Cluster-Based Optimization Method for Delivery Networks

Cheng-Hui Chen[1](✉), Yen-Shiuan Lin[2], Yung-Kuan Chan[2], and Shyr-Shen Yu[1]

[1] Department of Computer Science and Engineering, National Chung Hsing University, Taichung City 407224, Taiwan
star90154@gmail.com

[2] Department of Management Information Systems, National Chung Hsing University, Taichung City 407224, Taiwan

Abstract. Traditional logistics scheduling, which heavily relies on experienced personnel, can be time-consuming and prone to oversights, issues that are further amplified with the integration of new distributors. Addressing these challenges, this study proposes a unique cluster-based optimization method for delivery networks (COMDN). COMDN leverages extensive RFID signal data, incorporating delivery locations, spatial zones, and delivery priorities, among others. The process begins by collecting delivery locations and computing pairwise distances between distributors, followed by the clustering of suppliers based on these distances. The final stage involves constructing an optimal delivery route, assisting in distribution to diverse, dispersed, and complex locations on the map, thereby ensuring a balanced delivery to each location and establishing shorter delivery paths. This results in a significant reduction in order processing times. Using data from a prominent tobacco and alcohol distributor in central Taiwan, the study implements shipment scheduling and route optimization. Experimental results reveal that COMDN, when compared to previous manual methods, shows a significant 2.98% improvement over existing procedures, demonstrating its efficiency and applicability in a wide range of multi-objective delivery and logistics scenarios.

Keywords: Route Optimization · Delivery Networks · Vehicle Allocation

1 Introduction

Traditional logistics, focused primarily on the spatiotemporal displacement of goods, aids in bridging the gap between the production and consumption sites of products. This logistics modality encompasses storage, transportation, and auxiliary services associated with goods handling [1]. From a managerial perspective, cultivating proficient employees in warehousing and transportation demands extensive training, often spanning years. However, frequent personnel transitions impede the effective transfer and accumulation of experience, necessitating constant training for new recruits. Furthermore, internal coordination within enterprises presents considerable challenges, occasionally resulting in communicational discrepancies or conflicts.

D.-J. Deng and J.-C. Chen (Eds.): SGIoT 2023, LNICST 557, pp. 101–112, 2024.
https://doi.org/10.1007/978-3-031-55976-1_10

The transformation from traditional to modern logistics has been accelerated by economic growth and technological advancements. Modern logistics, which are underpinned by state-of-the-art information technology, consolidate transportation, handling, shipping, warehousing, distribution, recovery, and logistics information processing. As a departure from traditional logistics, which is anchored in manufacturing processes, modern logistics foregrounds customer service, underscoring the customer-centric orientation of logistics operations. Two core tenets underlie modern logistics: emphasis on customer service and prioritization of transportation and storage operations. In this context, enterprises are now tasked with addressing the critical challenge of swiftly responding to customer needs and abbreviating order processing lead times [2].

This study presents a cluster-based optimization method for delivery networks (COMDN). This methodology unfolds over three stages. The first stage involves the collection of delivery locations and the computation of pairwise distances between distributors. The subsequent stage focuses on clustering different suppliers based on these distances. In the final stage, an optimal delivery route is constructed. This method aids in distributing to numerous, dispersed, and complex locations on the map, enabling balanced delivery to each location and creating shorter delivery paths. This ultimately reduces order processing times. COMDN can be applied in a wide array of multi-objective delivery and logistics scenarios.

In the past few years, there has been a lot of research on time and capacity constrained vehicle scheduling and routing problems. The Vehicle Scheduling Problem saw the adoption of a Hybrid Genetic Algorithm (HGAV) [1]. This algorithm combines the greedy interchange local optimization algorithm, which requires that all nodes be assigned to vehicles for the minimum vehicle scheduling cost of moving goods from warehouse to arrival. The total travel time, the total delay time and the minimization of the number of trucks are considered, and a feasible solution is found after implementation.

Traditional delivery trucks, possessing a substantial load capacity, do not typically return to the warehouse before completing all customer visits [4]. However, factors such as terrain and traffic conditions may limit the number of customers that can be served. As a result, some e-commerce and logistics companies are now implementing a dual-delivery system using both drones and trucks for efficient distribution. The daily distribution of cartons from collection stations, a unique vehicle routing problem, was addressed by [5]. Because cartons are a special product with a short production cycle, they hope to provide a small number of cartons every day, for the carton factory, they need to get the maximum profit while satisfying customer demand. Due to the high delivery costs for customers who order a small number of cartons per day, carton factories assign delivery tasks to third-party logistics companies to reduce their operating costs. This problem can be solved by using Particle Swarm Optimization (PSO) to deliver cartons from multiple carton factories to a collection station, and then arranging vehicles from the collection station to deliver cartons to the customer. This problem can save about 28% of the total delivered cost, and compared with the actual example, it can significantly reduce the number of vehicles required.

In real life, certain types of transportation have strict time limits. When the delivery man sets out, not only the capacity of the vehicle should be considered, but also the demand of time. These two papers both belong to the Vehicle Routing Problem with Time Window (VRPTW), which is more complex than the traditional VRP and one of the most important branches of VRP. Research explored a scenario in which every customer receives a visit once to secure goods within a specified time limit [6]. The goal is to minimize the total cost, an improved artificial bee colony algorithm (IABC) is used to solve the problem. To tackle the issue of food delivery in Dalian, China [7]. In order to maintain the freshness of food, there are strict requirements on the delivery time. IABC is used as the solution to develop an integrated linear model, which is mainly to provide the lowest cost path for all customers within the time, and it needs to meet the constraints of service time and vehicles [8]. The results show that the IABC algorithm can effectively solve the vehicle routing problem with time window.

In the above-mentioned literature on all employee scheduling problems, whether VSP, VRP or VRPTW problems, they are all considered as non-deterministic polynomial time hard (NP-hard) problem [9]. When confronting complex or large-scale problems, the time taken to find a solution can increase exponentially due to constraints or variable length, making it challenging to find the optimal solution within a reasonable timeframe. In practical applications, swift results are often desired over the absolute optimal solution. Thus, the aim is to identify a superior feasible solution within an effective time window. Various methods such as genetic algorithms, artificial bee colony algorithms, or particle swarm optimization can be employed to tackle such problems.

The paper is structured as follows: Sect. 2 delves into the collection and processing of delivery data, and the methodology for optimizing delivery routes. Section 3 examines the application of the experimental method to real-world scenarios. Finally, Sect. 4 rounds off the discussion with the conclusion, highlighting the contributions of this delivery scheduling study and illuminating potential paths for future research.

2 Materials and Methods

This study presents a cluster-based optimization method for delivery networks (COMDN). COMDN unfolds over three stages. The first stage involves the collection of delivery locations and the computation of pairwise distances between distributors. The subsequent stage focuses on clustering different suppliers based on these distances. In the final stage, an optimal delivery route is constructed.

2.1 Problem Description

In this study, the logistics and distribution scheduling of the case company is discussed. In order to reduce the cost of transportation, the appropriate objective function is set according to different problems and needs, and the appropriate scheduling is finally planned, so that the transportation can be smoother and the overall time can be shortened. In traditional transportation and delivery, the delivery arrangement is made manually according to the customer's order and the delivery location, but it will cost too much manpower and time, and the artificial judgment undistributed is inevitable. Therefore, this study proposes a logistics delivery system to assist operational staff in managing dispatch-related tasks. After identifying the dealerships to which deliveries must be made each day, it is necessary to schedule vehicle dispatches. Crucially, decisions must be made regarding which dealerships each van should service each day. To optimize convenience and reduce fuel costs, nearby dealerships are grouped together and assigned to the same van. As such, this necessitates the design of a novel vehicle allocation and delivery routing method.

2.2 Data Acquisition and Content Description

In the proposed method, a wide range of delivery data is collected through RFID signals. The information acquired includes:

1. Dealer Location: The GPS coordinates (longitude and latitude) of each dealer are gathered.
2. Spatial Area Location: GPS positioning is employed to dynamically cluster different dealer areas.
3. Dealer ID: This unique identifier is used for each dealer.
4. Dealer Ranking: The dealers are categorized into four ranks based on their operational scale and sales performance. The ranks are: Level 1, Level 2, Level 3, and Level 4. The frequency of product delivery to these ranks is as follows: 1 delivery for Level 1, 3 deliveries for Level 2, 2 deliveries for Level 3, and 1 delivery for Level 4.

The final computations are transmitted to the individual delivery driver's mobile device via a WiFi network, facilitating the completion of the delivery tasks. This system allows for efficient scheduling and routing, ensuring that deliveries are made according to the priority levels of the dealers.

2.3 Vehicle Allocation

This section outlines the process of determining the distribution of dealerships to each delivery van and establishing the delivery schedules for each van. Typically, in the allocation of dealerships, those that are geographically close are assigned to the same delivery van to facilitate transportation efficiency. Conversely, dealerships that are situated at a greater distance from one another are handled by different vans to maximize the breadth of delivery coverage.

Shortest Distance

In this study, the shortest distance between the two dealers was automatically calculated by Google Map. There are many applications that use Google Maps for navigation or geographic information systems. With the Google Maps Application Programming Interface (API), which can link applications, data, and hardware. An integrated system that feeds back relative information and interacts with the user's purpose, providing many functions, including the Google Place API, which is used to build location-based service applications [10].

Suppose S_a is the dataset of dealers to be visited on the same day, where each element contains the dealer to be visited, represented by $S_a[i].Dealer$. The graph dist is described by an $n_w \times n_w$ matrix, where dist$[i, j]$ represents the shortest distance between the ith dealer and the jth dealer. The following pseudo code is the graphical dist created by using Google map to calculate the distance between each dealer, as shown in Algorithm Distance section.

Algorithm Distance

```
googlemaps.Client(key=api_key)  // google map api
dist = [][]  // create empty array
for i to range(length_data):
   for j to range(i+1, length_data):
      distance = calculation of distance  // The shortest distance between
   the ith and jth dealers
      dist[i][j] = distance
```

Cluster

In this study, after calculating the distance between two dealers, it is then necessary to assign each truck to deliver the products to those dealers. The purpose of clustering is to complete the delivery of all goods without spending too much delivery time and cost. Therefore, the delivery time and cost of delivery can be reduced if dealers with similar distance are delivered by the same truck. In contrast, dealers at the two locations furthest apart are usually assigned to different trucks for shipment during delivery schedules.

The COMDN algorithm is designed to find N_c reference points for N_c trucks and to find. First, two dealers with the farthest distance are found from dist as the benchmark. If N_c is greater than 2, which means there are more than two trucks, then the distance between the dealer and the base dealer must be greater than Th_d, and the sum of the distances between the dealer and the benchmark dealer is the largest, as the new benchmark dealer. The N_c dealers obtained through this algorithm are transported by N_c trucks, which are used as the benchmark for each truck. The pseudocode for COMDN is demonstrated in the Algorithm COMDN section.

Algorithm COMDN

```
benchmark = [ ]  // create an empty array to store the dealer as a benchmark
dist[i,j] = find the two dealers who are furthest away
Th_d = dist[i,j] × 0.9
benchmark = {i, j}
the distance between i and j is set to infinity
count = 2
while count < N_c :
    k = The distance between the dealer and the benchmark is greater than the Th_d, and the distance from the benchmark is the largest
    benchmark = benchmark ∪ {k}
    count = count + 1
Th_d = Th_d × 0.9
```

Next, the N_a dealers are divided into N_c groups, and each group represents the group of dealers who order products to be delivered by a truck. Therefore, the Partition algorithm designed in this study requires that the number of dealers responsible for delivering products of each truck can be evenly distributed. In addition, consideration is also given to the proximity of the dealer's products, as far as possible can be transported by the same truck. According to the base dealer identified by COMDN's algorithm and assigned to the dealers nearest to the benchmark until all the dealers are arranged. The pseudocode is presented in the Algorithm Partition.

Algorithm Partition

for i = 0 to $length_{benchmark}$:
    dealer[i] = benchmark[i]
    total[i] = 1
iteration = 0
while (iteration < $length_{N_a}$- $length_{benchmark}$):
    j = arg($Min_{i=0}^{N_c-1} total$[i])
    From N_a, the closest dealer k to dealer[j]
    total[j] = total[j] + 1
    dealer[j] = dealer[j] ∪ {k}
    iteration = iteration + 1

3 Experiment

3.1 Dataset Description

The data required for the experimental analysis were provided by a well-known tobacco and alcohol agent in central Taiwan for the purpose of the case study. The tobacco and alcohol dealer has more than 2,000 dealers, and there are more than 40 trucks responsible for the daily delivery work. Excel is used as the data set, as shown in Fig. 1. It consists of six fields, including the customer code, which is the code of all the dealers of the agent; grade, which is divided into different levels according to the size and sales performance of the dealers and represents the number of delivery times in a week; adjust the number of visits. The increase or decrease of orders due to activities or policies related to tobacco and alcohol, so users need to adjust the number of deliveries made by the dealer. If the space is blank, the delivery is made according to the number of times indicated by the level; block, the more than 2,000 dealers are divided into several sub-blocks, such as A, B, C, D, etc., each block has about more than 200 dealers; finally, the latitude and longitude are needed to calculate the shortest distance between dealers during grouping and to present them on a map after execution, so latitude and longitude need to be calculated and marked on the map. The system designed by this study was implemented in block to verify the vehicle allocation system. The experimental group was used as the system of this study, while the tobacco and alcohol agents manually arranged as the control group manually, and the experimental group is compared and analyzed with the control group.

In the proposed system, a wide range of delivery data is collected through RFID signals. The information acquired includes:

Dealer ID	Ranking	Number of visits	Spatial Area Location	Lngitude	Latitude
020033	Level1	3	A	120.960597	23.949971
020034	Level1	2	A	120.970005	23.964220
020042	Level1	2	A	120.980202	23.940583
020043	Level2	0	A	120.973451	23.957343
020046	Level1	0	A	120.973359	23.960298
020048	Level1	0	A	120.974125	23.967521
020052	Level1	1	B	121.077550	23.945799
020053	Level1	2	A	120.973029	23.959153
020057	Level1	3	A	120.973262	23.957505
020058	Level1	1	B	121.077646	23.945879
020059	Level2	1	A	120.989970	23.907736
020060	Level1	0	A	121.038958	23.997394
020061	Level1	3	C	120.973454	23.957336
020062	Level1	0	C	120.973962	23.965100
020069	Level1	2	A	120.977361	23.959031
020076	Level1	0	A	120.973621	23.956935
020077	Level1	1	A	120.973262	23.957505
020080	Level3	2	A	120.974022	23.965132
020081	Level1	1	A	121.105071	23.967832
020084	Level1	3	A	120.965437	23.963527

Fig. 1. Part of the experimental data

1. Dealer Location: The GPS coordinates (longitude and latitude) of each dealer are gathered.
2. Spatial Area Location: GPS positioning is employed to dynamically cluster different dealer areas.
3. Dealer ID: This unique identifier is used for each dealer.
4. Dealer Ranking: The dealers are categorized into four ranks based on their operational scale and sales performance. The ranks are: Level 1, Level 2, Level 3, and Level 4. The frequency of product delivery to these ranks is as follows: 1 delivery for Level 1, 3 deliveries for Level 2, 2 deliveries for Level 3, and 1 delivery for Level 4.

3.2 Experiment Setting

In this study, the data of block A were tested and the list of dealers who was scheduled to deliver daily by the agent was reprogrammed for vehicle allocation. First, two trucks were assigned to deliver the goods in this block. Next, the assumption is based on the dealership ratio originally assigned to the two trucks, and the test program's effectiveness succeeds in simplifying the labor cost and time spent on manual scheduling.

After the data is put into the system for execution, the data obtained is verified to be better than the manual method by the following steps:

Step 1: The two groups of dealers are counted separately and take the average of the sum of longitude, which is called the average of longitude (Avg_{long}). It is expressed by formula (1). $longitude_i$ represents the longitude of the ith dealer, and the latitude is also averaged by summation of latitude (Avg_{lat}), which is expressed in the way of formula (2), $latitude_i$ represents the latitude of the ith dealer.

$$Avg_{long} = \frac{\sum_{i=1}^{n} longitude_i}{n} \tag{1}$$

$$Avg_{lat} = \frac{\sum_{i=1}^{n} latitude_i}{n} \tag{2}$$

Step 2: Next, each dealer is averaged with the latitude and longitude into the formula (3). The straight-line distance between $(longitude_i, latitude_i)$ and $\left(Avg_{long}, Avg_{lat}\right)$, where $(longitude_i, latitude_i)$ is the longitude and latitude of the ith dealer.

$$\begin{aligned} d_i &= 6371 \times \arccos[\cos(Avg_{lat}) \times \cos(latitude_i) \\ &\times \cos\left(Avg_{long} - longitude_i\right) + \sin(Avg_{lat}) \\ &\times \sin(latitude_i)] \end{aligned} \tag{3}$$

Step 3: The dealers of the two trucks are respectively calculated by the above formula (3) and then summed up, which is represented by the symbol of sum_{car_j}, , refer to (4).

$$sum_{car_j} = \sum_{i=1}^{n} d_i \tag{4}$$

Step 4: Finally, add up the sum_{car_j} values of each truck to the formula (5), N_c represents the number of trucks. In other words, it's divided into N_c groups. A smaller value indicates that each group is more concentrated. Relatively, the dealership arranged for the same truck is closer. On the contrary, When the value is higher, it means that the dealers are more dispersed, and the distance between dealers is longer.

$$\sum_{j=1}^{N_c} sum_{car_j} \tag{5}$$

3.3 Results and Discussion

The results are shown in Table 1 and Fig. 2. In the table, the blue bottom is the result produced by the system. The vehicle allocation of manual method is substituted into formula (6), and the average distance is 117.88. The average distance of the program results is 114.37, 114.37 is less than 117.88, indicating that the distance between dealers in the weekly implementation of the system is relatively close and concentrated. It can be clearly understood that in the total of 6 groups of data from Monday to Saturday, the results of 4 cases of data are respectively that Tuesday, Wednesday, Friday and Saturday are better than manual allocation. Since this problem is an NP-hard problem, the remaining two pieces of data (Monday and Thursday) will be manually adjusted after execution to achieve better results. It takes a lot of time for the salesmen to allocate dealers based on their years of experience. When there is personnel transfer, the employees need to retrain and accumulate experience before they can clearly understand how to arrange the dealers. Through this study, it only takes 425 s to complete the assignment of the dealers, which can not only greatly reduce the time and labor costs, but also enable the business personnel to deal with the shipping related affairs more quickly.

$$\frac{\sum_{w=1}^{6} \sum_{j=1}^{N_c} sum_{car_j}}{6} \tag{6}$$

Table 1. Comparison of system and manual results in block A.

			The first truck sum_{car_1}	The second truck sum_{car_2}	$\sum_{j=1}^{N_c} sum_{car_j}$
block A	W1	Current Routes	37.77	47.19	**84.96**
		Program Results	45.74	42.04	87.77
	W2	Current Routes	26.38	30.03	56.41
		Program Results	24.07	26.93	**51.01**
	W3	Current Routes	191.76	84.50	276.26
		Program Results	158.74	102.76	**261.50**
	W4	Current Routes	38.46	50.40	**88.86**
		Program Results	46.61	44.90	91.51
	W5	Current Routes	25.76	27.27	53.03
		Program Results	27.68	20.53	**48.21**
	W6	Current Routes	18.38	129.39	147.77
		Program Results	16.72	129.51	**146.23**

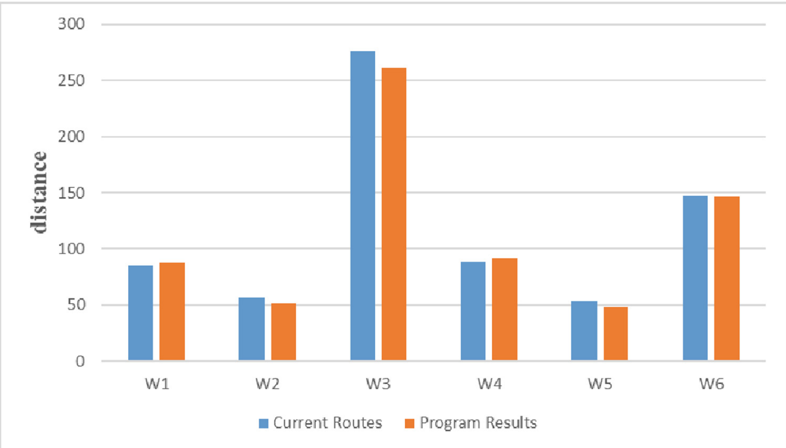

Fig. 2. A bar chart of system and manual comparison in area A

Statistical analysis revealed that the average total mileage from Monday to Saturday, under the existing planning, is 117.88 km. Conversely, the average total for the program-planned route amounted to 114.37 km, marking a reduction of 2.98%. Wednesday (W3) had the highest total mileage, with the most significant difference between the two strategies being 14.76 km.

Knowledge Accumulation for Task Allocation: The current logistical landscape heavily relies on manually acquired experience, which takes a significant amount of time to cultivate, often resulting in organizations having only a single experienced employee. Further compounding this issue is the societal shift towards smaller families, which gradually reduces the availability of apprentices - Factors that all contribute to an impending knowledge gap. However, our method has demonstrated an ability to deliver similar results or even improve efficiency by 2.98%. It facilitates the planning of complex logistical routes and lays the groundwork for the digitization of logistical knowledge.

Rapid Adjustment to Temporary Changes in Order: In the past, when manual task allocation was in place, any sudden **c**hange in distributor demands or inventory shortages required substantial time to readjust the routes, which could even cause significant disruption to the scheduling. Therefore, the implementation of our method allows for real-time adjustments of delivery orders and inventory, thereby enhancing the adaptive capacity of the delivery system.

4 Conclusions

In this study, we propose a novel cluster-based optimization method for delivery networks. Utilizing data provided by a renowned tobacco and alcohol distributor in central Taiwan, we implement shipment scheduling and route optimization. The effectiveness of the proposed algorithm is subsequently validated through comparison with preceding manual methods. Notably, the results indicate an improvement of 2.98% over the current procedures, thereby demonstrating the efficacy of the proposed algorithm.

This method aids in distributing to numerous, dispersed, and complex locations on the map, enabling balanced delivery to each location and creating shorter delivery paths. This ultimately reduces order processing times. COMDN can be applied in a wide array of multi-objective delivery and logistics scenarios.

Funding. This research was funded by Department of Industrial Technology, Ministry of Economic Affairs of the Republic of China grant number 112-EC-17-A-25-1243.

References

1. Li, B., Song, G.: Computational logistics for container terminal logistics hubs based on computational lens and computing principles. IEEE Access. **8**, 194820–194835 (2020). https://doi.org/10.1109/access.2020.3033849
2. Cao, Q., et al.: DDQN path planning for unmanned aerial underwater vehicle (UAUV) in underwater acoustic sensor network. Wirel. Netw. (2023). https://doi.org/10.1007/s11276-023-03300-0
3. Park, Y.-B.: A hybrid genetic algorithm for the vehicle scheduling problem with due times and time deadlines. Int. J. Prod. Econ. **73**(2), 175–188 (2001)
4. Yürek, E.E., Özmutlu, H.C.: A decomposition-based iterative optimization algorithm for traveling salesman problem with drone. Trans. Res. Part C: Emerg. Technol. **91**, 249–262 (2018)
5. Yao, B., et al.: An improved particle swarm optimization for carton heterogeneous vehicle routing problem with a collection depot. Ann. Oper. Res. **242**(2), 303–320 (2016). https://doi.org/10.1007/s10479-015-1792-x

6. Yao, B., et al.: Improved artificial bee colony algorithm for vehicle routing problem with time windows. PLoS One. **12**(9), e0181275 (2017). https://doi.org/10.1371/journal.pone.0181275
7. Yu, S., et al.: An improved artificial bee colony algorithm for vehicle routing problem with time windows: A real case in Dalian. Mech. Eng. **8**(8), 1–9 (2016)
8. Zhang, Z., et al.: Research on bi-level optimized operation strategy of micro grid cluster based on IABC algorithm. IEEE Access. **9**, 15520–15529 (2021). https://doi.org/10.1109/access.2021.3053122
9. Wang, Z., et al.: A joint and dynamic routing approach to connected vehicles via LEO constellation satellite networks. Wirel. Netw. (2021). https://doi.org/10.1007/s11276-021-02712-0
10. Mufid, M.R., et al.: Estimated vehicle fuel calculation based on Google map realtime distance. In: 2019 International Electronics Symposium (IES). IEEE (2019)