivery. Alternative management forms in relation to local population size. *European Journal of Operational Research*, 255 (2), 583–592.

Groß, P., Ehmke, J.F., and Mattfeld, D.C., 2020. Interval travel times for robust synchronization in city logistics vehicle routing. *Transportation Research Part E-Logistics and Transportation Review*, 143, 102058.

Haklay, M., 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design*, 37 (4), 682–703.

Han, M., and Wang, Y., 2018. A survey for vehicle routing problems and its derivatives. *In* IOP *conference series: materials science and engineering,* vol. 452. IOP Publishing, 042024.

Hill, A.V., and Benton, W.C., 1992. Modelling intra-city time-dependent travel speeds for vehicle scheduling problems. *Journal of the Operational Research Society*, 43 (4), 343–351.

Hina, S.M., *et al.*, 2020. Effective municipal solid waste collection using geospatial information systems for transportation: a case study of two metropolitan cities in Pakistan. *Research in Transportation Economics*, 84, 100950.

Hwang, C., 1988. Simulated annealing: theory and applications. *Acta Applicandae Mathematicae*, 12 (1), 108–111.
Imani, M., and Ghoreishi, S.F., 2021. Graph-BASED Bayesian optimization for large-scale objective-based experimental design. *IEEE Transactions on Neural Networks and Learning Systems*, 33 (10), 5913–5925.
Iturriza, M., *et al.*, 2020. Shifting to climate change aware cities to facilitate the city resilience implementation. *Cities*, 101, 102688.
Janic, M., 2007. Modelling the full costs of an intermodal and road freight transport network. *Transportation Research Part D-Transport and Environment*, 12 (1), 33–44.
Jiang, B., and Claramunt, C., 2004. A structural approach to the model generalization of an urban street network. *GeoInformatica*, 8 (2), 157–171.
Jiang, B., and Liu, C., 2009. Street-based topological representations and analyses for predicting traffic flow in GIS. *International Journal of Geographical Information Science*, 23 (9), 1119–1137.
Kim, G., *et al.*, 2015. City vehicle routing problem (city VRP): a review. *IEEE Transactions on Intelligent Transportation Systems*, 16 (4), 1654–1666.
Kumaravel, S., and Ponnusamy, V., 2020. An efficient hybrid technique for power flow management in smart grid with renewable energy resources. *Energy Sources Part A: Recovery Utilization and Environmental Effects*, 1–21.
Laporte, G., 1984. Optimal solutions to capacitated multidepot vehicle routing problems. *Congressus Nemerantium*, 4, 283–292.
Laporte, G., 2009. Fifty years of vehicle routing. *Transportation Science*, 43 (4), 408–416.
Laporte, G., Nobert, Y., and Taillefer, S., 1988. Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation Science*, 22 (3), 161–172.
Latora, V., and Marchiori, M., 2001. Efficient behavior of small-world networks. *Physical Review Letters*, 87 (19), 198701.
Laure, E., 2001. OpusJava: A Java framework for distributed high performance computing. *Future Generation Computer Systems*, 18 (2), 235–251.
Li, H., *et al.*, 2017. The vehicle flow formulation and savings-based algorithm for the rollon-rolloff vehicle routing problem. *European Journal of Operational Research*, 257 (3), 859–869.
Li, Q.Q., *et al.*, 2015. A hybrid link-node approach for finding shortest paths in road networks with turn restrictions. *Transactions in GIS*, 19 (6), 915–929.
Liu, B., and Rodriguez, D., 2021. Renewable energy systems optimization by a new multi-objective optimization technique: A residential building. *Journal of Building Engineering*, 35 (3), 102094.
Liu, G., *et al.*, 2018. Improving the throughput of transportation networks with a time-optimization routing strategy. *International Journal of Geographical Information Science*, 32 (9), 1815–1836.
Liu, G., *et al.*, 2021. A modified sparrow search algorithm with application in 3D route planning for UAV. *Sensors*, 21 (4), 1224.
Lu, D., and Gzara, F., 2019. The robust vehicle routing problem with time windows: solution by branch and price and cut. *European Journal of Operational Research*, 275 (3), 925–938.
Martin, F., Hemmelmayr, V.C., and Wakolbinger, T., 2021. Integrated express shipment service network design with customer choice and endogenous delivery time restrictions. *European Journal of Operational Research*, 294 (2), 590–603.
Mester, D., and Bräysy, O., 2005. Active guided evolution strategies for large-scale vehicle routing problems with time windows. *Computers & Operations Research*, 32 (6), 1593–1614.
Mester, D., Bräysy, O., and Dullaert, W., 2007. A multi-parametric evolution strategies algorithm for vehicle routing problems. *Expert Systems with Applications*, 32 (2), 508–517.
Mnih, V., *et al.*, 2013. Playing atari with deep reinforcement learning. *ArXiv Preprint arXiv: 1312.5602*.
Mouthuy, S., *et al.*, 2015. A multistage very large-scale neighborhood search for the vehicle routing problem with soft time windows. *Transportation Science*, 49 (2), 223–238.
Nasiri, S.D., 2014. *Vehicle routing on real road networks*. United Kingdom: Lancaster University.

Ng, A., Jordan, M., and Weiss, Y., 2001. On spectral clustering: analysis and an algorithm. *Advances in Neural Information Processing Systems*, 14, 849.
Nozari, H., Tavakkoli-Moghaddam, R., and Gharemani-Nahr, J., 2022. A neutrosophic fuzzy programming method to solve a multi-depot vehicle routing model under uncertainty during the covid-19 pandemic. *International Journal of Engineering*, 35 (2), 360–371.
Okamoto, K., Chen, W., and Li, X.Y., 2008. Ranking of closeness centrality for large-scale social networks. *Lecture Notes in Computer Science*, 5059, 186–195.
Ouyang, A.J., and Zhou, Y.Q., 2011. An improved PSO-ACO algorithm for solving large-scale TSP. *Advanced Materials Research*, 143, 1154–1158.
Ouyang, C., Qiu, Y., and Zhu, D., 2021. A multi-strategy improved sparrow search algorithm. *Journal of Physics*, 1848 (1), 12042.
Phiboonbanakit, T., *et al.*, 2021. A hybrid reinforcement learning-based model for the vehicle routing problem in transportation logistics. *IEEE Access.*, 9, 163325–163347.
Pourabdollah, A., *et al.*, 2013. Towards an authoritative OpenStreetMap: conflating OSM and OS OpenData national maps' road network. *ISPRS International Journal of Geo-Information*, 2 (3), 704–728.
Qin, S., *et al.*, 2021. Large-scale evolutionary multiobjective optimization assisted by directed sampling. *IEEE Transactions on Evolutionary Computation*, 25 (4), 724–738.
Rahim, M.K.I.A., *et al.*, 2016. Analyzing the effectiveness of vendor-managed inventory in a single-warehouse, multiple-retailer system. *International Journal of Systems Science*, 47 (8), 1953–1965.
Salman, R., Ekstedt, F., and Damaschke, P., 2020. Branch-and-bound for the precedence constrained generalized traveling salesman problem. *Operations Research Letters*, 48 (2), 163–166.
Shen, Y., 2020. Optimization of urban logistics distribution path under dynamic traffic network. *International Core Journal of Engineering*, 6 (1), 243–248.
Shen, Y., *et al.*, 2021. Operating policies in multi-warehouse drone delivery systems. *International Journal of Production Research*, 59 (7), 2140–2156.
Shi, L., *et al.*, 2022. Memory-based ant colony system approach for multi-source data associated dynamic electric vehicle dispatch optimization. *IEEE Transactions on Ransactions on Intelligent Transportation Systems*, 23 (10), 17491–17505.
Shimizu, S., Yamaguchi, K., and Masuda, S., 2020. A maximum edge-weight clique extraction algorithm based on branch-and-bound. *Discrete Optimization*, 37, 100583.
Shirabe, T., 2014. A path that buys time to decide where to go. *International Journal of Geographical Information Science*, 28, 314–325.
Toth, P., and Vigo, D., 2002. *The vehicle routing problem*. Society of Industrial and Applied Mathematics.
Tu, W., *et al.*, 2015. A novel spatial-temporal Voronoi diagram-based heuristic approach for large-scale vehicle routing optimization with time constraints. *ISPRS International Journal of Geo-Information*, 4 (4), 2019–2044.
Tu, W., *et al.*, 2017. A spatial parallel heuristic approach for solving very large-scale vehicle routing problems. *Transactions in GIS*, 21 (6), 1130–1147.
Vanhove, S., and Fack, V., 2012. An effective heuristic for computing many shortest path alternatives in road networks. *International Journal of Geographical Information Science*, 26, 1031–1050.
Wang, F., Wang, X., and Sun, S., 2022. A reinforcement learning level-based particle swarm optimization algorithm for large-scale optimization. *Inforamtion Sciences*, 602, 298–312.
Wang, H.F., and Chen, Y.Y., 2012. A genetic algorithm for the simultaneous delivery and pickup problems with time window. *Computers & Industrial Engineering*, 62 (1), 84–95.
Wang, M., *et al.*, 2013. Quality analysis on crowd sourcing geographic data with open street map data. *Geomatics and Information Science of Wuhan University*, 38 (12), 1490–1494.
Wang, Y., *et al.*, 2017. Cooperation and profit allocation in two-echelon logistics joint distribution network optimization. *Applied Soft Computing*, 56, 143–157.
Wang, Y., *et al.*, 2020. Collaborative multi-depot logistics network design with time window assignment. *Expert Systems with Applications*, 140, 112910.

Wang, Z.-J., *et al.*, 2021. Adaptive granularity learning distributed particle swarm optimization for large-scale optimization. *IEEE Transactions on Cybernetics*, 51 (3), 1175–1188.
Watkins, C.J., and Dayan, P., 1992. Q-learning. *Machine Learning*, 8 (3–4), 279–292.
Xue, J., and Shen, B., 2020. A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems Science & Control Engineering*, 8 (1), 22–34.
Xue, W., and Cao, K., 2016. Optimal routing for waste collection: a case study in Singapore. *International Journal of Geographical Information Science*, 30 (3), 554–572.
Yang, J., *et al.*, 2015. An improved ant colony optimization (I-ACO) method for the quasi travelling salesman problem (Quasi-TSP). *International Journal of Geographical Information Science*, 29 (9), 1534–1551.
Yao, Y., *et al.*, 2018. Estimating the effects of 'community opening' policy on alleviating traffic congestion in large Chinese cities by integrating ant colony optimization and complex network analyses. *Computers, Environment and Urban Systems*, 70, 163–174.
Zachariadis, E.E., and Kiranoudis, C.T., 2010. A strategy for reducing the computational complexity of local search-based methods for the vehicle routing problem. *Computers & Operations Research*, 37 (12), 2089–2105.
Zeng, W., and Church, R.L., 2009. Finding shortest paths on real road networks: the case for A*. *International Journal of Geographical Information Science*, 23, 531–543.
Zhou, L., *et al.*, 2018. A multi-depot two-Echelon vehicle routing problem with delivery options arising in the last mile distribution. *European Journal of Operational Research*, 265 (2), 765–778.