# A New Approach for Solving Location Routing Problems with Deep Reinforcement Learning of Emergency Medical Facility

Shaohua Wang†
Aerospace Information Research Institute
CAS Beijing 10094, China
wangshaohua@aircas.ac.cn

Junyuan Zhou
Faculty of Geomatics, Lanzhou Jiaotong University,
Lanzhou 730070, China
11220851@stu.lzjtu.edu.cn

Haojian Liang†
School of Artificial Intelligence
Jilin University
Changchun 130012, China
hjliang20@mails.jlu.edu.cn

Zhenbo Wang
Institute of Geographic Sciences and Natural Resources Research,
CAS Beijing 10094, China
wangzb@igsnrr.ac.cn

Cheng Su
Aerospace Information Research Institute
CAS Beijing 10094, China
sucheng23@mails.ucas.ac.cn

Xiao Li
Faculty of Geomatics, Lanzhou Jiaotong University
Lanzhou 730070, China
11220869@stu.lzjtu.edu.cn

## ABSTRACT

The Location Routing Problem (LRP) has extensive applications in emergency medical facility selection. It involves a complex spatial optimization problem of selecting the optimal facility locations from a set of demand points and candidate facility locations and planning the most efficient routes. This study introduces a two-stage solution approach based on deep reinforcement learning, which fully leverages the potential of deep reinforcement learning in handling sequential decision problems. Firstly, in the location selection stage, through the interaction between an agent and the environment, the agent needs to choose facility locations from numerous demand points to either maximize overall profit or minimize overall costs. This stage determines central points and allocates demand points to the nearest facility locations. Subsequently, in the route planning stage, the agent decides how to distribute deliveries among the pre-selected facility locations to minimize overall route length or costs. Experimental results demonstrate the effectiveness of this method in addressing location routing problems. It may provide valuable support for responding to emergencies and disaster events, helping decision-makers make more intelligent choices for facility locations and route planning, ultimately enhancing service response speed and efficiency. Furthermore, this method holds broad application potential in other domains, further advancing the development and utilization of deep reinforcement learning in spatial optimization problems.

## CCS CONCEPTS

•Computing methodologies~Artificial intelligence~Planning and scheduling~Planning for deterministic actions


EM-GIS '23, November 13, 2023, Hamburg, Germany

ACM ISBN 979-8-4007-0346-1/23/11.
https://doi.org/10.1145/3615884.3629429

## KEYWORDS

Location Routing Problem, Emergency Medical Facility, Deep Reinforcement Learning, Attention Model

**ACM Reference format:**

Shaohua Wang, Junyuan Zhou, Haojian Liang, Zhenbo Wang, Cheng Su and Xiao Li. 2023. A New Approach for Solving Location Routing Problems with Deep Reinforcement Learning of Emergency Medical Facility. In Proceedings of The 8th ACM SIGSPATIAL International Workshop on Security Response using GIS 2023 (EM-GIS 2023). ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3615884.3629429.

## 1 INTRODUCTION

With the continuous improvement of transportation capacity, many requirements are put forward for logistics management, and site selection and route planning are the two most important issues. Scientific and reasonable site selection and path planning are conducive to improving the efficiency of the entire logistics system while reducing costs. To meet the different needs of various logistics systems, different aspects of site selection and route planning are studied [1, 2].

At present, the rapid development of science and technology in the world and the continuous advancement of the strategy of building a powerful transportation country are still occurring, while sudden natural disasters, public health incidents, terrorism, and regional military conflicts still occur from time to time. How to reasonably choose the location of the logistics center, quickly and safely transport emergency materials to the demand point, and better control the total cost is a core problem of logistics. Integrating geospatial features, spatial optimization models, and cutting-edge deep learning technologies introduces innovative perspectives and revitalizes research in the field of location problems [3-5].

The most important thing is the Facility Location Problem (FLP) and the Vehicle Routing Problem (VRP), the rationality of the former has a certain impact on the efficiency of the latter, and the efficiency of the latter in turn affects the decision-making of the former, and the two complement each other, so if the two decisions are made separately, the optimal results cannot be produced, which is described in Salhi S et al. [6] It has been confirmed in studies. With Cooper [7] first proposing LRP, combining the problem of facility location with the problem of vehicle path, LRP has become a research hotspot in the field of logistics system optimization. Foroughi and Gokcen [8] measured the cost factors and analyzed the opportunity constraints when modeling random site selection and path planning problems, and the results of large-scale genetic algorithm solving based on multiple rules show that the proposed model has better performance and efficiency.

Location problem and path planning involve complex factors and dynamic characteristics, which are relatively difficult to solve the model of related problems. Therefore, many scholars use intelligent algorithms to solve models, which can be roughly divided into two types: accurate algorithms and heuristic algorithms. For solving small-scale LRP problems by accurate algorithms, Laporte and Nobert [9] established an integer programming model for single-facility LRP problems and solved them by branching delimitation method, which is a preliminary attempt to apply accurate algorithms to LRP problems. Later, many scholars used the branching delimitation method of 0-1 linear model [10] to solve small and medium-sized LRP problems based on complex, cutting and delimitation algorithms [11]. But accurate algorithms are only suitable for solving small and medium-sized problems, and more and more scholars are focusing on heuristics that can handle large-scale examples. Ferreira and Queiroz [12] propose two heuristic algorithms based on simulated annealing, which represent the problem solution with a set of integer vectors, and then use exchange and insertion operations to generate new problem solutions, in addition to the perturbation method and discretization to avoid falling into the local optimal solution. In order to solve the LRP problem, Schneider and Loffler [13] proposed a tree-based search algorithm, in which the location decision is represented by a tree structure to define the facilities to be opened or closed, and finally path planning is carried out based on the ideas of relocation, exchange, and splitting combined with the tabu search algorithm.

## 2 Location Routing Problem

LRP jointly considers the facility location problem (FLP) and the vehicle routing problem (VRP). Watson-Gandy and Dohrn [14], Salhi, S., & Rand, G. K. [15] have proven that the strategy for solving LRP by dismantling LRP into FLP and VRP and solving those problems sequentially is not optimal. Therefore, constructing the solving strategy that could consider FLP and VRP as a unity and solve them simultaneously to find the optimal solution is critical in processing LRP. Currently, the exact and heuristic algorithms are mainly applied to solve LRP.

LRP is a traditional strategic-tactical-operational problem that considers a set of potential facilities and a set of customers. The main decisions of LRP are:

- The number and location of facilities to open,
- the allocation of customers to the opened facilities,
- The design of routes to serve customers of each facility using a fleet of vehicles.

As with most other models, one cannot capture all aspects of a real-life LRP with one mathematical model. Considering the complex real-life scenarios, LRP has a lot of variant problems with different objectives and constraints. Therefore, identifying the constraints of each sub-problem, such as facility location, allocation, and routing problem, is essential for solving LRP.

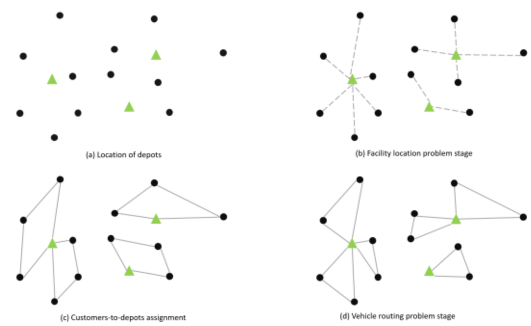


**Figure 1: An illustrative example of LRP**

Notations:
V: the set of the nodes, $V = I \cup J$.
I: the set of the potential depot nodes.
J: the set of the customer nodes.
E: the set of the edges.
K: the set of the vehicles
Impact factors:
$O_i$: Fixed cost of opening the depot node i.
$Q_i$: The depot capacity of node i.
$\mathrm{D_j}$: Customer requirements for node i.
$q$: The loading capacity of each vehicle
F: Fixed cost oer vehicle
$C_{ij}$: Travel cost at the edge $(i, j)$
$d$: The maximum distance allowed per vehicle.
$U_i$: Any real number.

Decision variables:

$$x_{ijk} = \begin{cases} 1, if\ Vehicle\ k\ covered\ the\ edge\ (i,j). \\ 0, otherwise. \end{cases}$$

$$y_i = \begin{cases} 1, if\ the\ depot\ located\ in\ node\ i. \\ 0, otherwise. \end{cases}$$

$$z_{ij} = \begin{cases} 1, if\ customer\ j\ is\ serviced\ by\ depot\ i. \\ 0, otherwise. \end{cases}$$

The MILP formulation of LRP is as follows [16]:

$$\min \quad \sum_{i\in I} O_i\, y_i + \sum_{i\in V}\sum_{j\in V}\sum_{k\in K} c_{ij}\, x_{ijk} + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K} F\, x_{ijk} \quad (1)$$

Subject to

$$\sum_{i\in V}\sum_{k\in K} x_{ijk} = 1 \quad \forall j \in J \quad (2)$$

$$\sum_{i\in I}\sum_{j\in J} x_{ijk} \le 1 \quad \forall k \in K \quad (3)$$

$$\sum_{j\in V} x_{ijk} - \sum_{j\in V} x_{jik} = 0 \quad \forall k \in K,\ \forall i \in V \quad (4)$$

$$U_i - U_j + (n-1)x_{ijk} \le n-2 \quad \forall i \in J,\ \forall j \in J, i \ne j \quad \forall k \in K \quad (5)$$

$$\sum_{u\in J} x_{iuk} + \sum_{u\in V\{j\}} x_{ujk} \le 1 + z_{ij} \quad \forall i \in I,\ \forall j \in J,\ \forall k \in K \quad (6)$$

$$\sum_{i\in V}\sum_{j\in J} D_j\, x_{ijk} \le q \quad \forall k \in K \quad (7)$$

$$\sum_{j\in J} D_j\, z_{ij} \le Q_i y_i \quad \forall i \in I \quad (8)$$

$$\sum_{i\in V}\sum_{j\in V} c_{ij}\, x_{ijk} \le \bar{d} \quad \forall k \in K \quad (9)$$

$$\sum_{i\in I}\sum_{j\in J}\sum_{k\in K} x_{ijk} \le |K| \quad (10)$$

$$x_{ijk} \in \{0,1\} \ \forall i \in V,\ \forall j \in V,\ \forall k \in K \quad (11)$$

$$y_i \in \{0,1\} \ \forall i \in I \quad (12)$$

$$z_{ij} \in \{0,1\} \ \forall i \in I,\ \forall j \in V \quad (13)$$

$$U_i \in R \quad \forall i \in J \quad (14)$$

The objective function (1) aims to minimize the total cost, which includes the opening cost of warehouses, variable distance costs of vehicles, and fixed usage costs of vehicles. Constraint (2) ensures that each customer can only be visited by one vehicle, enforcing the routing constraint. Constraint (3) ensures that all routes must start and end at the warehouse, while constraint (4) is a connectivity constraint to ensure that each vehicle leaves each customer after service. The sub-tour elimination constraint is defined by constraint (5). Constraint (6) ensures that a customer can only be assigned to a warehouse if there is a route to that warehouse. Constraints (7) and (8) respectively specify the capacity of each vehicle and the capacity of each warehouse should not be exceeded. Constraint (9) ensures that the length of each vehicle's route does not exceed the maximum distance constraint. Constraint (10) limits the number of vehicles used. Constraints (11) to (14) define the decision variables.

# 3 METHODOLOGY

In the Location Routing Problems (LRPs), there are two main problems to address: location problem and route planning. Both of these problems can be modeled as sequential decision processes. In this section, we propose a two-stage algorithm based on deep reinforcement learning (DRL) to solve LRP, comprising the location problem stage and the route planning stage. In the location problem and routing problem, there are three types of nodes: demand points, logistics centers, and distribution centers.

## 3.1 Location Problem Stage

In the location selection stage, we model the problem as a Markov Decision Process (MDP). The intelligent agent needs to choose one or more locations from all demand points to construct facilities, aiming to maximize overall revenue or minimize overall costs. To achieve this, we follow these steps:

- State: Define a state representation that includes the current chosen facility locations, the status of demand points, and other relevant information.
- Action: Define an action space in which the agent can choose the locations for constructing facilities.
- Reward: Design a reward function so that the agent can receive appropriate reward signals based on the chosen locations. Rewards can consider factors such as costs, revenues, and the degree of demand fulfillment.

Once this stage is completed, the facilities' deployment locations are determined, and these points are considered as central points, with each demand point assigned to the nearest facility point.

## 3.2 Route Planning Stage

In the route planning stage, the routing problem can also be modeled as a Markov Decision Process. In this stage, the intelligent agent needs to decide how to allocate deliveries between the selected facility locations to minimize the overall path length or costs.

The two-stage algorithm's solution framework is illustrated as follows:

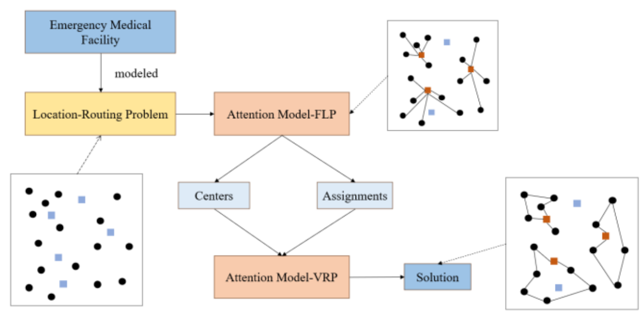


**Figure 2: The structure of the two-stage algorithm for solving emergency medical facility location problem.**

## 3.3 Deep Reinforcement Learning Algorithm

We are using the REINFORCE algorithm [17] to train our Attention Model-Facility Location Problem (AM-FLP) and Attention Model-Vehicle Routing Problem (AM-VRP). This algorithm is a type of policy gradient method. The algorithm steps are as follows:

**Define Policy Functions $\boldsymbol{\pi}(\boldsymbol{a}|\boldsymbol{s})$:** Define a policy function $\boldsymbol{\pi}(\boldsymbol{a}|\boldsymbol{s})$ that describes the probability distribution of taking action $\boldsymbol{a}$ given a state $\boldsymbol{s}$. We use $\boldsymbol{\pi}_{FLP}(\boldsymbol{a}|\boldsymbol{s})$ and $\boldsymbol{\pi}_{VRP}(\boldsymbol{a}|\boldsymbol{s})$ to represent the policy functions for the AM-FLP and AM-VRP, respectively.

**Sample Trajectories:** For each trajectory, calculate the cumulative return. The return reflects the quality of executing the policy during a single attempt.

**Compute Returns:** For each trajectory, compute the cumulative return. The return represents the quality of executing the policy in a single attempt.

**Update Policy:** Based on the computed gradient information, employ gradient descent to update the parameters of the policy function, with the goal of increasing the expected return.

**Repeat Training:** Continuously repeat the above steps until the policy function converges to a satisfactory policy or reaches a predefined number of training iterations**.**

# 4 EXPERIMENTS AND ANALYSIS

Deep reinforcement learning, through continuous interaction between an agent and its environment to maximize rewards, is effective in handling sequential decision problems. We propose a new approach to the two-stage location routing problem (LRP) based on deep reinforcement learning. In this section, we first discuss how to utilize deep reinforcement learning to solve LRP, then implement problem-solving using this framework, and finally conduct generalization experiments.

For the LRP, we constructed two different attention models, AM-FLP and AM-VRP, for solving the location problem and routing problem, respectively. We randomly generated 2000 instances for evaluating results. Each example consists of 50 demand points, and the task is to select 8 out of the 50 as central points. Then, following a greedy approach, the remaining demand points are assigned to the nearest central points, classifying all nodes into 8 categories. Finally, the shortest path is chosen for each category to traverse all the nodes. The experimental results are shown in Table 1. N represents the number of the demand points and p presents the number of the centers.

**Table 1 The comparison results of Gurobi vs. GA vs. AM for location routing problem on N=50, p=8**

| Method | Location stage | | Routing stage | |
|---|---|---|---|---|
| | obj | time | obj | time |
| Gurobi | 5.32 | 0.1292 | 10.28 | 0.3824 |
| GA | 5.45 | 0.1021 | 11.30 | 0.3214 |
| AM | 5.38 | 0.0210 | 10.62 | 0.1679 |

We divided the solution of the location-routing problem into two stages. Table 1 illustrates three different methods in terms of their objective functions and solution times in the location selection stage and the path planning stage. Gurobi, being recognized as the best solver, performs exceptionally well in this problem. It consistently obtains optimal solutions quickly, even in instances with 50 nodes. On the other hand, Genetic Algorithm (GA) is a heuristic algorithm, and its results are relatively less favorable compared to the other two algorithms. The gap between its objective values and the optimal solutions is the widest, which may be attributed to the parameter settings within the genetic algorithm. Our proposed method, which employs deep reinforcement learning (AM), exhibits a smaller gap from the optimal solutions compared to GA. Furthermore, it boasts the shortest solution times among the three methods. This suggests that our proposed approach achieves a balance between solution time and accuracy to some extent, making it a highly promising method. Overall, the results demonstrate that our proposed approach has the potential to be a valuable contribution in addressing the trade-off between solution time and precision in the context of the location-routing problem. The source code for this study is available at: https://github.com/HIGISX/hispot.

# 5 CONCLUSION AND FUTURE WORKS

The location routing problem holds significant importance in practical production. It directly impacts resource allocation, logistics efficiency, and service quality for businesses, making it crucial for their operations and competitiveness. This study has introduced an innovative solution—a two-stage approach based on deep reinforcement learning—to address the location routing problem, with a specific focus on emergency medical facility selection. This method leverages the full potential of deep reinforcement learning in tackling sequential decision problems. Through the interaction between the agent and the environment, it effectively optimizes the selection of emergency medical facility locations and route planning, enhancing overall efficiency and quality. This research provides a new and efficient approach for the selection of emergency medical facilities, holding significant promise for practical applications and offering robust support and decision-making tools for addressing emergencies and disasters. Future studies could further explore the method's generalization capabilities and its potential applications in other domains, thereby extending its practical value.

# ACKNOWLEDGMENTS

The research was financially supported by the innovation group project of the Key Laboratory of Remote Sensing and Digital Earth Chinese Academy of Sciences (E33D0201-5), the CBAS project 2023, and the Beijing Chaoyang District Collaborative Innovation Project (E2DZ050100).

# REFERENCES

[1] Kirsch K R, Newman G D, Zhu R, et al. Applying and integrating urban contamination factors into community garden siting[J]. Journal of Geovisualization and Spatial Analysis, 2022, 6(2): 33.
[2] Vahidnia M H, Vahidi H, Hassanabad M G, et al. A Spatial Decision Support System Based on a Hybrid AHP and TOPSIS Method for Fire Station Site Selection[J]. Journal of Geovisualization and Spatial Analysis, 2022, 6(2): 30.
[3] Church R L, Wang S. Solving the p-median problem on regular and lattice networks[J]. Computers & Operations Research, 2020, 123: 105057.
[4] Liang H, Wang S, Li H, et al. A Trade-Off Algorithm for Solving p-Center Problems with a Graph Convolutional Network[J]. ISPRS International Journal of Geo-Information, 2022, 11(5): 270. https://doi.org/10.3390/ijgi11050270
[5] Wang S, Liang H, Zhong Y, et al. DeepMCLP: Solving the MCLP with Deep Reinforcement Learning for Urban Facility Location Analytics[J]. 2023. https://doi.org/10.25436/E2KK5V
[6] Salhi S, Rand G K. The effect of ignoring routes when locating depots[J]. European Journal of Operational Research, 1989,39(2):150-156.
[7] COOPER L. The transportation-location problem[J]. Operations Research,1972,20 (1):94-108.
[8] Foroughi A, Gokcen H. An multiple rule-based genetic algorithm for cost-oriented stochastic assembly line balancing problem[J]. Assembly Automation,2019,39(1):124-139.
[9] Laporte G, Nobert Y. An exact algorithm for minimizing routing and operating costs in depot location[J]. European Journal of Operational Research, 1981,6(2):224-226.
[10] Belenguer J M, Benavent E, Prins C, et al. A branch-and-cut method for the capacitated location-routing problem[J]. Computers & Operations Research,2011,38(6): 931-941.
[11] Contardo C, Cordeau J F, Gendron B. An exact algorithm based on cut-and-column generation for the capacitated location-routing problem[J], INFORMS Journal on Computing,2014,26(1):88-102.
[12] Ferrira K M, de Queiroz T A. Two effective simulated annealing algorithms for the location-routing problem[J]. Applied Soft Computing,2018,70:389-422.
[13] Schneider M, Loffer M. Large composite neighborhoods for the capacitated location-routing problem[J]. Transportation Science,2019,53(1):301-318.
[14] Watson-Gandy, C. D. T., & Dohrn, P. J. (1973). Depot location with van salesmen—a practical approach. Omega, 1(3), 321-329.
[15] Salhi, S., & Rand, G. K. (1989). The effect of ignoring routes when locating depots. European journal of operational research, 39(2), 150-156.
[16] Almouhanna A, Quintero-Araujo C L, Panadero J, et al. The location routing problem using electric vehicles with constrained distance[J]. Computers & Operations Research, 2020, 115: 104864.
[17] Williams, R. J. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, *8*, 229-256.