Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/69/20241522

Optimization of logistics cargo tracking and transportation efficiency based on data science deep learning models

Ang Li[1,*], Shikai Zhuang[2], Tianyi Yang[3], Wenran Lu[4], Jiahao Xu[5]

[1]Business Analytics, University College Dublin, Dublin, Ireland
[2]Electrical Engineering, University of Washington, Seattle, WA, USA
[3]Financial Risk Management, University of Connecticut, Stamford CT, USA
[4]Electrical Engineering, University of Texas at Austin, Austin, TX, USA
[5]Master of Science in Financial Engineering, University of Southern California, CA, USA

*Corresponding author E-mail: lalala19940211@gmail.com

Abstract. With the digital transformation of the logistics industry, smart logistics algorithms have become a core technology to improve efficiency and reduce costs. This paper reviews the development history of traditional logistics technology and discusses the key role of technologies such as the Internet of Things, big data analysis, artificial intelligence, and automation in logistics technology innovation. It focuses on the application of intelligent logistics algorithms in path optimization, intelligent scheduling, data mining and prediction, and intelligent warehousing. To solve the problem of inconsistency between training and testing objectives, this paper proposes DRL4Route, a deep reinforcement learning-based path optimization framework, and designs the DRL4Route-GAE model. These research results provide important support to further promote the intelligent development of the logistics industry.

Keywords: Intelligent logistics algorithms, path optimization, deep reinforcement learning, data mining, transport efficiency optimization

1. Introduction

With the digital transformation of the logistics industry, intelligent logistics algorithms have become a core technology for improving efficiency and reducing costs. By leveraging advanced technologies such as Big Data and Artificial Intelligence, logistics companies can optimize logistics data for intelligent decision-making and efficient operations. [1] Intelligent logistics algorithms play a crucial role in several ways. Firstly, regarding route planning, these algorithms can significantly reduce transport time and costs and improve transport efficiency. Secondly, intelligent logistics algorithms can optimize supply chain management through demand forecasting and inventory management by reducing inventory costs and the risk of stock-outs.

In addition, real-time monitoring of logistics links and quick response to abnormal situations is another essential function of intelligent logistics. This paper aims to explore strategies and methods for optimizing logistics cargo tracking and transport efficiency based on data science and deep learning models to further promote the intelligent development of the logistics industry.

© 2024 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/69/20241522

2. Related Work

2.1. Traditional logistics technologies

With the acceleration of digital transformation and the rise of the intelligent logistics era, logistics technology innovation has emerged. Digital transformation enables companies to achieve a fundamental shift in their business model through digital technologies, greatly enhancing the logistics process's efficiency and accuracy. [2] In this, the availability of real-time data and collaboration between supply chain segments are critical to logistics technology innovation. For example, real-time collaboration mechanisms in digital platforms enable the sharing of supply chain information and improve overall operational efficiency.

Looking back at traditional logistics technology, historical data shows that logistics companies in the late 20th century mainly relied on manual records and basic computer systems to manage transport and inventory. Under this model, information transfer is slow and error-prone, leading to inefficiency [3].

Entering the 21st century, the advent of the intelligent logistics era has promoted the further development of logistics intelligence. In this, the application of advanced technologies such as artificial intelligence and the Internet of Things is an important direction of logistics technology innovation. Intelligent decision-making and automation have become the core features of the intelligent logistics era. For example, Amazon optimizes the transport path with its intelligent path planning system and realizes intelligent inventory management, which significantly reduces logistics costs and improves transport efficiency. [4] The introduction of barcode, RFID, EDI, and other technologies has laid a solid foundation for the logistics industry. In contrast, the application of emerging technologies such as artificial intelligence, the Internet of Things, blockchain, and edge computing has injected new vitality into logistics technological innovation and promoted the sustainable development of the logistics industry. This series of technological advances and innovations reflects the logistics industry's huge potential and infinite possibilities in the digital era.

2.2. AI-Drives Logistics Technology Innovation

Logistics technology innovation is driven by various technologies, of which IoT, big data analytics, artificial intelligence, and automation are the most critical factors. IoT technology enables real-time monitoring of logistics processes by connecting devices and sensors, providing data support for intelligent decision-making. For example, [5] IoT temperature monitoring systems ensure the safety of perishable goods during transport by tracking the temperature of goods in real-time. According to McKinsey's report, enterprises applying IoT technology can increase efficiency by 10-15 per cent in logistics operations.

Big data technology also plays a vital role in logistics technology innovation. Logistics companies can conduct predictive analysis and real-time monitoring by processing and analyzing massive amounts of data. Big data technology can use historical data to predict market demand, thereby optimizing transport plans and identifying and dealing with problems in real time in the logistics chain. For example, extensive data-driven demand forecasting systems enable logistics companies to predict market demand more accurately and improve supply chain efficiency.

By combining IoT, big data, artificial intelligence, and automation technologies, logistics companies can realize a full range of intelligent and automated operations, thereby maintaining a leading position in the fiercely competitive market. Actual data and literature reviews show that with these technologies' continuous development and improvement, the logistics industry will usher in a more efficient, safe, and intelligent future.

2.3. Intelligent Logistics Algorithms

Intelligent logistics algorithms are based on sensing technology and data processing. Sensing technology ensures real-time data collection, and data processing transforms the data into valuable information that provides the basis for the algorithm [6]. The importance of sensing technology is reflected in three aspects: real-time data acquisition, sensors capture the location of goods, temperature and humidity,

DOI: 10.54254/2755-2721/69/20241522

vehicle status, and other data to provide a transparent view; environmental monitoring, such as temperature sensors to ensure the quality of cold chain logistics goods; location tracking, GPS real-time tracking of goods and vehicle location, optimize path planning, to ensure on-time delivery.

1. Path optimization algorithm

The path optimization algorithm is a critical technology in graph theory and network analysis, aiming to find the optimal or shortest path in each network. By defining the path encoding and fitness function, the initial path solution population is randomly generated, the crossover and mutation operations produce new path solutions, and the excellent individuals are selected for reproduction according to the fitness function. [7]The iteration continues until the stopping condition is satisfied, and the globally optimized path is obtained.

2. Intelligent Scheduling Algorithm

Intelligent scheduling algorithms refer to optimizing resource allocation and task scheduling to improve logistics efficiency significantly; commonly used methods include genetic, simulated annealing, and ant colony algorithms. [8]Logistics scheduling helps optimize distribution routes or task-scheduling orders in a global search. The ant colony algorithm is inspired by the foraging behaviour of ant colonies and guides the search process through pheromone release and update. In logistics scheduling, it optimises distribution routes and simulates cooperative driving between vehicles. The strengths of these algorithms are global search capability, adaptability, and ability to handle complex problems.

3. Data Mining and Predictive Logistics Algorithms

Data mining and predictive logistics algorithms are crucial in modern logistics management. They explore data associations, reveal time series patterns, and use machine learning to provide powerful prediction and optimizations tools for logistics [9]. Optimism transport routes considering traffic, weather, and other factors to improve delivery efficiency and identify supply chain risks. Data mining and predictive logistics algorithms improve logistics operation efficiency, reduce costs, provide more accurate forecasts, and enhance supply chain competitiveness.

4. Intelligent Warehousing Algorithms

Intelligent warehousing algorithms are a key technology in modern logistics, and they significantly improve warehousing efficiency by integrating advanced technologies and algorithms. Automation technologies such as robots and automated lifting tables enable automated storage and retrieval, reducing manual handling and increasing speed and accuracy. Algorithms optimise warehouse layout to maximise space, reduce costs, and increase storage density.

3. Methodology

Learning-based approaches have been used to solve collection and delivery route prediction problems. These methods typically use deep neural networks to learn courier route patterns from large amounts of historical data in a supervised manner. Despite the results achieved by these deep learning-based methods, all have some limitations: the training and testing objectives need to be aligned, limiting the performance of these methods. Inconsistency between the training and testing objectives will limit the performance of a trained model, so we will seek a practical solution that allows the model to differentiate between the two cases of inconsistency in the testing objectives mentioned above during training.

3.1. Methodological model

An intuitive solution is to transform the test targets into loss functions to update the model parameters. However, this approach is infeasible because the test targets in this task, such as the position squared deviation and the Kendall rank correlation coefficient, are non-microscopic. In recent years, reinforcement learning has been widely used to optimise non-differentiable test objectives in various tasks, such as machine translation, text summarization, image description, etc., achieving better results than supervised deep learning methods.

DOI: 10.54254/2755-2721/69/20241522

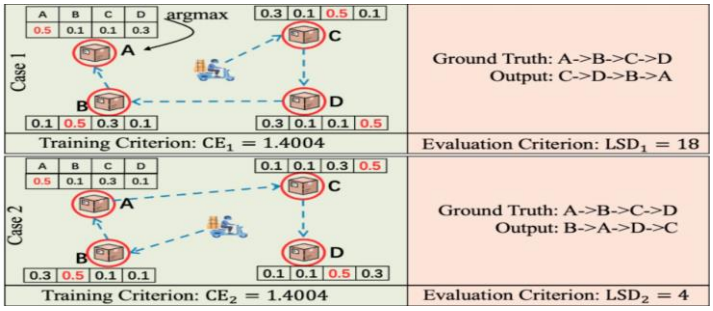


Figure 1. Illustration of the mismatch between the training and test objectives. The transition probability corresponding to A, B, C, and D is the vector near each location.

This paper is the first to model the route prediction problem of dispatching from the perspective of reinforcement learning and to construct the training framework DRL4Route using a policy-based reinforcement learning approach, which is based on the rewards computed from non-differentiable test objectives optimizing the deep neural network through a policy gradient approach to solving the problem of inconsistency between the training and the test objectives [10].

Overall, the contributions of this paper are as follows:

1. the first time the task of collection and delivery route prediction is solved from a reinforcement learning perspective and proposes a reinforcement learning-based framework, DRL4Route, which combines the ability of reinforcement learning in optimising an undifferentiable objective function and the ability of deep neural networks to learn historical behavioural patterns, compared to previous supervised learning approaches.

2. propose the DRL4Route-GAE model for route prediction in logistics scenarios, which utilises an actor-critic framework to compute rewards at each decoding step designed by the test objective to guide model training and computes an approximation of the dominance using a generalised dominance estimation to balance the bias and variance during gradient estimation.

3. Extensive offline experiments on real datasets as well as online deployments validate the effectiveness of the method proposed in this paper. Compared to the optimal baseline method, DRL4Route-GAE improves the location bias squared metrics by 0.9%-2.7% and the top three location prediction accuracy metrics by 2.4%-3.2%.

3.2. DRL4Route

The DRL4Route (Deep Reinforcement Learning for Route Optimization) model is a route optimization algorithm based on deep reinforcement learning, which is specifically used to solve complex logistics route planning problems. The model can autonomously learn and optimize transport routes by using reinforcement learning strategies, thus effectively reducing transport time and cost. This allows us to optimize the non-differentiable objective function for more accurate route prediction.

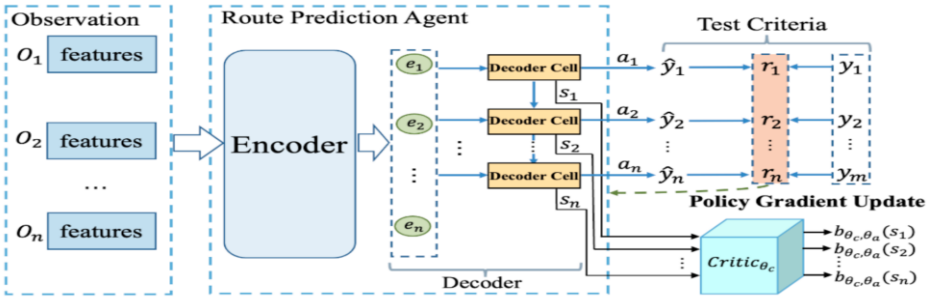


Figure 2. DRL4Route Framework.

DOI: 10.54254/2755-2721/69/20241522

In this paper, based on the DRL4Route framework, a model named DRL4Route-GAE is proposed for parcel collection services in logistics scenarios to illustrate the effectiveness of the proposed framework.The model training is guided by the strategy gradient so we can optimize the non-differentiable test objective to solve the problem of inconsistency between the training and test objectives, furthermore, we use a generalized dominance estimation method to compute an approximation of the dominance to balance the bias and variance during gradient estimation, so we can get better strategies as well as better results.

3.3. Experimental data sets and methods

We conduct offline experiments on the logistics parcel-acquisition dataset provided by Cainiao, and the sample ratio of the training set, validation set, and test set is about 6:2:2, and the data statistical information is shown in Table 1.

Table 1. Experimental Data Details for Logistics Parcel-Acquisition Dataset

Type	Time Range	City	ANUT	#Workers	#Samples
Logistics-I1Z	07/10/2021-10/10/2021	Ilangzhou		1,117	373,072
Logistics-SI1	03/29/2021-05/27/2021	Shanghai		2,344	208,202

Baseline methods

This paper implements a variety of baseline methods, as well as some state-of-the-art deep models in different scenarios (food distribution and last mile logistics) for comparison:

Time Greedy: Generate paths by ranking all locations to be visited based on the time remaining until the task timeout.

Distance Greedy: Selects the closest location to the courier to visit each time and completes the path generation step by step.

OR-Tools: a heuristic algorithm to find the shortest path.

OSquare: an XGBoost-based method that generates the route one node at a time.

FDNET: A path and time prediction model for takeaway scenarios using LSTM and attention mechanisms.

DeepRoute: A deep learning route prediction model based on a Transformer encoder and attention decoder.

DeepRoute+: Based on DeepRoute, a courier decision preference modeling module is added to the coding.

Graph 2Route: For the first time, a graph is used to model the locations to be visited by couriers, and a GCN-based encoder and attention mechanism are used to complete the path prediction.

3.4. Experimental results

Table 2. Experimental result

Method	Logistics-HZn E(0.25]	Logistics-HZ n ∈ (0.11]	Logistics-SH n E(0.25]	Logistics-SH n ∈ (0.11]
	HR@I	ACC@3	KRC	LMD
Time-Greedy	33.15	20.32	41.92	1.7
Distance-Greedy	33.13	51.82	136	5.73
OR-Tools	53.93	1.23	4.68	1.46
OSquare	54	33.1	58.5	1.16
FDNET	52.76	33.22	55.47	1.18
DeepRoute	54.76	34.64	58.61	1.1

DOI: 10.54254/2755-2721/69/20241522

Table 2. (continued).

DeepRoute+	55.42	35.63	59.32	1.08
Graph2Route	56.45	36.12	60.63	1.05
DRLAROute-REINFORCE	55.88	29.02	59.97	1.38
DRL4Route-AC	56.36	36.16	60.86	1.05
DRL4Route-GAE	57.72	37.23	61.47	1.03
Improvement	2.20%	3.10%	1.40%	1.90%

As can be seen from the results on the two datasets in Table 2, the distance-based greedy or time-greedy methods, as well as OR-Tools, only consider distance or time alone, and their results are only optimal solutions under a particular strategy, ignoring many patio-temporal constraints. The tree-based model Square cannot model spatiotemporal correlations, and in addition, its goal is to maximize the output probability of the next position rather than the entire route. Sequence-based models Deep Route and FDNET have difficulty in modeling neighborhood relationships between locations, and thus they may provide unreasonable outputs.

To address the above problems, we apply reinforcement learning methods to the collection and delivery path prediction tasks for the first time and propose a new framework called DRL4Route. It combines the current behavioral pattern learning capabilities of deep neural networks with reinforcement learning's ability to optimize non-micro-objectives. Furthermore, DRL4Route can be used as a plug-and-play component to improve the performance of existing deep models. Based on this framework, we implement a reinforcement learning-based model, DRL4Route-GAE, for solving the route prediction problem in logistics, which is based on an actor-critic architecture and uses a generalized dominance estimation method that balances the bias and variance of the strategy gradient estimation. We experimentally demonstrate that DRL4Route achieves significant improvements over the most competitive baseline models on real datasets.

4. Conclusion

This paper reveals the importance of intelligent logistics algorithms in modern logistics systems by reviewing the development history of traditional logistics technologies and exploring the key roles of the Internet of Things (IoT), big data analytics, artificial intelligence (AI), and automation technologies in logistics technology innovation. In particular, the proposal and application of DRL4Route, a deep reinforcement learning-based path optimization framework, demonstrates its significant advantages in improving logistics path optimization, intelligent scheduling, and prediction accuracy through extensive offline experiments and online deployments on real datasets.

The adoption of edge computing will accelerate real-time response, while human-machine collaboration technologies will enhance cooperation between humans and intelligent systems to jointly improve logistics efficiency. Emerging technologies such as 5G, IoT, and edge computing will be applied on a large scale in the logistics sector. At the same time, the logistics industry will be committed to sustainability. In the future, intelligent green port management systems and circular economy logistics networks will become a reality, driving smarter and more efficient logistics systems and promoting the industry toward more sustainable and environmentally friendly development.

References

[1] Oliveira, R. R., Cardoso, I. M., Barbosa, J. L., Da Costa, C. A., & Prado, M. P. (2015). An intelligent model for logistics management based on geofencing algorithms and RFID technology. Expert Systems with Applications, 42(15-16), 6082-6097.

DOI: 10.54254/2755-2721/69/20241522

[2] Kumar, R. S., Rani, C., & Kumar, P. G. (2018). Design of smart logistics transportation system using MapReduce intelligent water drops algorithm in Hadoop environment. International Journal of Logistics Systems and Management, 31(2), 249-266.
[3] Mohamed M. Toward Smart Logistics: Hybridization of Intelligence Techniques of Machine Learning and Multi-Criteria Decision-Making in Logistics 5.0[J]. Multicriteria algorithms with applications, 2023, 1: 42-57.
[4] Singh M K, Parhi D R. Path optimization of a mobile robot using an artificial neural network controller[J]. International Journal of Systems Science, 2011, 42(1): 107-120.
[5] Dhand, A., Lang, C. E., Luke, D. A., Kim, A., Li, K., McCafferty, L., ... & Lee, J. M. (2019). Social network mapping and functional recovery within 6 months of ischemic stroke. Neurorehabilitation and neural repair, 33(11), 922-932.
[6] Sarkis, R. A., Goksen, Y., Mu, Y., Rosner, B., & Lee, J. W. (2018). Cognitive and fatigue side effects of anti-epileptic drugs: an analysis of phase III add-on trials. Journal of Neurology, 265(9), 2137-2142.
[7] Pióro M, Szentesi Á, Harmatos J, et al. On open shortest path first related network optimization problems[J]. Performance evaluation, 2002, 48(1-4): 201-223.
[8] Arulkumaran K, Deisenroth M P, Brundage M, et al. Deep reinforcement learning: A brief survey[J]. IEEE Signal Processing Magazine, 2017, 34(6): 26-38.
[9] Rosner, B., Tamimi, R. M., Kraft, P., Gao, C., Mu, Y., Scott, C., ... & Colditz, G. A. (2021). Simplified breast risk tool integrating questionnaire risk factors, mammographic density, and polygenic risk score: development and validation. Cancer Epidemiology, Biomarkers & Prevention, 30(4), 600-607.
[10] Li, Shengbo Eben. "Deep reinforcement learning." Reinforcement learning for sequential decision and optimal control. Singapore: Springer Nature Singapore, 2023. 365-402
[11] Allman, R., Mu, Y., Dite, G. S., Spaeth, E., Hopper, J. L., & Rosner, B. A. (2023). Validation of a breast cancer risk prediction model based on the key risk factors: family history, mammographic density, and polygenic risk. Breast Cancer Research and Treatment, 198(2), 335-347.