ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 13, Issue, 07, pp. 63290-63295, July, 2023
https://doi.org/10.37118/ijdr.26940.07.2023

RESEARCH ARTICLE **OPEN ACCESS**

A SYSTEMATIC REVIEW ON THE USE OF MACHINE LEARNING IN LAST MILE DELIVERY VEHICLE ROUTING PROBLEMS

[1]Patricia Silva, [2]Regina Barwaldt, [3]Helida Santos, [4]Giancarlo Lucca and *[5]Claus Haetinger

[1]Center for Computational Sciences, Universidade Federal do Rio Grande - FURG, Rio Grande, Brazil; [2]Graduate Program in Science Education, Center for Computational Sciences, FURG, Rio Grande Brazil; [3]Center for Computational Sciences, FURG, Rio Grande Brazil; [4]Visiting Professor at Universidade Federal do Rio Grande - FURG, Rio Grande, Brazil; [5]Visiting Professor at Universidade Federal do Rio Grande – FURG, Rio Grande Brazil

ARTICLE INFO

Article History:
Received 11th April, 2023
Received in revised form
18th May, 2023
Accepted 23rd June, 2023
Published online 30th July, 2023

KeyWords:
Vehicle Routing Problem, Last Mile Delivery, Machine Learning, Metaheuristics.

Corresponding author: Claus Haetinger

ABSTRACT

Vehicle Routing Problem has been a classic optimization deadlock, which has been present in operational research for the last 50 years. It aims to find solutions to minimize distances to be covered in delivery demands. Over time and considering the development of logistics, some variants have emerged, such as Last Mile Delivery. These comprise the final stretch, from the regional distribution center to the end user. In this article we have shown a systematic review of the literature on the Last Mile Delivery Vehicle Routing Problem, aiming to find Machine Learning methods and techniques used for optimization. The research is focused on the analysis of scientific publications made between 2016 and 2021 in the following databases: ACM Digital Library, IEEE Xplorer, Science Direct, Springer Link, Wiley Online Library and Brazilian Digital Library of Theses and Dissertations. We obtained a sample of 103 papers, which were categorized according to the solutions presented and the methods applied. The results indicate that metaheuristics and hybrid or mixed methods are the most used. The percentage of participation of Machine Learning is not significant compared to traditional methods, being present in only 6% of publications.

Citation: ***Patricia Silva, Regina Barwaldt, Helida Santos, Giancarlo Lucca and Claus Haetinger. 2023.*** "A systematic review on the use of machine learning in last mile delivery vehicle routing problems". *International Journal of Development Research*, 13, (07), 63290-63295.

INTRODUCTION

With the e-commerce growth and the pressure on logistics operations, together with the increased demand for the customers' service level, there is a demand for innovative solutions in the process of delivering goods. Thus, the search for process optimization, aided by Artificial Intelligence (AI), has become increasingly recurrent (PANTAROTTO, 2018). Costs in the logistics sector are high and transport in general represents the highest percentage of these expenses. Companies are starting to think more about logistics processes, implying the need for applications to reduce costs and better supplying the demands. The distribution of goods is influenced by multiple factors, resulting from the demands of carriers, customers, and the external environment. These factors are transformed into constraints of the problem and consequently lead to the creation of different Vehicle Routing Problem (VRP) variants (KONSTANTAKOPOULOS *et al*., 2020). The VRP, widely studied since the 1950s, aims to route a set of homogeneous vehicles with limited capacity to meet the demand of a set of customers. Its goal is to minimize the total distance traveled by vehicles so that each one starts and ends its route at a given depot, each customer must be visited once by a single vehicle, and the total demand of a route must not exceed the capacity of the vehicle (SALAVATI-KHOSHGHALB *et al*., 2019). The routing problem consists of a network with traffic flows from a source node to a destination node. Vehicles in traffic flows continuously enter the network at the origins, choose their routes to the destination and leave the network at the destination (VARGA, 2022). VRP are proven to be NP-hard problems, although there are exact algorithms capable of finding optimal solutions in a reasonable time for relatively small instances. These problems are computationally difficult and require high computational cost (KIM *et al*, 2015). Since its first formulation, other variants of the VRP have emerged, which present some additional restrictions to the problem, and represent conditions present in logistical operations. In school transport, the scenario is no different, operating costs are also high. Well-planned routes calculated with the greatest possible precision are important allies in the process of reducing costs, mitigating the emission of polluting gases and improvement in the flow of vehicles. VRP establishes the minimum cost routes to be traveled by vehicles, which depart from the same point of origin, and aim to meet the demand of customers in geographically dispersed locations. It

has become a classic problem of combinatorial optimization and is one of the most important and studied since its inception (CARI and GOLD, 2008). There are studies in the literature searching for solutions for the VRP. Many variants of this problem remain without effective outcomes since the studies focus on analytical properties of different types and the corresponding solution methods. Some of these solutions mathematically formulate uncertainties related to VRPs, as they are limited to theoretical or small-scale empirical studies. The implementation of these models and proposals, as well as solution methods in real-world applications, are unusual and still face considerable challenges (BAI *et al*., 2021).

One of the current VRP aggravating factors is the Last Mile Delivery (LMD). It consists of the vehicle routing problem in the last leg of a company package delivery to the final recipient (consumer) who must receive the goods at home or a collection point (GEVAERS *et al*, 2009). Last mile logistics is considered one of the most expensive segments of the entire merchandise distribution process, and the least efficient. This cost is mainly due to the high rate of failed deliveries, caused by the absence of the customer at the indicated address. Operational research that seeks to optimize the VRP Last Mile has grown in recent years, with a variety of works with different approaches. There are some competing definitions of the LMD, but a widely accepted understanding is that the term refers to all logistical activities related to online delivery mode, to families of private customers in urban areas. As per this definition, LMD begins when a shipment arrives at a departure point in an urban area, such as a central warehouse, after a long-distance transport, and ends when the shipment successfully arrives at the destination point to the end customer (BOYSEN *et al*. 2020). Aiming to contribute to the expansion of literature on the subject, the systematic review of the literature investigated the methods and solutions through Machine Learning (ML) techniques proposed for the problem of Routing Vehicles in the LMD, contained in scientific works from 2016 to 2021.

Seeking to solve the VRP, in 2018 Cai *et al*. (2018) used the Algorithm of Iterated Local Search (ILS). The proposed algorithm consists of a sequence of local searches interspersed by perturbations. The perturbations are based on the Adaptive Large Neighborhood Search (ALNS) metaheuristic. Four test instances were used, obtaining an efficiency of 75%. The Ant Colony Optimization (ACO) algorithm was employed by Zhang and Lee (2016) to solve last mile VRP integrated with delivery options to home delivery and shared collection point, obtaining savings in relation to the distance to be traveled and trip duration of 20.87% and 33.99%, respectively. Jiang, Zang, *et al*. (2021) proposed a two-phase heuristic based on Variable Neighborhood Search (VNS). The first phase aimed to build a viable solution, and the second was aimed at optimizing the solution found. The algorithm was evaluated in 54 instances divided into two sets, each with 27 instances. The first group with small instances contained 51 to 101 nodes, and the second, medium sized group with 130 to 200 nodes. The proposed algorithm had a 98.15% efficiency, managing to find the best solution for 53 instances. A real dataset-based optimization approach that combines ML techniques with empowered vehicle routing optimization was proposed by Chu, Zhang, *et al*. (2021), aiming to face the challenges of LMD.

MATERIALS AND METHODS

The systematic process was based on the method proposed by Denyer and Tranfield (2009), which divide the Systematic Literature Review (SLR) into five steps: (i) Formulation of the Research Question, (ii) Location of the studies, (iii) Evaluation and Selection of the Studies, (iv) Analysis and Synthesis, and (v) Results' Report and Use. Fig. 1 shoes the phases of SLR, which make up the protocol of this work, discussed in details in the following sections.

RESULTS

In this section, the research results have been shown. They were based on the methodological process steps of the SLR.

Identifying the research question: We have proposed a research question and some related sub-questions, responsible for guiding the study, as set forth: "What methods and solutions have been used to optimize the VRP in LMD and which are their limiting factors?". The sub-questions were formulated as:

- **Question 1:** What methods have been suggested by the scientific literature to solve the VRP?
- **Question 2:** What are the main methods for solving the VRP in LMD?
- **Question 3:** Which are the main limitations of the proposed methods?

Thereafter we have selected, among the electronic scientific research repositories, the most relevant and the ones related to the researched topics.

Locating the studies: The chosen databases consist of Brazilian and international repositories of Journal and Meeting Papers, PhD and Master Theses. They can be seen in Table 1.

Table 1. Repositories used as a basis for research

International Repositories	Access link
ACM Digital Library	https://dl.acm.org/
IEEE Xplorer	https://ieeexplore.ieee.org/Xplore/home.jsp
Science Direct	https://www.sciencedirect.com/
Springer Link	https://link.springer.com/
Wiley Online Library	https://onlinelibrary.wiley.com/
Brazilian Repositories	Access link
BDLTD	https://bdtd.ibict.br/vufind/

We made use of the following keywords: solution, vehicle routing problem, and last mile to perform the automatic search in these repositories. Furthermore, we utilized the synonyms: application, optimization, vehicle routes prob- lem, and vehicle itinerary problem, as well as the logical connectors (*AND* and *OR*) to optimize the seek. The simplified and the resulting string are given as follows:

- ("solution" OR "optimization" OR "application") AND ("vehicle routing problem" OR "vehicle routes problem" OR "vehicle itinerary problem");
- ("vehicle routing problem" OR "vehicle routes problem" OR "vehicle itinerary problem") AND "last mile";
- Resulting String: (((("solution" OR "optimization" OR "application") AND ("vehicle routing problem" OR "vehicle routes problem" OR "vehicle itinerary problem")) AND "last-mile")

Picking out and assessing the studies: The looking-for included works written in Portuguese and English, during 2016 till 2021. Were only included Journal and Meeting papers and PhD and Master Theses. The research string was applied in each of the selected data repositories, taking into account the following metadata: title, abstract, keywords, as well as the previously established criteria. They were obtained an amount of 1387 studies in the initial search, without applying filters and/or criteria, seen in Table 2. It should be noted that in the Brazilian Digital Library of Theses and Dissertations (BDLTD, for short) repository, we have performed searches with simplified strings aiming to expand the results. In order to guarantee the search trustworthiness, we have applied some inclusion/exclusion criteria, shown below.

- Inclusion of papers written in Portuguese and English;
- Exclusion of papers without authorship;

- Exclusion of introductory papers for special questions;
- Exclusion of works in books, book chapters and workshops;
- Deletion of duplicate papers in different data sources; and
- Exclusion of papers not related to the research topic.

Table 2. Initial search result in the data repositories

International Repositories	Total Studies
Science Direct	593
Springer Link	588
Wiley Online Library	65
IEEE Xplorer	39
ACM Digital Library	17
Brazilian Repositories	Total Studies
Brazilian Digital Library of Theses and Dissertations	85

After refining the search using the inclusion/exclusion criteria, we obtained a total of 103 studies. The Science Direct database brought forward the greatest contribution with 36 works. The list of papers which were found in other repositories can be seen in Table 3.

Table 3. Search result in the repositories after applying some criteria

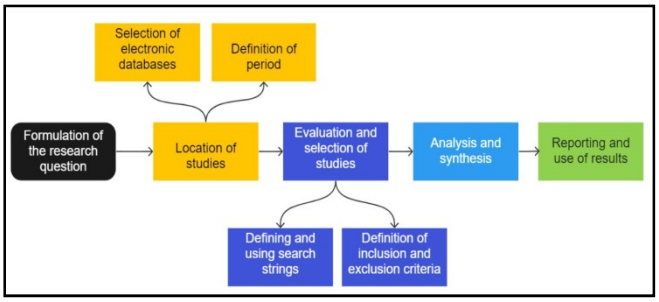
International Repositories	Total Studies
Science Direct	36
Springer Link	24
IEEE Xplorer	14
ACM Digital Library	07
Wiley Online Library	00
Brazilian Repositories	Total Studies
Brazilian Digital Library of Theses and Dissertations	22

Analysis: The results obtained through the search string in the scientific data repositories and the subsequent selection of studies were exported to *Zotero*[1] software tool. The analysis of the works has shown the percentage of publications per repository, the number of publications per year, the types of publication sources, the keywords, the number of authors, the affiliation institutions and the countries of origin. We have divided the productions into categories, according to the answers to the problem questions, allowing the systematization of the solutions addressed in the papers.

REPORTING THE RESULTS

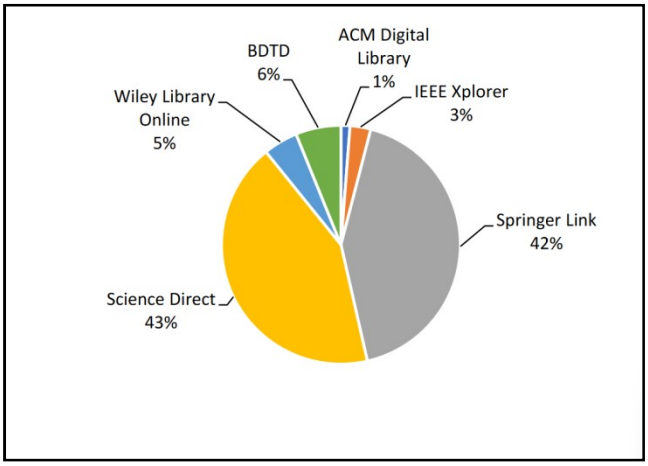
As already mentioned, the initial search in scientific databases totaled 1387 studies. Most of them belong to the Science Direct Database, representing 43% of the total, as can be seen in Fig. 2. After applying the inclusion/exclusion criteria, we had found 103 selected works, of which 35% were from the Science Direct database and 23% from Springer Link. The remaining repositories add up to 42%, as shown in Fig.3. The studies were classified according to the type of scientific publication into: Journal Papers, Meeting Articles, PhD and Master Theses. In line with this classification, we have found 55 Scientific Journal Papers, 26 Meeting Articles, 16 Master's Dissertations and six PhD Theses, as seen in Fig. 4. These works were published by 309 authors, linked to 138 educational institutions, spread over 28 countries. The countries with the greatest contribution in the analyzed period are China and Brazil, with a combined total of 44 publications. Fig. 5 depicts a word cloud with the participant countries. The following describes the number of educational institutions distributed by countries in the publications. As shown in Fig. 6, China hás the highest number of institutions investing in scientific research related to logistics problems, followed by Brazil, France, Germany, and Italy, respectively.

[1] Free research assistant tool for collection, organization, citation and sharing, available at: https://www.zotero.org/.



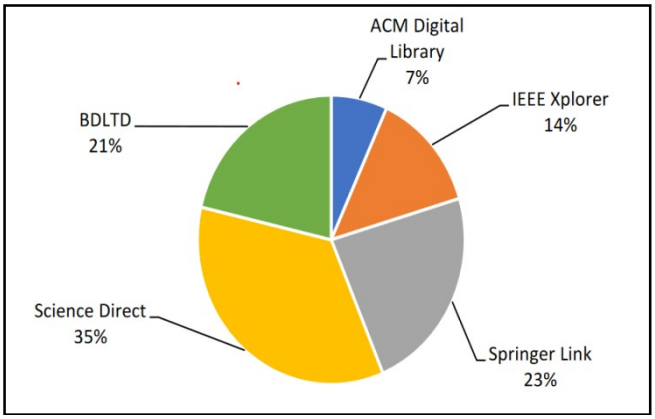
Source: authors

Figure 1. Systematic literature review (SLR) steps used in the work



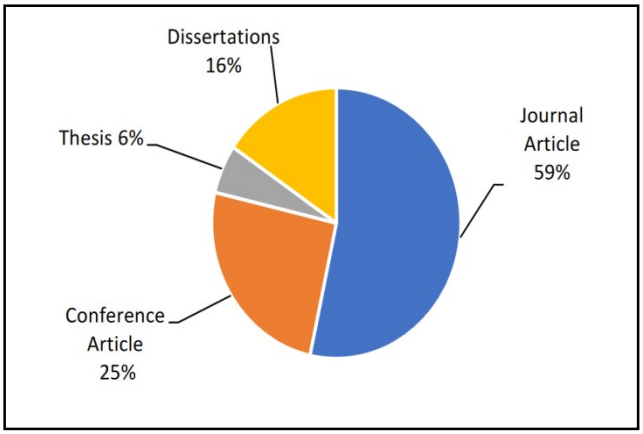
Source: authors

Figure 2. Result of the search string in the repositories



Source: authors

Figure 3. Results after applying the inclusion/exclusion criteria

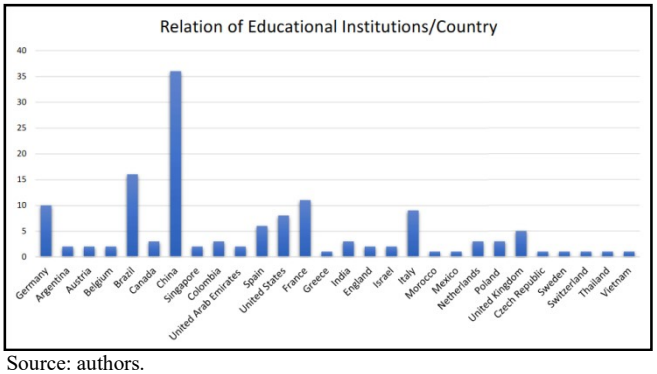


Source: authors.

Figure 4. Percentages of types of publications found

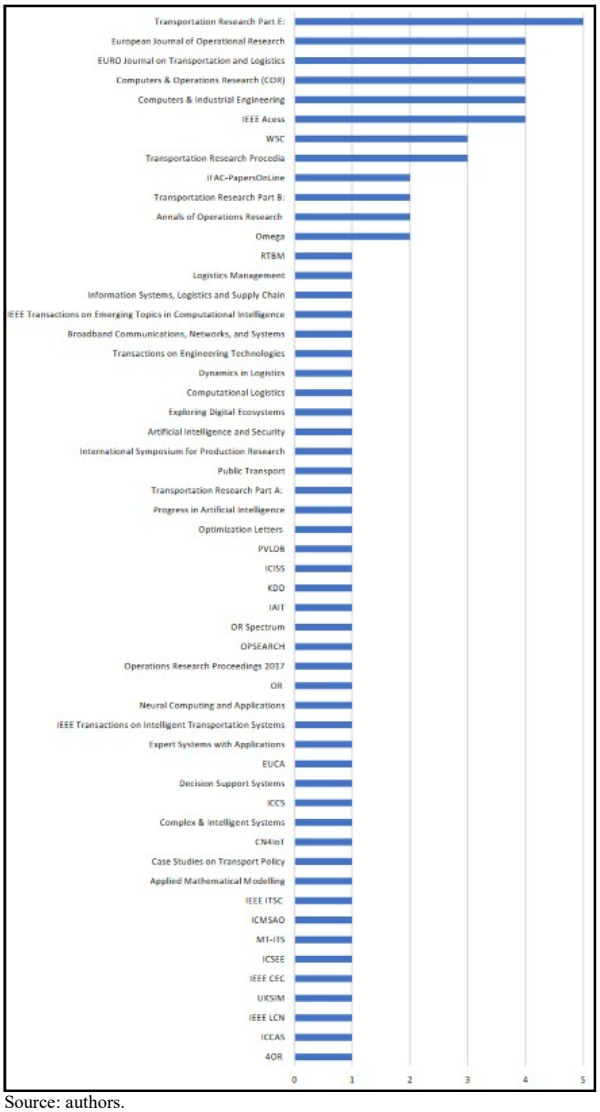
Source: authors.

Figure 5. Cloud of words with the countries related with the selected studies



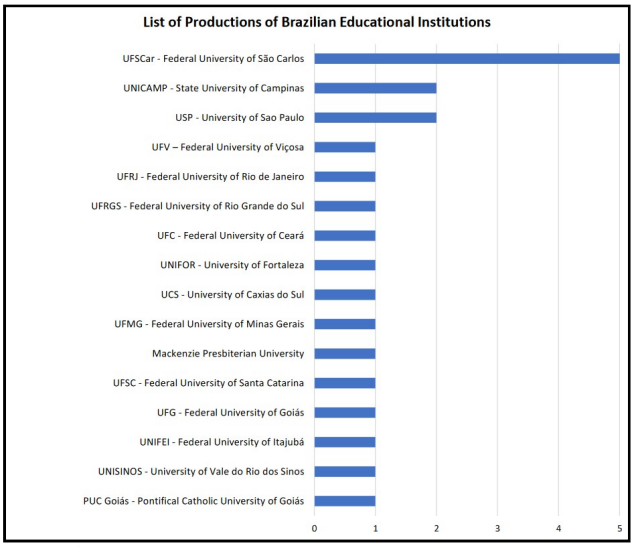
Source: authors.

Figure 6. Number of educational institutions by country



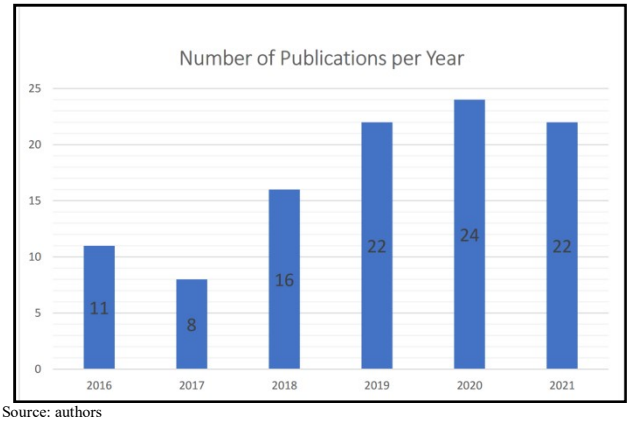
Source: authors.

Figure 7. List of studies by international publication sources



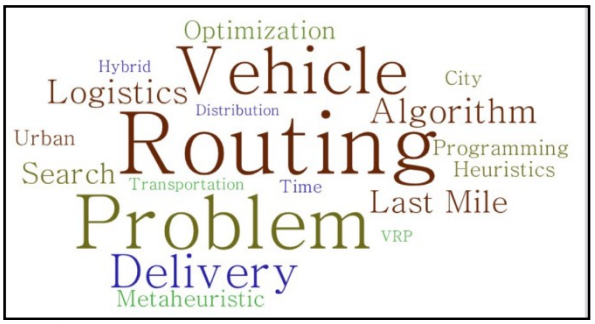
Source: authors.

Figure 8. List of productions of Brazilian educational institutions



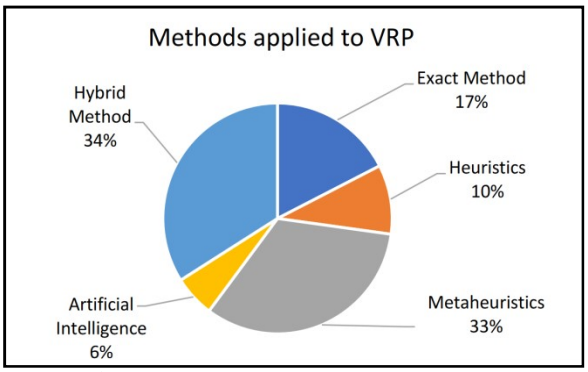
Source: authors

Figure 9. Number of publications per year



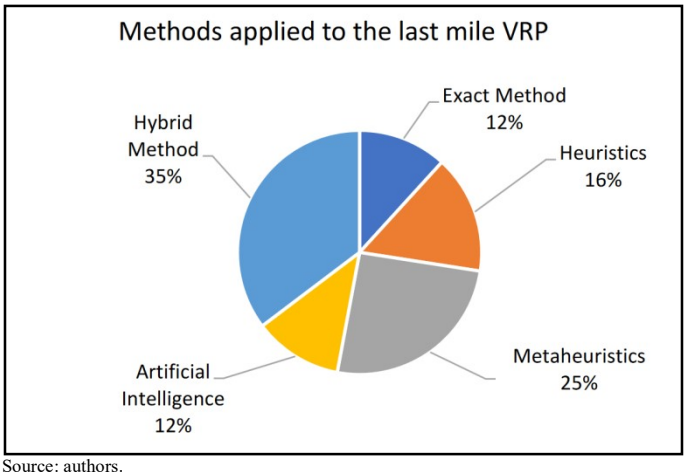
Source: authors.

Figure 10. Cloud of words related with the keywords in the selected studies



Source: authors.

Figure 11. Categorization of VRP solutions



Source: authors.

Figure 12. Categorization of last-mile VRP solutions

Fig. 7 provides the number of works published by international publication source. The journal entitled Transportation Research Part E: Logistics and Transportation Review2 is the one with the highest number of studies. In Fig. 8, we highlight Brazilian educational institutions involved in the studies, and the amount of works from each institution. UFSCar has the largest number of publications, followed by UNICAMP, USP, UFV, and UFRJ. All of them are located in the Southeast region, in the wealthiest and most densely populated Brazilian states (S˜ao Paulo, Rio de Janeiro, and Minas Gerais). The selected studies were published from 2016 to 2021. The last three years have the largest number of published studies. Fig. 9 shows the list of published works per year. For the analysis of the studies, 522 keywords were used, among which the term Routing had 61 occurrences, Problem with 46, Vehicle with 35 occurrences, Delivery with 33, Logistics with 33, Algorithm with 19, and Logistics with 18, highlighted in the word cloud in Fig. 10. All selected studies found gave us answers to the first question: Which methods are proposed in the literature from 2016 to 2021, to solve the VRP? The methods for VRP solutions indicated in those publications were synthesized into five categories: Exact Methods, Heuristics, Meta-Heuristics, Artificial Intelligence, and Hybrid Methods. Exact methods are mathematical methods used to solve optimization problems. They seek the best solution to the problem, not meaning that it is the most satisfactory one. They are mainly used in small-order problems, since the computational processing time becomes costly in higher order problems (GÓES, 2005). As examples of exact methods it can be mentioned the Branch and Bound, Benders Decomposition, and Branch-and-Cut. Heuristics consist of a set of rules and methods that lead to discovery, invention and problem solving. As exact methods, mathematical methods are used, although the most satisfactory solution is sought (WILHELM, 2013). A meta-heuristic is a set of concepts used to define heuristic methods, with potential application to a wide range of problems. In other words, it can be seen as a general framework of algorithms to be applied to different optimization problems, with relatively few modifications, to make them adapted to a specific problem (BENEVIDES, 2011). It was characterized as a hybrid approach those studies in which more than one solution method was applied, that is, with the use of two or more methods, aiming to find an optimized solution. It can be observed that from the 103 selected studies, roughly 67% of the proposed solutions have used Meta-heuristics and Hybrid Methods, 17% applied exact methods, only 6% utilized some artificial intelligence technique and 10% made use of Heuristics, as shown in Fig. 11. Taking into account the second question: What are the main methods for solving the VRP applied in LMD? fifty one studies address this discussion. The most used methods were Meta-heuristics and Hybrid Methods, being present in 60% of the studies, as seen in Fig. 12. Regarding the third question: What are the main limitations of the proposed methods? the majority of researches faced problems when working with large databases. The tests were performed with few instances, obtaining, in these cases, good levels of accuracy and cost reduction. Nevertheless, in studies where larger databases were used, the results indices were less significant.

DISCUSSION

It is understood through the analysis and synthesis of the research results, that the most used methods to solve classic vehicle routing problems, as well as, in the last mile, consist of meta-heuristics and hybridization. It should be noted that the entire sample with AI-based solutions appeared only in last-mile VRP solutions, not being present in the classic problem. It can be concluded that most existing solutions that have achieved better results and optimality have used small instances and datasets for testing, which makes them unfeasible. Since the demand is increasing, one of the biggest challenges today is working with large databases.

Acknowledgments: We are grateful for the encouragement and support provided by Loggi Tecnologia Ltda. company, our partner in the development of this research.

REFERENCES

Bai, R., Chen, X., *et al*. Analytics and machine learning in vehicle routing research. Cornell University, arXiv:2102.10012v1 [cs.LG] 19 Feb 2021.

Benevides, P. F. Aplicação de heurísticas e metaheurísticas para o problema do caixeiro viajante em um problema real de roteirização de veículos, UFPR, Curitiba, PR, 2011.

Boysen, N., Fedtke, S., *et al*. Last-mile delivery concepts: a survey from an operational research perspective. OR Spectrum, 43:1–58, 2020.

Cai, Z., *et al*. Solving the last mile delivery problem using iterated local search approach. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018.

Cari, C.T. and Gold, H. Vehicle Routing Problem, Intech Open, 2008.

Chu, H., Zhang, W., *et al*. Data-driven optimization for last-mile delivery. Complex & Intelligent Systems, 2021.

Denyer, D. and Tranfield, D. Producing a systematic review. Buchanan, D. (ed.). The Sage Handbook of Organizational Research Methods. London: Sage, pages 671–689, 2009.

Gevaers, R., Voorde, E. V., *et al*. Characteristics of innovations in last mile logistics - using best practices, case studies and making the link with green and sustainable logistics. Association for European Transport and contributors, 2009.

Góes, A.R.T. Otimização na distribuição de carga horária de professores - método exato, método heurístico, método misto e interface. UFPR, Curitiba, PR, 2005.

Jiang, L., Zang, X., *et al*. A covering traveling salesman problem with proft in the last mile delivery. Optimization Letters, 2021.

Kim, G., Ong, Y., *et al*. City vehicle routing problem (city vrp): A review. IEEE Transactions on Intelligent Transportation Systems, 16, 2015.

Konstantakopoulos, G.D., Gayialis, S.P. *et al*. Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification. *Oper Res Int J,* 2020.

Pantarotto, C. Otimização do roteamento e programação de veículo e drone aplicada a last mile delivery. USP, Escola Politécnica, Departamento de Engenharia da Produção, São Paulo, 2018.

Salavati-Khoshghalb, M., *et al*. A hybrid recourse policy for the vehicle routing problem with stochastic demands. *EURO Journal on Transportation and Logistics,* 8(1): 269–298, 2019.

2 https://encurtador.com.br/mtA15

Varga, L.Z. Solutions to the routing problem: towards trustworthy autonomous vehicles. Artif Intell Rev (2022). https://doi.org/ 10.1007/s10462- 021-10131-y, 2022.

Wilhelm, V. E. Problema de roteamento de veículos. Departamento de Engenharia de Produção – UFPR, Curitiba, PR, Brasil, 2013.

Zhang, S. Z. and Lee, C. K. M. Flexible vehicle scheduling for urban last mile logistics: The emerging technology of shared reception box. 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2016.

****