Proceedings of the 4th International Conference on Computing Innovation and Applied Physics
DOI: 10.54254/2753-8818/2025.23342

A New Integer Linear Programming Model of Transportation Problem

Yan Dong[1*], Yixuan Li[2], Jiaming Zuo[3], Jin Yi[4]

[1]*University of Birmingham, School of Mathematics, Edgbaston, Birmingham, United Kingdom*
[2]*Wuhan Britain-China School, Wuhan, China*
[3]*Lanzhou University of Technology, School of Science, Lanzhou, China*
[4]*Wuhan No.15 Middle School, Wuhan, China*
[*]*Corresponding Author. Email: yxd406@student.bham.ac.uk*

Abstract: The transportation problem has been a central focus in operational research, to determine the most cost-effective routes for moving goods or people. With the continuous development of modern logistics and the increasing complexity of supply chain management, optimizing transportation systems has become more crucial than ever. This paper introduces a novel approach that combines classical transportation problems with Minimum Spanning Tree (MST) techniques, enabling a more efficient way to address real-world logistics issues. Our model incorporates a topological graph that includes not only factories and retail stores but also intermediate warehouses to balance the flow of goods. To solve this problem efficiently, we propose a new Intger linear programming (ILP) model with flow balance constraints and using a CPLEX solver to get the numerical results. Through numerical experiments based on China's automobile transportation industry, we demonstrate the practical applicability of our approach in minimizing transportation costs. Additionally, we discuss future research directions, including the integration of dynamic demand and production capabilities.

Keywords: transportation problem, ILP model, Minimum Spanning Tree, CPLEX solver

1. Introduction

1.1. Background

The transportation problem, which originated from the need to move goods or people from one location to another in daily life, has become a fundamental issue in operations research. The challenge is to determine the most economical or least costly transportation route or scheme. With the continuous development of the global economy and the rapid expansion of the modern logistics industry, maximizing the efficient use of time, information, warehousing, distribution, and intermodal transportation systems to create additional value has posed an increasing challenge to operations research. Although the complexity of organizing the sourcing, transport, and distribution of goods has grown, the basic objective remains the same: achieving optimal allocation of available resources.

© 2025 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: 10.54254/2753-8818/2025.23342

With the rapid development of the market economy, China has actively introduced foreign capital and supported the growth of local industries, leading to remarkable achievements in the country's manufacturing sector. As highlighted by Li [1], the transportation industry has not only been thoroughly examined in terms of its current state but also through an analysis of the challenges facing automobile transportation, with proposed management countermeasures. As a fundamental, leading, and strategic industry within the national economy, China's transportation sector is now facing challenges of transformation, upgrading, and improving quality and efficiency. Furthermore, efforts to enhance ecological civilization and accelerate green development have become critical components of the supply-side structural reform in the transportation industry.

The relationship between transportation and economic development can be summarized as follows: (1) The transportation economy is a vital part of the national economy, and the development of transportation infrastructure plays a key role in advancing China's socialist economic construction. With the continuous expansion of modern transportation networks, road transport has become an essential medium for integrating and facilitating exchanges between cities. In developing cities, transportation is more flexible and convenient compared to other modes due to its unique role, enabling the smooth circulation of goods within the city by Liu [2]. (2) Transportation is also the foundation for socio-economic development. In many economically underdeveloped remote areas, poor transportation is often a significant issue, whereas in economically developed coastal regions, robust traffic infrastructure has brought tremendous economic benefits to local populations. Clearly, transportation plays an irreplaceable role in national economic growth, and solving transportation problems is a prerequisite for regional development.

As one of the most widely used transportation methods, road transport is characterized by its flexibility, convenience, and speed. China's ongoing efforts to improve and expand its transportation network underscore the increasing importance of transportation in the country's overall economic development by Yue [3].

The general transportation problem revolves around determining how to transfer goods from multiple production areas to multiple sales areas while minimizing total transportation costs. This is achieved by ensuring that the supply from each production area, the demand in each sales area, and the transportation costs between locations are known. Several conditions are essential for addressing transportation issues: (1) clearly defining the origin, destination, supply, demand, and unit cost; (2) assuming that each origin has a set supply and all of it must be delivered to the destination, while each destination has a fixed demand that must be met entirely by the origins (i.e., "total supply = total demand"); and (3) assuming that transportation costs are linearly proportional to the quantity delivered, meaning that the total cost equals the unit cost of delivery multiplied by the quantity transported, making the objective function linear by Fan [4].

DOI: 10.54254/2753-8818/2025.23342

To address the transportation problem, the researchers developed and solved it by formulating a linear model by Kacher [5]:

$$\min \quad F = \sum_{j=1}^{n} \sum_{i=1}^{m} c_{ij} x_{ij} \tag{1a}$$

$$\sum_{j=1}^{n} x_{ij} = a_i \quad \forall i = 1, 2, ..., m \tag{1b}$$

$$\sum_{i=1}^{m} x_{ij} = b_j \quad \forall j = 1, 2, ..., n \tag{1c}$$

$$\sum_{i=1}^{m} a_i - \sum_{j=1}^{n} b_j = 0 \tag{1d}$$

$$x_{ij}^{f} \in \mathbb{N} \quad \forall i \& j \tag{1e}$$

In this model, constraint (1a) represents the objective, which is to minimize costs. Constraint (1b) specifies that the total quantity of products delivered from all factories must equal the factory's total output. Constraint (1c) ensures that the total quantity received by each retail store meets its demand. Constraint (1d) guarantees that the total factory output matches the overall demand. Finally, constraint (1e) requires that the minimum unit of product transmission is 1.

According to data provided by the China Association of Automobile Manufacturers (CAAM) , in June 2024, China's automobile production and sales reached 2.507 million and 2.552 million units, representing year-on-year declines of 2.1% and 2.7% [14], respectively. Of these, passenger vehicle production and sales totaled 2.177 million and 2.215 million units, down by 1.9% and 2.3% year-on-year. Commercial vehicle production and sales were 330,000 and 337,000 units, decreasing by 3.5% and 4.9% year-on-year.Despite the overall decline in automobile sales, new energy vehicles (NEVs) showed remarkable growth. In June, NEV production and sales reached 1.003 million and 1.049 million units, reflecting year-on-year growth of 28.1% and 30.1%. NEVs accounted for 41.1% of all new vehicle sales during the month. In the first half of 2024 (January to June), cumulative NEV production and sales amounted to 4.929 million and 4.944 million units, both up by more than 30% year-on-year. Additionally, exports showed strong performance as well. In June, China exported 485,000 complete vehicles, representing a year-on-year increase of 26.9%.

1.2. Related Work

So far, Many researchers have proposed models and algorithms for different transportation problems. We will briefly introduced the related work in those problems.

Kacher [5] tackles the broad issue of optimizing transportation systems and proposes models for resource allocation and routing. They propose several ILP models in order to minimize both single-objective and multi-objective problems under deterministic and stochastic conditions. They also establish some algorithms including dynamic programming algorithm and genetic algorithms. These methods are particularly useful for large, complex problems where exact solutions are impractical due to high computational demands. However, a major limitation of these models is that they lack real-world applications in dynamic environments. Kacher [5] also falls short in addressing how new technologies, such as autonomous vehicles and big data analytics, could improve the efficiency and adaptability of these transportation models. Bridging this gap between theory and practice, particularly in the context of smart transportation systems, remains an important area for future research.

DOI: 10.54254/2753-8818/2025.23342

Abounacer [6] focuses on transportation problem of disaster response by proposing a multi-objective ILP model. The model aims to determine the optimal number and location of humanitarian aid distribution centers (HADCs) while also optimizing the transport routes to minimize delays. It set ,balances and minimize three key objectives: transportation time, number of personnel required, and meeting the demand in disaster-stricken areas. The authors propose the epsilon-constraint method to generate an exact Pareto front which is a element of Pareto optimal set, providing decision-makers with a range of trade-offs between objectives. However its computational intensity makes it impractical for real-time use in large disaster scenarios. So, a heuristic algorithm is introduced to reduce computation time for large-scale instances in order to adapt with rapidly changing disaster conditions.

In additional, Archetti [7] introduces some programming models designed to solve practical multi-modal freight transportation problems, where goods are transported by different combinations of modes. The main focus of [7] is on service network design and optimization in multi-modality which lead to challenges on routing selection and competition behaviours in multi-modal transportation. In that paper, several algorithms have been selected to solve the ILP model, like genetic algorithms and tabu search, which are effective for solving large, complex logistics networks. However, a key limitation identified is the lack of real-time data integration, making these models less adaptive to dynamic situations like traffic disruptions or weather changes. Moreover, the paper notes that most models prioritize cost efficiency, with insufficient consideration of environmental sustainability and carbon emissions. The absence of practical case studies further limits the real-world applicability of these models, highlighting a gap between theoretical development and practical implementation in modern logistics networks.

1.3. Our Contribution

In this paper, we integrate the transportation problem with a well-known graph-based problems called minor spinning trees (MST) problems to get a new Graph-based transportation problem . In this new model, we set a topological graph with the nodes not only the set of factory and retail store, but also some warehouses situated between those nodes. These warehouses can temporarily store products during the transportation process, effectively acting as additional sources of products for delivery from the factories to retail stores. which means warehouses could be some new source for products to transport into retail store from factory. When a warehouse is activated as a transportation node, we must ensure that, during each transportation phase, the quantity of products entering the warehouse equals the quantity leaving it. Thus, by the end of the transportation process, the warehouse holds no inventory, and all products originating from the factory must be delivered to the retail store.

To address this problem, we reformulate the original ILP model. In that ILP model, we use a flows balance constraint to ensure warehouses product balance which is established by Bertsimas [9]. And Our focus is on a specific case study from China's traditional vehicle industry, which will serve as the basis for our numerical experiments. To solve the ILP models, we employ the CPLEX solver in Python. While our study focuses on vehicle transportation, it is important to note that the proposed model and algorithm are generalizable to the transportation of other types of products.

The rest of the paper is organized as follows. In Section 2, we propose the ILP model with some notation and constraints explaining. In Section 3, we will show the details of our numerical experiments, including our dataset, topological graph, algorithm and results. Finally, Section 4 provides conclusions.

DOI: 10.54254/2753-8818/2025.23342

2. ILP model of transportation problem

In this section, we present the detail of our ILP model for the Graph-based transportation problem. We assume a given undirected topological graph with nonnegative weights which represent the cost for one unit product for each edges . And for any edges ij, we ensure weight $c_{ij} = c_{ji}$. Our goal is to minmaize the total cost of a transportation from factories to retail stores. We formulate the ILP models as follow.

2.1. Notations

$G(V, E)$: A given topology graph $G(V, E)$ is a pair of sets such that $E \subseteq [V]^2$; the elements of V are the nodes of the Graph G and the elements of E are the edges .

F: is the set of factory. We denote $n(f)$ is the node for factory $f \in F$.

R: is the set of retail. We denote $n(r)$ is the node for factory $r \in R$.

$a_f \in \mathbb{Z}^+$: production capacity of the factory f.

$b_r \in \mathbb{Z}^+$: weekly demand for retail outlet r.

$C(G) = (c_{ij})_{n\times n}$: is the cost matrix of graph G. c_{ij} is the cost for one unit product between nodes i and j. If i and j are not adjacent, that we denote $a_{ij} = +\infty$.

$x_{ij}^f \in \mathbb{N}$: is a variable. we denote $x_{ij}^f = k$, if for factory f's production, the edge $ij \in E$ has been established to transmit k unit products. If the edge $ij \in E$ is not use for factory f's production, we denote $x_{ij}^f = 0$.

2.2. ILP Model with Constraint

As we shown before, the variables x_{ij}^f are the units we transport from factory f in edges ij and c_{ij} illustrate the cost for one unit in edges ij. so (2a) calculates the total cost for all factory $f \in F$ in all edges $ij \in E$.

$$\min \quad F = \sum_{f\in F}\sum_{i,j\in V} c_{ij}x_{ij}^f \tag{2a}$$

$$\text{s. t.} \quad \sum_{i\in V} x_{ij}^f - \sum_{k\in V} x_{jk}^f = \begin{cases} -a_f & j = n(f) \\ 0 & others \ \forall j \in V, \forall f \in F, \forall r \in R \\ b_r & j = n(r) \end{cases} \tag{2b}$$

$$x_{ij}^f + x_{ji}^f \le \max\{x_{ij}^f, x_{ji}^f\} \quad \forall f \in F, \forall i, j \in V \tag{2c}$$

$$\sum_{f\in F} a_f - \sum_{r\in R} b_r = 0 \tag{2d}$$

$$x_{ij}^f \in \mathbb{N} \quad \forall f \in F, \forall i, j \in V \tag{2e}$$

For constraints (2b) , it ensure the flow balance for all the nodes in that graph. At the node for factory f, the quantity of products entering is a_f less than the quantity leaving it, which is the total output of the factory. At the node for retail store r, the quantity of products entering exceeds than the quantity leaving by b_r, which is the demand for retail store. And for all other nodes in that graph, inflow and outflow are equal, ensuring a balance for any warehouse nodes. Constraints (2c) ensures that there is no circle in all nodes, if one of the x_{ij}^f and x_{ji}^f is positive , than the other are equal to 0.

DOI: 10.54254/2753-8818/2025.23342

Constraints (2d) ensures that the output of all the factory $f \in F$ is equal to the demand for all the retail store $r \in R$, hence there is a balance for the whole graph. Constraints (2e) shows that all the variable x_{ij}^{f} are nonnegative integers which ensure the minimum unit of our problem is 1.

2.3. Discuss of our model

As introduced in our model, we can construct a tree from one factory to each retail store through the flow balance constraint. This tree is an acyclic and connected subgraph, ensuring that no cycles exist by Diestel [8]. Our objective is to find a tree which has the minimum cost known as the minimum spanning tree (MST) problem. When there are multiple factories, this problem can be reduced to the single-factory case without loss of generality, as illustrated by Bertsimas [9].

There are numerous algorithms available for solving the Minimum Spanning Tree (MST) problem, each with varying efficiencies depending on the nature of the graph. Some of the most well-known algorithms include Kruskal's algorithm according to Stam[10], which uses a greedy approach to iteratively add the smallest edges while avoiding cycles, and Prim's algorithm by Martel[11], which builds the MST by progressively expanding the tree from an arbitrary starting node. Additionally, more advanced methods in ILP models such as CPLEX solver, branch and cut algorithm can be employed by Marin[12]. These algorithms can be adapted by transfer MST problem to ILP problems to solve.

3. Numerical Experiments

To evaluate the performance and practical applicability of our proposed ILP model, we conducted a series of numerical experiments using real-world data from the traditional vehicle transportation industry. In these experiments, we constructed topological Figure 3. based on specific study with Lv [13], where factories, warehouses, and retail stores formed the nodes, and transportation routes between them represented the edges. By solving these instances using the CPLEX solver in Python, we were able to assess the model's effectiveness in optimizing product flow, ensuring balance at warehouse nodes, and minimizing overall transportation costs. The following provide detailed descriptions of the data, experimental setup, and results, illustrating the model's capability in addressing complex transportation problems.

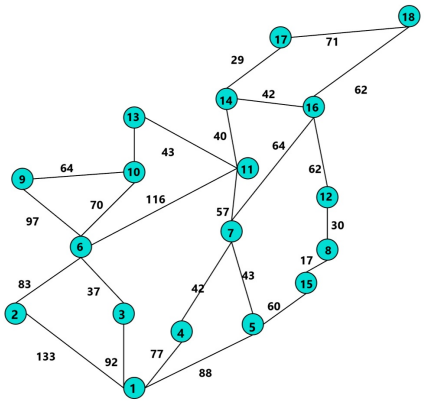


Figure 3: Graph of Major Cities in China(GMCC)

Given that the number of independent variables in our model increases significantly with the number of factories, we selected four groups of experiments, each containing one factory and six

DOI: 10.54254/2753-8818/2025.23342

retail stores. This approach ensures that the total number of independent variables remains below 1,000, making the problem manageable and solvable by most commercial solvers. The detailed data for these four experimental groups, including factory outputs, retail store demands, and locations, are presented in Table.1. In particular, the transportation cost we set in the model is based not only on distance but also on road conditions. It represents the monetary expense required to transport one unit of product, reflecting both the physical distance and the quality of the transportation infrastructure.

Table 1: Factory Information

	GVC	BGV	SVC	AGC
factory number	1	1	1	1
factory output	47300	57000	40600	39400
weekly demand of LN	8000	9900	6200	6200
weekly demand of FJ	8600	10500	6800	6700
weekly demand of BJ	9500	11200	7100	7100
weekly demand of YN	7200	7800	9700	8200
weekly demand of SX	6900	8700	5300	5600
weekly demand of GZ	7100	8900	5500	5600

As outlined in our model, the primary objective is to solve for an optimal spanning tree that ensures minimal transportation costs while satisfying the demand at each retail store. This is accomplished through integer linear programming (ILP), which allows us to handle the complex constraints imposed by the flow of goods between factories, warehouses, and retail stores. By applying the CPLEX solver, a powerful optimization tool capable of handling large-scale ILP problems, we obtain a solution that consists of two key components: the matrix of independent variables and the corresponding optimal value of the transportation cost.

The matrix of independent variables is sparse, meaning that most of its entries are zero. Non-zero entries in the matrix correspond to specific paths between nodes (factories, warehouses, and retail stores) where transportation occurs. Each entry represents the quantity of products transported along a particular path, and it is a non-negative integer if and only if the edge ij is used to transmit goods.

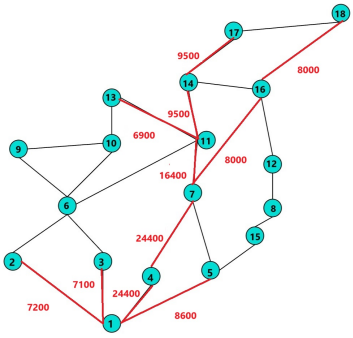
To gain insights into the solution and interpret it visually, we can convert the matrix into a graphical representation. Each non-zero element of the matrix corresponds to an edge in the graph, indicating that the path is part of the optimal transportation network. By visualizing this graph, we can identify the optimal spanning tree that minimizes transportation costs while ensuring the flow of goods meets the demand at each retail store.

For illustrative purposes, we present the optimal spanning tree in the form depicted in Figure.4, where nodes represent the factories, warehouses, and retail stores, and edges represent the transportation paths along which goods are moved. The thickness of the edges can be used to indicate the amount of product transported along each path, providing a clear visual summary of the transportation plan.

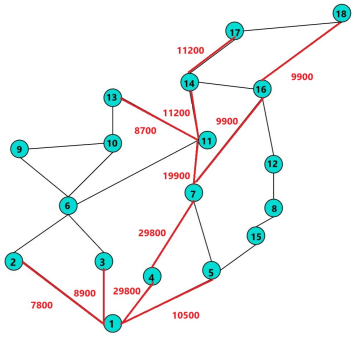
4. Conclusion

In this paper, we introduced a novel approach that combines the graph flow problem with the traditional transportation problem, proposing an innovative ILP model to address the challenge. By

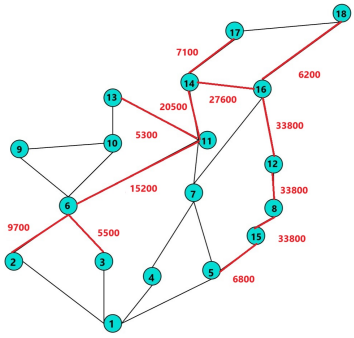
DOI: 10.54254/2753-8818/2025.23342



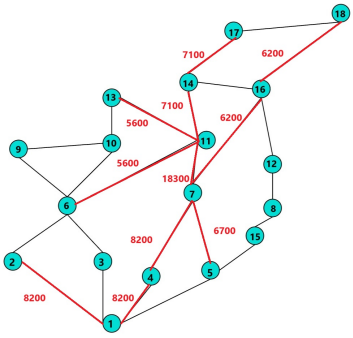
(a) Guangzhou Vehicle Co., Ltd.(GVC)



(b) Beyond Guangzhou Vehicle(BGV)



(c) Shanghai Vehicle Co., Ltd.(SVC)



(d) Anhui Vehicle Co., Ltd.(AGC)

Figure 4: Result of ILP model

incorporating warehouses as intermediate nodes and ensuring flow balance at each node, we provided a more flexible and realistic representation of product transportation. We demonstrated the effectiveness of our model by utilizing the CPLEX solver, showing its applicability to real-world scenarios such as the vehicle transportation industry. Our results highlight the potential of this approach to optimize transportation networks, minimizing costs while maintaining the balance of product flow through intermediate storage facilities.

While our proposed model and solution offer promising results, there are several avenues for future exploration. Introducing additional constraints, such as setting maximum transport limits for each edge, could provide more robust and detailed solutions. Incorporating time series data to account for variable production rates and fluctuating demand over time would further enhance the model's applicability in dynamic environments. Moreover, extending the model to handle multiple product types and more complex processes within the transportation network could expand its use across various industries. These enhancements will make the model more versatile and adaptable to a broader range of practical transportation problems.

Contribution

Yan Dong : idea, paper, slices, ILP model, algorithm, conclusion, coding , experiments results and image editing.

Yixuan Li (Mandy): Background, motivation and coding.

Jiaming Zuo: related work and coding.

Jin Yi: Data collection .

DOI: 10.54254/2753-8818/2025.23342

References

[1] Li Chun. Analysis and research on the operation and management of automobile transportation[J]Technology and Market,2018(11)

[2] Liu Shuangyuan: Analysis of the current situation and development strategy of China's automobile transportation, Industrial Technology, 2017(06):76-77

[3] Chen Yue, The Current Situation and Development Countermeasures of Transportation Economic Development, 2017 (05): 168-169

[4] Fan Yuyun,Transport issues and assignment issues,2013

[5] Kacher, Yadvendra, and Pitam Singh. 'A Comprehensive Literature Review on transportation problems'. International Journal of Applied and Computational Mathematics 7, no. 5 (October 2021): 206.

[6] Abounacer, Rachida, Monia Rekik, and Jacques Renaud. 'An Exact Solution Approach for Multi-Objective Location–transportation problem for Disaster Response'. Computers & Operations Research 41 (January 2014): 83–93.

[7] Archetti, Claudia, Lorenzo Peirano, and M. Grazia Speranza. 'Optimization in Multimodal Freight transportation problems: A Survey'. European Journal of Operational Research 299, no. 1 (May 2022): 1–20.

[8] Diestel, Reinhard. Graph Theory. Vol. 173. Graduate Texts in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017.

[9] Bertsimas, Dimitris, and John N. Tsitsiklis. Introduction to Linear Optimization. Athena Scientific Series in Optimization and Neural Computation. Belmont, Mass: Athena Scientific, 1997.

[10] Stam, C.J., P. Tewarie, E. Van Dellen, E.C.W. Van Straaten, A. Hillebrand, and P. Van Mieghem. 'The Trees and the Forest: Characterization of Complex Brain Networks with Minimum Spanning Trees'. International Journal of Psychophysiology 92, no. 3 (June 2014): 129–38.

[11] Martel, Chip. 'The Expected Complexity of Prim's Minimum Spanning Tree Algorithm'. Information Processing Letters 81, no. 4 (February 2002): 197–201.

[12] Marín, Alfredo. 'Exact and Heuristic Solutions for the Minimum Number of Branch Vertices Spanning Tree Problem'. European Journal of Operational Research 245, no. 3 (September 2015): 680–89. https://doi.org/10.1016/j.ejor.2015.04.011

[13] Lv, J., B.D. Yang, Y.J. Yang, Z.H. Zhang, F. Chen, and G.J. Liu. 'Spatial Patterns of China's Major Cities and Their Evolution Mechanisms during the Past Decades of Reform and Opening Up'. Procedia Engineering 198 (2017): 915–25. 7

[14] Equipment Industry Division 1, Economic Performance of the Automotive Industry in June 2024