Jestr

Journal of Engineering Science and Technology Review 16 (5) (2023) 19 - 24

Research Article

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Retail Distribution using Georeferenced Systems and Genetic Algorithms for Product Delivery. Case study.

José Antonio Taquía Gutiérrez[1,*] and Yvan García Lopez[2]

[1]*Universidad de Lima, Instituto de Investigación Científica, Lima, Peru*
[2]*Universidad de Lima, Lima, Peru*

Received 23 March 2023; Accepted 13 September 2023

Abstract

This paper presents the difference in the performance of the genetic algorithm and the nearest neighbor algorithm applied in a case study of a food distribution company in the city of Lima. A descriptive methodology was used to evaluate the impact on the number of distribution points and the delivery routing obtained. It was possible to show how the evolutionary technique and the nearest neighbor algorithm obtained equivalent results of 1760 on a scaled distance metric and a processing time difference of 78 % between both methods.

Keywords: Analytical distribution, genetic algorithm, urban logistics.

1. Introduction

After the Covid pandemic that reconfigured many business models in the retail consumer sector, companies found it necessary to redesign their processes and transform them digitally. This complexity is observed in the fields of logistics operations, urban logistics, and the field that involves multiple dimensions such as the impact of fuels on the environment, and the increase of operational efficiencies among others. This makes it necessary to use integrated tools within frameworks to establish a delivery design to be evaluated with various scenarios [1].

To achieve customer satisfaction in the last mile product supply chain, routing algorithms play an important role in the logistic distribution frameworks. These algorithms have been studied in the last thirty years, starting with the classical vehicle routing problem to minimize the total distance traveled. Subsequently in the following years, many variants have been studied, with combinatorial problem approaches related to increasing the complexity of the delivery scenario (with time windows, with capacity constraints on the delivery means, with selection of types of delivery means, among others). A systematic literature review on this topic can be found in [2][3][4].

The novelty of the topic addressed in this work lies in how to implement solutions for delivery in complex contexts that have particular characteristics due to the type of geography or consumer characteristics that require a case study [5] where priority is given to customer satisfaction and operational efficiency dimension of the business [6]. It should be considered that there are studies that consider the cost of delivery with a share of 20-40% of the outbound logistics cost. [7].

Currently, an important focus in the study of urban retail distribution is to consider the various collateral aspects related to the delivery of products, public regulations, increased customer orders, and the effect on greenhouse gas emissions pollution due to the increased mass of means of transport on the streets [8], [9]. Likewise, the importance of considering performance indicators involving time and distance is considered relevant in the design of decision support involving retail distribution. [10].

Among the techniques used to model last-mile supply chains, the use of optimization and simulation tools stands out [11], [12], [13]. As reported by [14], there are reductions of 22% in time windows of one hour with the dispatch center model (micro Hubs) for customers with a distance of less than 2 kilometers in urban areas. Addressing the distribution problem by improving the efficiency of delivery capacity use because last-mile outlets are dispersed and do not have storage capacity generating the replenishment effect that makes the solution more complex [15]. This is one of the causes why in these studies the use of analytical application services is implemented due to the complexity of the heuristics used to obtain feasible solutions [16]. The increase in vehicular flow, delivery time restrictions, and the regulations that must be respected about the use of public spaces generate a complexity that involves various types of techniques. Among the most used ones are the use of geospatial tools and discrete simulation [17]. Also in the decision of delivery configurations in the last mile, fuzzy logic can be used to evaluate whether to use autonomous vehicles, drones, bicycles, or traditional delivery vehicles [20],[18], [19]. On the delivery time side, the problem of delivery with time windows is a well-studied topic. Methods such as heuristic optimization are part of this field obtaining interesting results [20].

In a context such as the one described above, two of the digital tools that support and are used in the process of managing product delivery services are georeferenced systems and algorithmic routing techniques. Our research questions are: Is it possible to obtain an optimal level of deliveries in the operational delivery risks perspective? how the distribution sequence of deliveries would be? these questions are addressed in the article.

*E-mail address:jtaquia@ulima.edu.pe
ISSN: 1791-2377
doi:10.25103/jestr.165.03

2. Methodology

Companies dedicated to the delivery of products need to manage the utility in the distribution considering that there are reasons that cause a product not to arrive to be delivered. Among these reasons is that the delivery address cannot be found, this may occur because the location is in a new or inaccessible area. It may also happen that there are delivery times that are not the most appropriate for security reasons, usually called time windows to access some places. The risk of meeting a delivery can be high or very low depending on the location. If we consider the shipping cost dimension, we have used a cost function of carrying the product and not delivering it, becoming the opportunity cost of failing to deliver another load due to delivery capacity constraints. In this logic, two scenarios appear.

Scenario 1: (p): the probability that the number of deliveries will be greater than the actual number of deliveries.
Scenario 2: (1-p): the probability that the number of deliveries is less than the actual demand for possible shipments.
In both scenarios we find different costs, in scenario 1, we will have the operating cost of delivery.
In scenario 2, the opportunity cost or lost margin appears.

Considering this situation as a variant of the marginal sales uncertainty problem [21]. The total utility function will be:

$$\cup = \int_0^R (D * C_u - (R - D) * C_o) * p\, dr + \int_R^\infty R * C_u * (1 - p) dr \quad (1)$$

D: Demand and R: Number of scheduled delivery schedules
Cu: Cost of R<D (being below demand). Margin foregone (Price - Cost).
Co: Cost R>D (to be above actual demand) Distribution operating costs.

We seek to maximize profit:

$$\frac{\partial \cup}{\partial R} = 0, \qquad minimizing$$

$$\frac{\partial \cup}{\partial R} = \int_0^R (D * C_u - (R - D) * C_o) * p\, dR + \int_R^\infty R * C_u * (1 - p)\, dR \quad (2)$$

$$-R * C_o * p + R * C_u * (1 - p) = 0 \quad (3)$$

$$C_o * p + C_u * p = C_u \quad (4)$$

$$p = \frac{C_u}{C_u + C_o} \quad (5)$$

According to the experience of the distribution process managers interviewed, this value should be a ratio that should fluctuate in a value greater than 0.8, according to the historical behavior of delivery costs and the effectiveness of distribution operations. This value was obtained from historical data on the allocation of deliveries and scheduled deliveries.

For the order-taking stage, digital media today have many options [22]. This paper presents a Google application such as Appsheet that can enable access to GPS position with the location of delivery points, as shown in Figure 1.

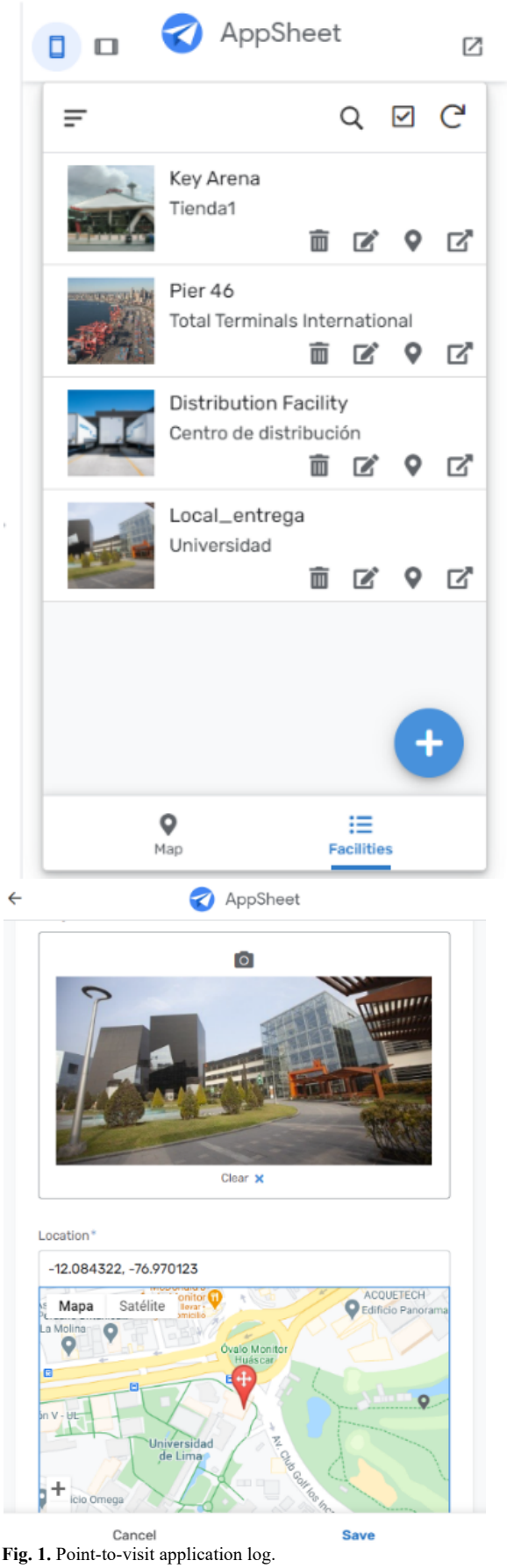


Fig. 1. Point-to-visit application log.

For the simulation of the marginal utility function, we generated the data of a company dedicated to the distribution of orders, considering between four and twelve number of

delivery points. An interface has been designed with the Shiny library of R Studio, to show the generation of the sequence of visits and depending on the scaling of the generated coordinates of the points, obtain a distance matrix that is processed by the "solver" of the TSP library. In combination with the Leaflet library, the result obtained is shown in Figure 2. The design repository can be accessed at (https://bit.ly/3JLVpgL).

Fig. 2. Interface in R - Shiny using Leaflet libraries.

To obtain the sequence of visits, the nearest neighbor method, and the genetic algorithm method will be compared. The distances used to compare these methods are obtained between twelve distribution points extracted in the optimal route analysis of the vendor's problem [23]. The objective is to minimize the distance traveled between the points by complying with visiting each one finding this solution as a Hamiltonian path obtained as a result of a combinatorial optimization problem [24].

3. Case Study

The Peruvian retail market according to the Peruvian Institute of Statistics and Informatics (INEI), still did not show the levels of employability by 2019, one of the segments that have grown the most due to this need to obtain an income in times of pandemic is the distribution by delivery. This modality is used to deliver products in the restaurant, pharmaceutical, and financial services sectors, among others (see Table 1).

Table 1. Which of these pick-up alternatives do you use the most when shopping online?

Most used order reception	2019	2021
Store delivery	58%	80%
Pickup using a mobile application	10%	19%
Customer picks up in-store	30%	31%
Meeting with the seller at a point to be coordinated.	11%	12%

Note: Ipsos. Online shopping in Peru (2019-2021).

Taking into account the INEI employment survey, in the last quarter of 2020 the number of personnel related to product delivery increased compared to the number registered before the pandemic. [25]. This is an indicator of the increase in the dynamics of distribution to the final customer, As mentioned in that report, in other economic sectors there is still a drop of 17% on average. The types of distribution consider the classic delivery driver, with other modalities appearing such as the cab driver by application and mini distribution centers located in geo-segmented conglomerates by consumption indexes. As shown in Figure 3, the online sales channel in Peru tends to grow exponentially in the coming years. [26]. Before the pandemic, the digital channel was just one of many ways to attend the market; today, in this new scenario of multiple channels, online retail sales companies (business to consumer) have made an effort to involve the largest number of their customers where the digital component has greater prominence, not only through the use of tools but also through delivery methods with appropriate security protocols and distribution policies that do not increase the direct costs of the service but allow making it a profitable operation due to the volume of operations and the proper optimization in the allocation of resources.

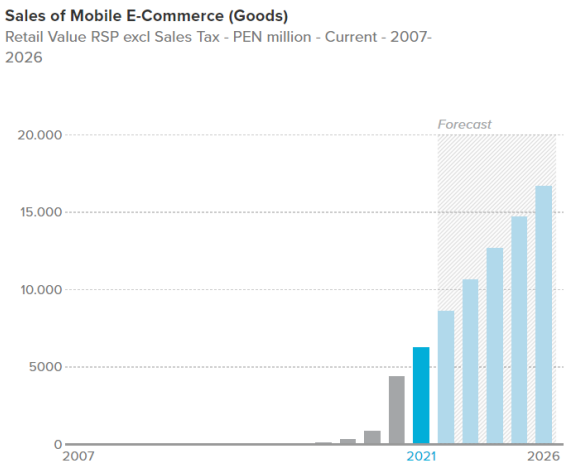
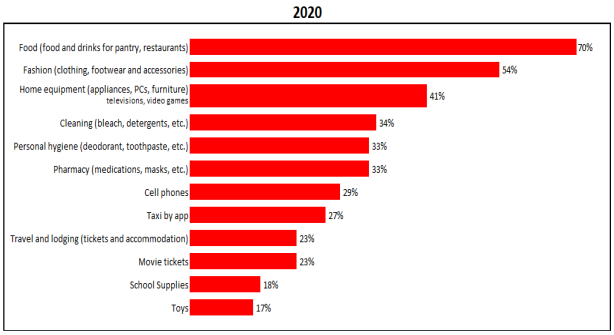


Fig. 3. Internet consumption trend in the Peruvian market.

This trend has been favored by changes in the way of designing the distribution strategy, which responds to omnichannel sales models, For example, the impact of the pandemic on online transactions in retail stores represents about 20% of turnover, and in supermarkets 5%. [27]. Delivery to the customer's door has been a challenge within the distribution strategy and service experience. Some companies have seen the need to set up pick-up points for their products in areas closer to their customers [28]. In direct sales businesses, this modality has been transformed because the information of the delivery point of the customers, in the pre-pandemic model, was not part of the commercial process because the delivery was to the independent consultant who played the role of distributor to its customers. Today this has changed because the companies in the sector take care of the delivery at the final point of sale of the buyer, generating a lot of information and opportunities to optimize the business model. In the modern supermarket channel, the role of intermediary between purchase and delivery at the door has appeared, with companies that manage the purchase and deliver the products to the customers. Customer information and purchase frequency are no longer only in the supermarkets' databases. This mode of collaborative work between manufacturing companies, distribution centers, and distributors is enabled due to the increase of digital applications that use technology capable of growing rapidly due to the volume of operations where it is in the number of transactions where profit margins are generated. On the customer side, the profile of consumer products during the pandemic period in Peru shows in men women a presence of 67% and 79% respectively, Lima, with 77%, where more purchases were made by this means, compared to major cities in the interior of the country (67%). [29]. The trend to use a segmented communication oriented to provide a high level of satisfaction should be part of the strategy of all companies

competing to build loyalty and increase penetration in this type of sales channel. Especially in the food category (Figure 4).



Notes: Ipsos. Online shopping. 2020

Fig. 4. What products or services have you purchased online in the last 12 months?

4. Results

In the sequence of deliveries, it is possible to identify the amount of processing time and the value of the minimum distance obtained for each algorithm considering between 4 and 12 delivery points, which is the number of deliveries obtained from the interviews with experts in retail distribution companies. Table 2 shows the maximum profit for each quantity with default cost parameters (lost margin) of 0.7 and excess cost (storage cost and operating cost) of 0.8 (https://bit.ly/3FTegVZ). For the retail company, this is the maximum profit it can achieve with a delivery unit meeting planned deliveries.

Table 2. Sequence of apportionments

Number of scheduled lines	Profit (monetary units)
4	2.8
5	3.5
6	4.2
7	4.9
8	5.6
9	6.3
10	7
11	7.7
12	8.4

Tables 3 and 4 show the results obtained for the distance matrix used with the algorithms used. These results were obtained with an Intel Core i7-8700 processor.

Table 3. Distances obtained with the genetic algorithm.

Delivery points	Processing time (seconds)	Minimum distance obtained	Sequence
4	1.18	1969	3 1 2 4
5	1.21	1437	5 1 3 4 2
6	1.41	976	1 5 2 6 3 4
7	1.56	1703	5 1 7 6 3 4 2
8	1.75	1648	6 3 4 1 7 8 5 2
9	2.19	1926	4 3 6 2 5 1 7 8 9
10	2.31	1997	7 9 4 3 6 2 10 1 5 8
11	1.88	2116	7 9 11 6 3 4 1 10 2 5 8
12	3.13	1760	7 12 10 1 5 2 6 11 3 4 9 8

Table 4. Distances obtained with nearest neighbor.

Delivery points	Processing time (seconds)	Minimum distance obtained	Sequence
4	0.29	1969	4 3 1 2
5	0.37	1437	5 2 4 3 1
6	0.40	976	6 3 4 1 5 2
7	0.39	1703	7 1 5 2 4 3 6
8	0.44	1398	8 5 2 6 3 4 1 7
9	0.49	1652	9 1 4 3 6 2 5 8 7
10	0.52	2145	10 1 7 8 5 2 6 3 4 9
11	0.65	2116	11 6 7 8 5 2 10 1 9 4 3
12	0.67	1760	12 10 1 5 2 6 11 3 4 9 8 7

The evolutionary analysis is obtained using the GA library [30] which adds to the method the optimization by recombination of the random initial sequence values, using permutations and several 1000 simulations. Distance for this case is obtained using the classical Multidimensional Scaling (MDS) applied for distances [31]. Time reduction on computing processing time shows a 78% less time reduction with the nearest neighbor in comparison with the genetic algorithm.

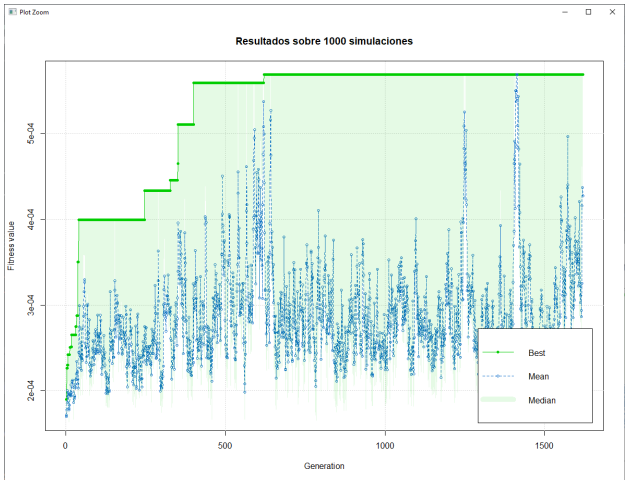


Fig. 5. Routing sequence obtained from 1000 simulations.

Figure 5 shows the randomness in the permutation process of the genetic algorithm using the GA library of R. The graph gives an indication of the number of jumps that the recombination of values in the sequence generates and the optimal minimization frontier of the sum of distances.

5. Discussion

About the techniques used, it can be observed that both models converge to an optimal solution, being the time of the genetic algorithms longer than that of the next neighbor algorithm considered greedy due to the type of path of the feasible solution options. The processing time in the genetic algorithms is longer due to the nature of the algorithm to perform randomness in the recombination of the sequences due to the permutations that the processing takes between the points to be visited.

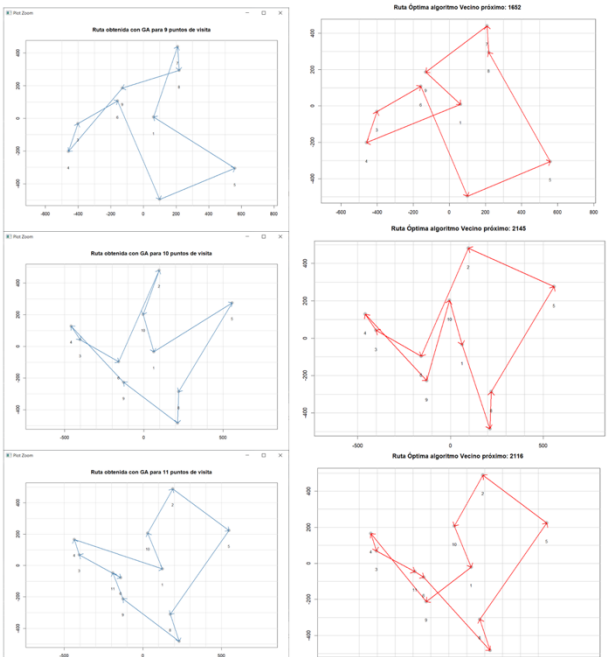


Fig. 6. Differences in results for 9, 10, and 11 distributions.

Regarding the sequence of deliveries, we observe in Tables 3 and 4, the comparison of the results between the next neighbor and genetic algorithm methods, both obtain similar results of distance for 6 of the 9 cases. In three of them, the algorithms had different results (see Figure 6). It can be observed that the time difference between both methods using the same processor, reflects that the nearest neighbor algorithm reaches the minimization objective faster but a limitation of this type of algorithm is the complexity when the number of points to visit increases.

As the increasing e commerce tendency to drop purchased good on time windows delivery time these algorithms could be part on a predictive delivery design to plan a dispatch solution. Moreover, it can give input to more complex mixed integer programming supply chain effort to increase efficiency on multiple echelon supply chain.

Potential challenges on implementing these algorithms are related on time latency to respond while accessing on a concurrent application with thousands of users, which can be addressed with a system architecture engineering base on designing a cloud solution.

6. Conclusions

Delivery companies can take advantage of applications such as those described in this article to improve the sequence of deliveries and develop a strategy of maximum utility in the allocation of delivery units based on a utility function criterion. From an administrative management approach, the processing time using both methods lead to the conclusion that it can be used in a routing solution such as the one described in this case.

Additionally, in the final part of the process, the routing sequence is obtained with the interaction of the objective function of ordering the routes of the distribution in a heuristic way. Genetic algorithms, due to their random design, are computationally intensive, but the experience in the parameterization of evolutionary operators such as the mutation probability, the number of replications, and the threshold function to accept a solution vector, makes this technique a tool that, in conjunction with other algorithms such as a binary or next neighbor model, can give optimal results in an acceptable time.

Future research direction will involve adding monitoring dispatch tracking using frontier embedded systems and gain more insights on urban logistics to better planning and organizing each of the routes, stops, container openings, and some other tasks from real-time generated information.

Acknowledgments

This paper has received support from the Universidad de Lima - Scientific Research Institute (www.ulima.edu.pe/en/scientific-research).

Data availability statement

Datasets and R code can be accessed at (https://bit.ly/42Jynj8)

References

[1] J. G. Urzúa-Morales, J. P. Sepulveda-Rojas, M. Alfaro, G. Fuertes, R. Ternero, and M. Vargas, "Logistic Modeling of the Last Mile: Case Study Santiago, Chile," *Sustain*, vol. 12, no.2 p. 648, Jan. 2020, doi: 10.3390/su12020648.

[2] Ç. Koç, T. Bektaş, O. Jabali, and G. Laporte, "Thirty years of heterogeneous vehicle routing," *Eur. J. Oper. Res.*, vol. 249, no. 1, pp. 1–21, Feb. 2016, doi: 10.1016/j.ejor.2015.07.020.

[3] R. Escamilla, J. C. Fransoo, and C. S. Tang, "Improving Agility, Adaptability, Alignment, Accessibility, and Affordability in Nanostore Supply Chains," *Production & Oper Manag*, vol. 30, no. 3, pp. 676–688, Mar. 2021, doi: 10.1111/poms.13309.

[4] A. Anderluh, P. C. Nolz, V. C. Hemmelmayr, and T. G. Crainic, "Multi-objective optimization of a two-echelon vehicle routing problem with vehicle synchronization and 'grey zone' customers arising in urban logistics," *Eur. J. Oper. Res.*, vol. 289, no. 3, pp. 940–958, Mar. 2021, doi: 10.1016/j.ejor.2019.07.049.

[5] Y. Boulaksil, J. C. Fransoo, E. E. Blanco, and S. Koubida, "Understanding the fragmented demand for transportation – Small traditional retailers in emerging markets," *Transp. Res. Part A Policy Pract.*, vol. 130, pp. 65–81, Dec. 2019, doi: 10.1016/J.TRA.2019.09.003.

[6] E. Gutierrez-Franco, C. Mejia-Argueta, and L. Rabelo, "Data-driven methodology to support long-lasting logistics and decision making for urban last-mile operations," *Sustain.*, vol. 13, no. 11, Jun. 2021, doi: 10.3390/su13116230.

[7] A. Muñoz-Villamizar, E. L. Solano-Charris, L. Reyes-Rubiano, and J. Faulin, "Measuring Disruptions in Last-Mile Delivery Operations," *Logistics*, vol. 5, no. 1, p. 17, Mar, 2021, doi: 10.3390/logistics5010017.

[8] L. Ranieri, S. Digiesi, B. Silvestri, and M. Roccotelli, "A review of last mile logistics innovations in an externalities cost reduction vision," *Sustain.*, vol. 10, no. 3, pp. 1–18, Mar. 2018, doi: 10.3390/su10030782.

[9] E. Croci, M. Donelli, and F. Colelli, "An LCA comparison of last-mile distribution logistics scenarios in Milan and Turin municipalities," *Case Stud. Transp. Policy*, vol. 9, no. 1, pp. 181–190, Mar. 2021, doi: 10.1016/j.cstp.2020.12.001.

[10] H. Chu, W. Zhang, P. Bai, and Y. Chen, "Data-driven optimization for last-mile delivery," *Complex Intell. Syst.*, vol. 9, no. 3, pp. 2271–2284, Jun. 2023, doi: 10.1007/s40747-021-00293-1.

[11] M. Leyerer, M.-O. Sonneberg, M. Heumann, T. Kammann, and M. H. Breitner, "Individually optimized commercial road transport: A

decision support system for customizable routing problems," *Sustain.*, vol. 11, no. 20, Oct. 2019, doi: 10.3390/su11205544.

[12] D. Chen, S. Pan, Q. Chen, and J. Liu, "Vehicle routing problem of contactless joint distribution service during COVID-19 pandemic," *Transp. Res. Interdiscip. Perspect.*, vol. 8, Nov. 2020, doi: 10.1016/j.trip.2020.100233.

[13] J. A. Orjuela-Castro, J. P. Orejuela-Cabrera, and W. Adarme-Jaimes, "Multi-objective model for perishable food logistics networks design considering availability and access," *OPSEARCH*, vol. 59, no. 4, pp. 1244–1270, Dec. 2022, doi: 10.1007/s12597-022-00594-0.

[14] E. Aktas, M. Bourlakis, and D. Zissis, "Collaboration in the last mile: evidence from grocery deliveries," *Int. J. Logist. Res. Appl.*, vol. 24, no. 3, pp. 227–241, May 2021, doi: 10.1080/13675567.2020.1740660.

[15] B. Kin, T. Ambra, S. Verlinde, and C. Macharis, "Tackling Fragmented Last Mile Deliveries to Nanostores by Utilizing Spare Transportation Capacity—A Simulation Study," *Sustainability*, vol. 10, no. 3, p. 653, Feb. 2018, doi: 10.3390/su10030653.

[16] V. Naumov and M. Pawluś, "Identifying the optimal packing and routing to improve last-mile delivery using cargo bicycles," *Energies*, vol. 14, no. 14, Jul. 2021, doi: 10.3390/en14144132.

[17] Z. Lyu, D. Pons, Y. Zhang, and Z. Ji, "Freight Operations Modelling for Urban Delivery and Pickup with Flexible Routing: Cluster Transport Modelling Incorporating Discrete-Event Simulation and GIS," *Infrastructures*, vol. 6, no. 12, p. 180, Dec. 2021, doi: 10.3390/infrastructures6120180.

[18] V. Simić, D. Lazarević, and M. Dobrodolac, "Picture fuzzy WASPAS method for selecting last-mile delivery mode: a case study of Belgrade," *Eur. Transp. Res. Rev.*, vol. 13, no. 1, pp 1-43, Dec. 2021, doi: 10.1186/s12544-021-00501-6.

[19] Z. Chen, M. Yang, Y. Guo, Y. Liang, Y. Ding, and L. Wang, "The split delivery vehicle routing problem with three-dimensional loading and time windows constraints," *Sustain.*, vol. 12, no. 17, pp, 6987 - 6998, Aug. 2020, doi: 10.3390/su12176987.

[20] W. C. Yeh and S. Y. Tan, "Simplified swarm optimization for the heterogeneous fleet vehicle routing problem with time-varying continuous speed function," *Electron.*, vol. 10, no. 15, pp. 1775 - 1784, Jul. 2021, doi: 10.3390/electronics10151775.

[21] W. L. Winston, *Operations research: applications and algorithms*, 4. ed., Student ed. with Infotrac Coll. ed., [Nachdr.]. Belmont, Calif.: Brooks/Cole, Cengage Learning, 2009.

[22] R. C.-B. y B. Figueroa- Zamudio, "Plan De Negocio: Servicio Delivery Por Conveniencia '2Go," Universidad de Piura, Lima, Perú, 2019. [Online]. Available: https://pirhua.udep.edu.pe/bitstream/handle/11042/4233/MDE_1906.pdf?sequence=2&isAllowed=y

[23] S. Nasini, "The Travelling Salesman Problem: Introductory notes and computational analysis," Dept. of Statistics and Operations Research: Universitat Politécnica de Catalunya, 2015, pp. 1–29.

[24] M. Hahsler and K. Hornik, "TSP - Infrastructure for the traveling salesperson problem," *J. Stat. Softw.*, vol. 23, no. 2, pp. 1–21, Jul. 2007, doi: 10.18637/jss.v023.i02.

[25] "Encuesta Permanente de Empleo (EPE) 2021 Trimestre Móvil (Dic-Ene-Feb)," Instituto Nacional de Estadística e Informática - INEI, Apr. 2021. [Online]. Available: http://iinei.inei.gob.pe/microdatos/

[26] "eCommerce market in Peru," Euromonitor, 2021. [Online]. Available: https://www.euromonitor.com/retail-e-commerce-in-peru/report

[27] D. Falcon, "Omnicanalidad, de compra en dos mundos," *Revista G de Gestión*, Dec. 2020. [Online]. Available: http://willardmanrique.pe/wp-content/uploads/2020/12/Revista-G-de-Gestion_18DIC2020.pdf

[28] "Yanbal repartos a domicilio," Yanbal. [Online]. Available: https://zonareparto.yanbal.com

[29] O. A. Mercados, "Compras por Internet. Perfil del comprador on line.," Siatista. [Online]. Available: https://es.statista.com/temas/3751/el-comprador-online-en-espana/#topicOverview

[30] L. Scrucca, "GA: A package for genetic algorithms in R," *J. Stat. Softw.*, vol. 53, no. 4, pp. 1–37, 2013, doi: 10.18637/jss.v053.i04.

[31] K. Severn, "Chapter 6 Multidimensional Scaling (MDS)." 2023. [Online]. Available: https://rich-d-wilkinson.github.io/MATH3030/index.html