Faculdade de Economia da Universidade do Porto

IN-STORE ORDER PICKING ROUTING:

A BIASED RANDOM-KEY GENETIC ALGORITHM APPROACH

Tiago Miguel Ferreira das Neves Salgado

Tese de Mestrado em Análise de Dados e Sistemas de Apoio à Decisão

Orientada por

Dalila B. M. M. Fontes

José Fernando Gonçalves

2015

to every one of whom I was deprived in the process,

myself included.

Biographical Sketch

Tiago Salgado was born on April 7 1983 in Guimarães, Portugal. He graduated in Economics from the University of Porto in 2006. He also holds a Post-Graduation in Tax and an MBA in Finance, for which he received an academic achievement award.

For the past 9 years, he has been working in corporate finance and management consulting. Since 2011, he is also an invited assistant lecturer at the Faculty of Engineering of the University of Porto, teaching corporate management.

Currently, he is the Founder of *Lean Finance*, a new consulting project developed at HM Consultores.

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisors, Dalila Fontes and José Fernando Gonçalves.

Dalila Fontes has been with me throughout this long journey, always bringing good judgement and clarity to many ungrounded impulses. Without her focused and realistic guidance, this thesis would have been an eternal work in progress.

José Fernando Gonçalves was absolutely instrumental for the core of this work, and without his decisive contributions this thesis would not have been possible.

Abstract

Online grocery shopping has been steadily growing and is expected to account for 130 billion USD worldwide by 2025. However, its path so far proved to be more turbulent than initially anticipated, due to severe logistic and marketing challenges. Among these issues is in-store order picking – the fulfillment of customers' orders when performed in a supermarket or grocery store. This is still a very inefficient process, which is responsible for a large share of an online grocer's logistic costs.

We developed a mixed integer linear programming model to assist online grocers in finding efficient routes for order picking, since the pickers' traveling time is the primary source of waste in this process. We propose an approach based on *biased random-key genetic algorithms* (BRKGA) to find solutions for this problem. The genetic algorithm evolves a population of solutions for a number of generations with a bias towards the fittest individuals in each generation.

Experiments with benchmark instances showed that the BRKGA performs well for order picking routing problems, since it finds good quality solutions with little computational effort. The solutions achieved improve upon the ones reported in the literature.

Aided by this BRKGA model, online grocers would be one step closer to a profitable and sustainable business model, by reducing the time and resources required to source customers' orders from the stores.

Resumo

A atividade dos supermercados online tem vindo a crescer sustentadamente e, em 2025, deverá corresponder a 130 mil milhões USD. Não obstante, o seu trajeto até aqui tem-se mostrado mais turbulento do que inicialmente previsto, devido a difíceis desafios logísticos e de marketing. Entre estes encontra-se o *in-store order picking* – a recolha dos produtos das encomendas dos clientes online, quando efetuada no supermercado. Este processo é ainda muito ineficiente, sendo responsável por uma grande fatia dos custos logísticos de um supermercado online.

Desenvolveu-se um modelo de programação linear inteira mista com o objetivo de apoiar os supermercados online na definição de rotas eficientes para o *order picking*, dado que o tempo do trajeto dos operadores é a principal fonte de desperdício neste processo. Neste sentido, propõe-se uma abordagem baseada em algoritmos genéticos viciados com chaves aleatórias (*biased random-key genetic algorithms* – BRKGA) para encontrar soluções para este problema. O algoritmo genético evolui uma população de soluções por um número de gerações, favorecendo os indivíduos mais aptos de cada geração.

Experiências com instâncias de referência mostraram que o BRKGA apresenta um bom desempenho para a geração de rotas de *order picking*, dado que obtém soluções de boa qualidade com reduzido esforço computacional. As soluções encontradas representam uma melhoria face às reportadas na literatura.

Com o auxílio deste modelo BRKGA, os supermercados online estarão um passo mais perto de um modelo de negócio lucrativo e sustentável, ao reduzir o tempo e os recursos necessários para satisfazer as encomendas dos clientes a partir das suas lojas.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

This chapter begins with an overview of the problem addressed in this work: in-store order picking. Next, the research questions are stated, as well as the motivation for developing this work. This chapter ends with a description of the structure of this thesis.

1.1. Problem overview

Online grocery shopping has been continuously growing for the past two decades. However, its development has not been as fast and seamless as initially expected, leading to some skepticism towards this business (De Vuist et al., 2014). Still, the industry's current resurgence seems clear, pulled first and foremost by the customers' increasing needs and also their growing acceptance of the online channel. In France, 33% of customers who never bought online would do so if the service was available in their area, whereas in Spain that figure rises to 49% (Galante et al., 2013).

One of the main obstacles for this business and its viability is logistics (Yrjola and Tanskanen, 2005; De Vuist et al., 2014). Last mile delivery (from the store to the customer) and in-store order picking are the central logistic challenges online grocers need to face. Inefficient processes lead to intolerable operational expenses that can amount to as much as 10 USD to 20 USD per order, which the customers are not willing to bear (Delaney-Klinger et al., 2003; Galante et al., 2013; De Vuist et al., 2014).

In-store order picking refers to the tasks performed by the store's staff in order to fulfill a customer's online order. This is one of the most significant logistic processes for online grocers, and one that is particularly inefficient (Yrjola and Tanskanen, 2005; Boyer and Hult, 2005; Vanelslander et al., 2013). This is why the efficiency of in-store order picking was designated as the focus of our work. To a great extent, the process's inefficiencies are the result of picking routes that do not regard time or distance optimization and thus lead to the waste of pickers' time and effort.

However, the particular case of in-store order picking routing for online grocers is a problem to which we found no relevant reference in current literature. The online grocery business only started to boom in the second half of the 1990's decade, benefiting from the attention given to *dotcom* businesses. Still, its inception was anticlimactic, since the viability of this business was initially overestimated and many companies failed, in some cases dramatically. Hence, the business's widespread adoption has been slow and supported by diffuse business models (López et al., 2014; De Vuist et al., 2014). This may help to explain why there is so little attention given to the operating processes of a business with enormous potential and numerous adversities.

Given its importance for online grocers, finding solutions to make this process more efficient is determinant to allow for the faster and healthier development of this activity. We attempt to design a model to assist in-store order picking, which will hopefully contribute to the establishment of profitable business models for online grocers. Good in-store picking solutions can boost picking speed up to three times, which can be the difference between profitable operations or being out of business (Galante et al., 2013).

In order to achieve good quality solutions, we developed a Genetic Algorithms approach. These are highly problem-independent heuristics that are remarkable at scanning the entire search space, so we believe such an approach to be well suited for this problem. By developing and testing a model based on Biased Random-key Genetic Algorithms (Gonçalves and Resende, 2011), we aim to have a clearer understanding about how well this approach solves the aforementioned problem.

1.2. Research questions

We address the above-mentioned problem by formulating the following research questions:

Question 1: *Can a Genetic Algorithm approach help improve in-store order picking efficiency?*

Question 2: *Are Biased Random-key Genetic Algorithms competent for solving order picking routing problems, when compared to benchmark solutions?*

Question 3: *Will the adoption of an efficient in-store order picking system result in effective financial gains for online grocers?*

Question 1 results from the methodological approach adopted to solve the in-store order picking problem. Although Genetic Algorithms do not appear in the literature as the primary option to solve routing problems, they have interesting qualities: they adjust to complex problems and they are able to properly explore the search space. Biased Random-key Genetic Algorithms seem promising in finding good quality solutions in short times, due to their bias towards the best solutions. We also wish to understand the quality of these heuristics for order picking problems, hence *Question 2*. *Question 3* covers the final goal of this work: to improve the economic viability of online grocers and reduce the barriers to entry this challenging business.

Throughout this work, we aim to find proper answers to these questions.

1.3. Motivation

There are several reasons why online commerce can be an extremely convenient buying channel. For instance:

i. we can access virtually any existing product, including niche products, regardless of where we are and their local availability;
ii. we can compare product prices and features quickly and cheaply without having to travel to multiple stores;
iii. we can access previous customers' scores and reviews, which helps understanding upfront how the product will perform against our expectations;
iv. we can shop at any time, on any day, whenever it suits our schedule;
v. we can shop at home, at work or anywhere we wish to, through mobile devices.

Since food and beverage stand as the largest retail category and one that is very frequently bought (almost twice a week, on average), the potential benefits of online commerce seem huge (Smith, 2014; Ellickson, 2015). However, when we oppose this to the persistent lack of options for online grocery shopping, the gap remains difficult to understand.

Much has already been written about the strategic and marketing challenges of online grocers (e.g., Morganosky and Cude, 2000; Bevan and Murphy, 2001; Delaney-Klinger et al, 2003; Perea y Monsuwé et al, 2004; Müller-Lankenau et al., 2005; Boyer and Hult, 2006; Janson et al., 2007). Nevertheless, little can be found that relates to the logistic processes of this activity, however vital these issues seem to be (Boyer and Hult, 2005; Yrjola and Tanskanen, 2005).

Considering this lack of attention given to the logistic processes of online grocery shopping, it is our goal that this work contributes to help potential and existing online grocers find better solutions for efficient order picking processes, particularly when performed in existing stores, where the layout is not tailor-made for picking and pickers need to share the area with physical customers (Yrjola, 2001).

Also, as a regular user and advocate of online commerce, the author considers that online grocery shopping could significantly benefit from increasing the awareness for these issues, which will hopefully shed new light on these problems. Moreover, it is the author's belief that fruitful research on this subject will encourage potential online grocers to pursue this business model, accelerating its widespread adoption.

1.4. Structure of the Thesis

After this introductory chapter, Chapter 2 provides a general depiction of the online grocery industry and argues on the importance of developing efficient solutions for its logistic obstacles.

Then, we review some core issues regarding the problem of in-store order picking in Chapter 3. We present an overview of the main logistic challenges for online grocers and then dive deeper into the order picking process. Since the order picking problem can be represented as a Vehicle Routing Problem, this well-known class of problems is also reviewed, along with some of its variants.

Chapter 4 overviews Genetic Algorithms, since this is the methodological approach we use to address the order picking problem. Some Genetic Algorithm approaches to similar problems are presented and the chapter concludes with a comprehensive description of Biased Random-key Genetic Algorithms.

Next, Chapter 5 presents a detailed description of the problem we address, order picking routing, and the mathematical formulation developed to solve this problem.

Chapter 6 is dedicated to the explanation of the proposed methodological approach. The operation of the Biased Random-key Genetic Algorithm model is explained in detail, including the description of the solution encoding and the algorithm's evolutionary strategy.

The numerical experiments and the results achieved by our approach, as well as a description of the benchmark instances used for this purpose, are presented and discussed in Chapter 7.

Lastly, in Chapter 8 the conclusions of this thesis are drawn. This includes the main contributions of this work, its limitations, and also some suggestions for future research.

Chapter 2: Online Grocery Industry

In this chapter we discuss the evolution of online grocers and explain both the industry's appeal and its main trials.

2.1. Overview of the online grocery industry

Supermarkets and grocery stores are mature businesses, considering that these have been around since as early as the 1910's (Ellickson, 2015). Still, there have been some relevant evolutions concerning these businesses throughout the decades, always with the purpose of getting products to the consumers more conveniently, at lower prices and with higher efficiency: from grocery chain stores to self-service supermarkets and supercenters; from computerization to the scanning register and data-based marketing.

The decade of 1990 gave birth to one of these major advancements, as the internet pressed the advent of online shopping. In reality, shopping from home was not a novelty, since it actually precedes supermarkets. The railway, national postal services and telegraph allowed for the development of catalog sales, whose first accounts date back to as early as the 1850's, in the United Kingdom and the United States. These mail-order sales consisted of both staple products and luxury items, but their development was distinct in each side of the Atlantic. American firms operated in predominantly rural areas with limited access to shopping facilities; in the UK, customers largely consisted of urban working-class people with easy access to shops and department stores (Coopey et al., 1999). Beginning in the 1970's, radio and television brought a new model of home shopping, based on direct response marketing.

Still, online shopping became a new paradigm in home shopping and the late 1990's saw a surge of internet-based retailers. This trend also hit the grocery business. The demand for more convenient ways to shop for groceries already existed, due to a greater labor-force participation by women, an increase in dual-income households, and a greater number of single-parents and elderly households with time and mobility constraints. However, online grocery shopping only became a viable option when that demand was complemented by an increasingly universal adoption of personal computers, modems, and home subscriptions of internet services (Morganosky and Cude, 2000). Between 1996 and 2000, numerous new online grocers were battling for customers, offering groceries at the same price of existing grocery stores, without the inconvenience of leaving the house, driving to the store and facing crowds of other customers. Within this period, a large number of online grocers were founded in North America and the UK: Peapod, HomeGrocer, Webvan, Streamline, Grocery Gateway, SimonDelivers, Sainsbury's, Asda, Tesco, Ocado, among others. These were usually segmented into two distinct strategies: bricks-and-clicks (based on the infrastructure of an existing offline grocery business) or pure-play (new and exclusively online grocers). Like many web-based businesses originated at the time, only a few online grocers did succeed, with a large number of them failing to find a successful business model. This was more evident among pure-play online retailers, such as Webvan, Streamline and HomeGrocer. Webvan is frequently cited as the archetype of the dotcom bust. The company invested more than 1 billion USD in fully automated distribution centers and a fleet of delivery trucks, before having any real traction in the market. Then, assuming the need to scale, it spent 1.2 billion USD in the acquisition of HomeGrocer, a bigger pure-play competitor who was also struggling to stop its losses. In the meantime, it achieved high market awareness and successfully went public, reaching a peak stock market value of 7.9 billion USD after its IPO. However, its sales never got near the projected values and were always too distant from the breakeven levels required to capitalize on its huge infrastructure. The company went bankrupt in 2001. On the opposite side, UK's Tesco.com steadily became a successful venture and the leading online grocer in the world. It did so by leveraging its existing stores in the early days, and gradually investing in dedicated warehouses as the sales grew in some regions and fulfilling orders from stores was no longer practical. Albertson's (in the US) and Sainsbury's (UK) also achieved a successful business model by sourcing the orders from its stores. More recently, the successful trajectory of companies such as UK's Ocado or US's FreshDirect has been proving that pure-play can be a viable strategy as well (Delaney-Klinger et al., 2003; Yrjola and Tanskanen, 2005; Boyer and Hult, 2005; Janson et al., 2007; Cattani et al., 2007).

Currently, a growing number of large retailers, e-commerce sites and consumer brands have the online grocery business under their radars: Amazon (with AmazonFresh), Coca-Cola and PepsiCo (with exclusive online products), Procter & Gamble, Unilever, Mondelez International, Wal-Mart, among others (Kantar Worldpanel, 2015).

2.2. The trend towards online grocery

> *Food retailers can't afford not to take e-commerce seriously in the long run. The cynics will say, 'Even after 15 years of e-commerce in food retailing, we're talking about at best 3 to 5 percent market share, compared with 50 percent in travel or 35 percent in electronics in mature markets.' To this, I would have two replies: first, (...) given much lower margins in grocery, if you lose 5 percent of customers to a competitor's online proposition, that makes a big difference in both your profit and loss (P&L) and your competitor's P&L. What's more, online grocery typically attracts the most profitable customers: dual-income households, customers who prioritize convenience over price or promotions, big-spending customers – these are the type of customers you'll be making more loyal to the franchise.* (…) *[Retailers] need to begin transforming their organizations now; otherwise, they will have a rude awakening when outsiders like Amazon start entering their market.* (López et al., 2014, p.1-2).

The quote above is from Christian Wanner, cofounder of LeShop.ch, Switzerland's first online grocery shop (López et al., 2014). It illustrates that, although noticeably lagging behind the dynamic of the overall online commerce landscape, online grocery shopping is an accomplished reality. It emerges as an inescapable trend for the grocery industry, and although it is not expected to dominate grocery shopping in the foreseeable future, it is increasingly representing a more relevant portion of this huge market. In the US alone, where online grocery is yet to achieve 1% of the grocery market, it already accounts for almost 7 billion USD in sales. By 2025, the online grocery market is expected to amount to 130 billion USD worldwide (Kantar Worldpanel, 2015).

While online currently represents only 3.9% of the worldwide grocery shopping market, it grew by 28% in 2014 alone. In the same year, overall supermarket sales dropped by 0.1% in Europe (Kantar Worldpanel, 2015). Thus, with traditional grocery business decaying, growth can only come by stealing market share from competitors, offline or online.

Country	**Penetration** *% of households buying grocery products online at least once in 2014*	**Frequency** *Online grocery purchase acts per online shopping household in 2014*
South Korea	58.9%	9.6
Taiwan	39.1%	3.9
China	35.9%	4.0
USA	29.1%	4.8
Spain	24.7%	2.5
UK	24.2%	13.6
France	23.0%	8.4
Portugal	4.9%	2.9

Table 2.1: Penetration and frequency for online grocery shopping.
(source: Kantar Worldpanel 2015)

South Korea is the most successful case of the adoption of online grocery shopping, with 58.9% of the households buying groceries online at least once a year (see Table 2.1) and the online channel market share reaching 13.2% in 2014. The online market share is expected to reach 30% in ten years (see Table 2.2). This results from the fact that South Korea is one of the world's most technologically advanced societies, having young households with lack of spare time and heavy traffic on its urban areas. The United Kingdom, with 6% online market share, is the second leading country in online grocery shopping adoption, and the country where customers most frequently buy online (13.6 times per online shopping household). However, China is on the path to become the second country with largest online market share and the biggest online grocery market in the world. At the same time, it is curious to notice the low adoption of online grocery shopping in some developed countries, such as Germany or the USA, as well as in emerging markets such as Brazil.

Country	Online Share (2014)	Online Share Forecast (2025)
South Korea	13.2%	30%
UK	6.0%	10%
France	4.3%	10%
China	2.6%	15%
Taiwan	3.9%	-
Netherlands	2.1%	-
Germany	1.4%	-
Russia	1.2%	-
Spain	1.2%	-
Denmark	0.9%	-
USA	0.8%	-
Poland	0.7%	-
Belgium	0.6%	-
Portugal	0.6%	-
Turkey	0.5%	-
Brazil	0.2%	-

Table 2.2: Online grocery shopping market share, 2014 and 2025.
(source: Kantar Worldpanel 2015)

The typical profile of the online shopper is a family with young children, urban-suburban and middle/upper class. They will spend on average 66 USD per order in the UK, 26 USD in South Korea, and 76 USD in France, as opposed to 16, 12, and 36, respectively, for the average offline order. In Figure 2.1 we see that, in comparison to its offline counterpart, the online customer will spend at least twice as much per order, reaching 4 times as much in the UK and 3 times as much in Portugal. Moreover, when a consumer moves from store to online shopping, he is expected to increase his spending (by 4% in the UK). Thus, the average online customer is even more valuable than the average offline shopper (Kantar Worldpanel, 2015).

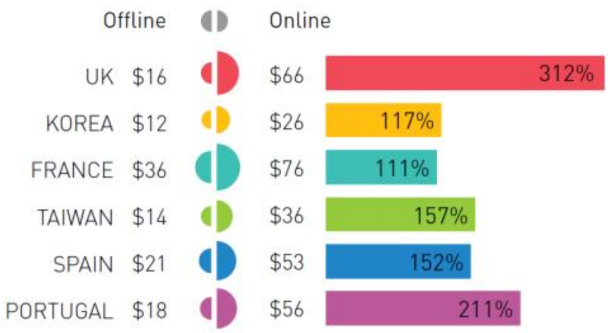

Figure 2.1: Average spending per order, offline vs. online.
(source: Kantar Worldpanel 2015)

According to Goodman (2008), the average American spends 41 minutes in the supermarket store each time he goes shopping. Yrjola (2001) estimates that the time a customer spends shopping represents as much as 20% of the value of the products he buys. A large portion of this time is unnecessary: there is strong evidence that the customers' routes are highly inefficient, since customers frequently enter and leave aisles through the same end and do backtracking, returning to aisles and zones where they had previously been (Larson et al., 2005). Supermarket pickers are able to shop more efficiently than customers due to better knowledge of the store, ability to group orders, and the use of picking systems (Boyer and Hult, 2005). Apart from time savings, several other benefits for consumers of online shopping have been pointed out: saving trips to supermarkets, avoiding crowded stores in rush hour, avoiding queues, not having to carry heavy products home, reduced stress, greater control of the expenses, wider product range, access to specialty items, online reviews, price transparency and comparability, and online promotions (Galante et al., 2013; Smith, 2014; De Vuist et al., 2014; Kantar Worldpanel, 2015).

Thus, the online shopping option configures a better service for some customers, but also for grocers, since it is a way to gather additional data from the shoppers. By analyzing their demographics, clickstreams, online browsing and search histories, online grocers gain insight into the habits and preferences of customers when they put their shopping baskets together, as well as the important factors in the final purchasing decision. This data can be valuable to identify up-selling and cross-selling opportunities, and is becoming more and more a strategic need. Although this information can be profitably leveraged for the offline business, its use for online basket-building features alone, such as contextual product recommendations, can boost average order size by 5% (Kantar Worldpanel, 2015; Breuer, 2013; Galante et al., 2013).

Even if we would disregard that, the truth is that providing online grocery shopping may become inevitable in the future for any supermarket chain. As any other market, online commerce and electronic market systems are likely to become part of the industry's infrastructure and, thus, a strategic necessity (Bakos, 1991). Considering the current importance of e-commerce and the slow but steady increase of online grocery shopping, not having this option may correspond to alienating a relevant portion of the company's market in the future. When moving to the online channel, the company's customers will switch to a competitor, if necessary, and their loyalty to online grocers will be greater than the loyalty to offline grocers. In addition, the online grocery market shares show that first movers are presented with clear advantages and the market is remarkably unkind to latecomers (Kantar Worldpanel, 2015; Galante et al., 2013). Therefore, even if this activity does not become profitable on its own in the short term, it is possible that many established grocery companies will be forced into this business by their customers and competitors.

2.3. The challenges of online grocers

As with online shopping in general, the online grocery business has evidenced a positive trend over the last decades, although with a more modest growth. The main reasons for this are the specific challenges this activity has to face when compared to online commerce of other products such as books, clothing, or technology. These challenges include selling perishable goods, like fresh and frozen items, which need to be stored, picked, transported, and delivered to the customer assuring their freshness and with no quality loss. Because of this, the orders usually need to be delivered in person to the customer, or collected by the customer. In turn, this results in the need for the supplier and the customer to meet, which implies a relatively short time window to be arranged (Yrjola, 2001). Additionally, grocery is a mature and competitive business that operates with low margins and weakening growth. The convenience of online shopping implies additional logistic costs for which the customer is not always willing to pay: it costs the online grocer 10 to 20 USD, whereas the customer is only willing to pay 4 to 8 USD (Delaney-Klinger et al, 2003; Galante et al., 2013; De Vuist et al., 2014; Kantar Worldpanel, 2015). Also, grocery items are bought very frequently. Thus, the customer will have to go through the process of ordering and receiving its groceries regularly, so it needs to be particularly convenient. Lastly, in most cases the online grocery needs to integrate the more stable, bigger and more profitable offline business without impacting its performance (Vanelslander et al, 2013).

As it happens with participants in other electronic markets, potential online grocers face substantial uncertainty about the actual benefits they will achieve, inducing them to adopt a "wait and see" strategy, in the hopes of learning from the experience of early adopters (Bakos, 1991). This could help explain why, in many countries, the offer of online options is so scarce when compared to the number of traditional grocers. In Portugal, for instance, only a few of the major supermarket chains currently offer e-grocery services. None of the small-store supermarket chains have yet, however, made the transition from *bricks-and-mortar* (traditional offline grocer) to a *bricks-and-clicks* (merging offline and online) business model.

This illustrates the fact that large-store chains, usually composed of a lower number of large supermarkets, are best suited for online grocery shopping. Tesco's portfolio of large surface retailers was instrumental for its swift dominance of UK's online grocery market (Bell and Cuthbertson, 2004). In fact, small store chains seem to face a set of additional challenges, due to the stores' smaller dimensions and the larger number of stores. Large stores have a wider product variety, which translates also into a wider online shopping offer. A smaller store must be more selective and deal with a smaller number of stock keeping units (SKUs), reducing both the range of products and the customer's choices on each type of product, in order to make a more efficient use of its area. This will result in reduced customer choice and lower value online orders.

On the other hand, small store chains' specificities may also bring a few advantages. One of the main features of their business model is proximity to the customer: smaller stores are usually placed in more accessible locations, precisely to reduce some of the inconveniences of grocery shopping. As much as this somewhat reduces the need for online shopping (as the inconvenience of driving to the store is less significant), it also makes it easier to offer online shopping without home delivery, in a store pickup (or click-and-collect) model. In this case, the customer would place the order online, select the time at which he would collect it and drive by the store to pick it up. People spend on average 41 minutes in the store shopping each time they visit a supermarket (Goodman, 2008). A pickup model still allows customers to save a large portion of this time and, when compared to home delivery, it has clear financial benefits for the grocers as well, as it can lead to as much as 30% higher profit margins (Galante et al., 2013). This has been put into practice with remarkable success in France, where E.Leclerc, the leading online grocer, has more than 3 500 collection points all over the country, and more than double the online market share of its main competitor. Many European consumers drive by a grocery store on their daily commute and they don't like having to wait for home delivery, which makes them favorably disposed towards the pickup model (De Vuist et al., 2014; Kantar Worldpanel, 2015).

The impact on the order picking performance of issues such as facility layout, storage assignment or the depot location are evident and have been subject to extensive research (see Chapter 3). Also, these different policies interact and influence each other’s efficiency (Petersen and Aase, 2004). However, when order picking is performed in-store, altering the layout or the storage assignment to benefit the order picking process would result in a store that compromises the commercial performance of its main business (Yrjola, 2001). Despite its strong growth, online groceries will only account for a small percentage of the total business, in the foreseeable future, and thus the supermarket stores must still be optimized for the offline business model. Consequently, the store layout, storage assignment and depot location must be taken for granted, which will necessarily limit the overall efficiency of the order picking in supermarkets and grocery stores.

Most of the order picking work consists in traveling throughout the store. According to Bartholdi and Hackman (2014), travel time is the largest component of labor in a typical distribution center. In spite of this, and even more importantly, it is waste: “it costs labor hours but does not add value”, in the words of the authors. Since in a supermarket the facility layout and the products location assignment are not optimized for order-picking, the time wasted travelling will be even more significant: an in-store picker can collect around 70 items per hour, whereas in an efficient dedicated warehouse this number would be between 300 and 400 items (Kantar Worldpanel, 2015). Those restrictions will limit the efficiency of the whole picking process, but will also make even more determinant the decisions concerning order picking and routing policies, as they are the only decisions that could reduce the resources and time needed for fulfilling orders.

The present work aims at helping grocery chains to develop adequate business models for online service, giving their customers more options to shop more conveniently. An efficient order picking model would make it viable to perform this process in stores, thus providing customers with the option of having their orders delivered at home or to collect them at the store on their way home, with no delays.

Chapter 3: Literature Review

In this chapter, we present a review of works on several issues concerning online grocery logistics and order picking. We also review works regarding the Capacitated Vehicle Routing Problem, since it is a closely related problem. Genetic Algorithms, the methodological approach chosen for this work, is reviewed in Chapter 4.

3.1. Online grocery logistics

The online grocery business model consists in allowing the customers to electronically shop for grocery items, and then assuring the necessary processes for the ordered products to be delivered to the customer. This business is composed of a number of distinct activities that the grocer has to perform, some above the customer´s line of visibility (like the customer’s online interface or the value proposition) and others below the line of visibility (Müller-Lankenau et al., 2005). The latter include most notably operational activities, in which logistics play a vital role. The main logistic processes include inbound logistics (from the supplier to the grocer) and outbound logistics (from the grocer to the customer). Since inbound logistic processes for an online grocer are similar to those of the traditional (offline) counterpart, this section will focus on outbound logistics.

A significant part of the outbound logistics for a traditional offline grocer is performed by the customers themselves (Murphy 2003). They are expected to drive to the store, pick their goods, check-out, and transport them home. In some cases, it is even the customer who scans the products at the check-out. An online grocer proposes to internalize some or all of these inefficient and time-consuming activities without adding significantly to the customer's expenses or the company's operating costs (Yrjola, 2001).

According to Hays et al. (2005), differentiation among major online grocers is based on the types of products they offer, their services and the geographical markets in which they operate. In terms of service, the method of order fulfillment and delivery is particularly important. These decisions result from the particular business model of each online grocer:

i. *Pure-play online*: grocers that only operate online and have no physical stores;
ii. *Bricks-and-clicks* (also known as *clicks-and-mortar*, both designations deriving from *bricks-and-mortar*): consists of supporting the online business with a traditional offline grocery business structure;
iii. *Offline-online partnership*: a merger or business agreement between *bricks-and-mortar* (a purely offline business with physical structure) and a pure-play online grocer.

Boyer and Hult (2005) study the integration of marketing and operations for four online grocers, each representing a particular operational strategy, as established by their order fulfillment and delivery decisions. Store-based order fulfillment corresponds to the circumstance in which the order picking is performed in an existing grocery store, as opposed to picking in a dedicated distribution center. With direct delivery, the company will deliver the order directly to the customer's home or office, while with indirect delivery this happens by customer pickup or a third-party logistics provider. Figure 3.1 characterizes the four strategies that result from these decisions: *semi extended*, *fully extended*, *de-coupled*, and *centralized extended*.

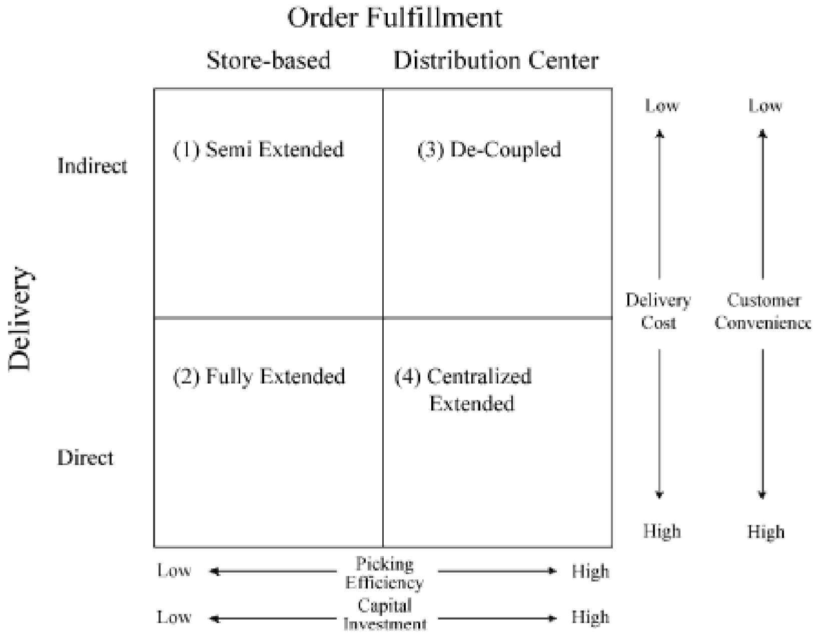

Figure 3.1: Online grocers' operational strategies.
(source: Boyer and Hult, 2005)

Boyer and Hult (2005) state that the online experience, product quality, and service quality have a significant effect on customer behavioral intentions for repeated purchases. They also conclude that grocers with indirect delivery systems require greater efforts to deliver excellent service quality.

According to Vanelslander et al (2013), from the multiple supply chain setups, there are three types that are more commonly employed in Western Europe:

i. Pure play with van delivery;
ii. Pure play with parcel delivery;
iii. Bricks-and-clicks with van delivery.

Pure play with van delivery is the model used by retailers that only use the online model and have a dedicated fleet of vehicles for delivery. Since pure play retailers don't have stores, the picking is performed in dedicated distribution centers. *Pure play with parcel deliveries* is the model used by pure play retailers that rely on a parcel carrier's distribution network for the customer delivery. With *bricks-and-clicks*, the delivery to the customer is usually done with a dedicated fleet of vehicles.

Bricks-and-clicks retailers sell online, as well as in physical stores. Thus, they leverage their existing *bricks-and-mortar* distribution network, avoiding the significant capital and operational expenses required to build and run a dedicated distribution center. Picking operations are performed in-store, with pickers collecting items directly from the shelves and competing with offline shoppers. On the other hand, this also means that the orders are fulfilled closer to the customers, reducing the delivery distance and costs when compared to pure play retailers (Vanelslander et al, 2013).

Murphy (2003) realizes that in order to recognize the potential and the difficulties of online grocers, it is vital to understand customer fulfillment logistics issues (picking, packing, and delivery), which create conflicting space and time management challenges. For *bricks-and-clicks*, the design of the store layout is not well suited for efficient picking and packing by store employees, the pickers have to compete with customers for store space and product availability, and reduced backroom space affects the efficient and seamless processing of orders. Delivery planning also faces severe problems: in low-density areas, delivery will be too expensive; in high-density areas, it may become too complex. The delivery time flexibility given to customers is also a very sensible subject: high flexibility will result in high distribution costs if there is not enough customer density; low flexibility compromises service quality and customer satisfaction. Fulfilling the orders from a centralized automated distribution center is efficient, but results in high operational and capital expenses; picking the orders in-store has significant advantages in terms of product freshness and upfront investment, but carries picking inefficiencies, as stores are designed for and filled with offline customers. The ability to scale is another critical issue. For online grocers with in-store fulfillment, increasing demand will lead to increased contact and competition between store customers and pickers. In contrast, an existing network of stores can be leveraged to fulfill additional customers without major upfront investments.

Boyer and Hult (2005) identify a diverse number of areas concerning online grocery marketing and logistics that require further study. Among these, they cite operational decisions such as algorithms and systems for controlling order picking. Also, Vanelslander et al. (2013) find that the most significant logistic costs for online grocers are the costs of delivering to the customer and the cost of the order picking process when performed in-store, which is consistent with the conclusions of most authors. Thus, the importance of finding efficient solutions for in-store order picking becomes evident.

3.2. Order-picking

Order picking is a logistic warehouse process in which items are collected from their storage locations to fulfill customers' orders. According to Drury (1988), order picking is the most laborious of the warehouse processes, accounting for up to 60% of all labor activities in the warehouse. Order picking's processing time is composed of four major elements (Henn and Schmid, 2013):

i. *Travel time*: the time the picker spends moving between pick locations and the depot;
ii. *Search time*: the time needed for the picker to identify the items;
iii. *Pick time*: the time the picker takes to move the item from its location to the picking cart;
iv. *Setup time*: time spent on administrative tasks between tours.

Travel time is the most substantial of these four elements, since it accounts for at least 50% of the total processing time (Henn and Schmid, 2013).

The study of order picking has addressed a number of different problems, concerning distinct order picking design and control decisions. These include picking methods, but also layout design, storage assignment, picker routing, order batching, zoning, and order accumulation & sorting (De Koster et al., 2007). Gray et al (1992) group these problems into three major decision levels: facility layout & technology selection; item allocation; operating policy. They argue that solving the entire problem formally with a single model is impractical even for small sized cases. Thus, these issues are typically studied independently, although there are also efforts to address more than one simultaneously, as evidenced in the following chapters.

3.2.1. Order-picking routing

Ratliff and Rosenthal (1983) address the problem of order-picking routing in a rectangular warehouse with crossovers only at the end of the aisles. They use a graph representation for the warehouse, in which the vertices correspond to the shipping area, the pick locations, and the endpoints of the aisles. Their model assumes that the picker collects items for just one order at a time and has the capacity to transport all the items in that order. For this configuration, they designed an algorithm based on dynamic programming whose computational effort increases linearly with the number of aisles. Thus, they conclude this algorithm to be fast enough to be applied to any realistic size warehouse.

Petersen (1997) compares the routes that result from the *optimal strategy* of Ratliff and Rosenthal (1983) with five other order picking routing policies (see Figure 3.2):

i. *Traversal/transversal strategy (also known as the S-shape algorithm)*: if an aisle contains items to pick, the picker will always cross the entire aisle from end to end.
ii. *Return strategy*: if an aisle contains items to pick, the picker will enter and leave an aisle from the same end.
iii. *Midpoint strategy*: similar to return strategy, but in which the picker does not cross the aisles' midpoint.
iv. *Largest gap strategy*: similar to midpoint strategy, but the picker will enter an aisle and go as far as the largest separation between two adjacent picking items in that aisle.
v. *Composite strategy*: combines features of transversal and return strategies.

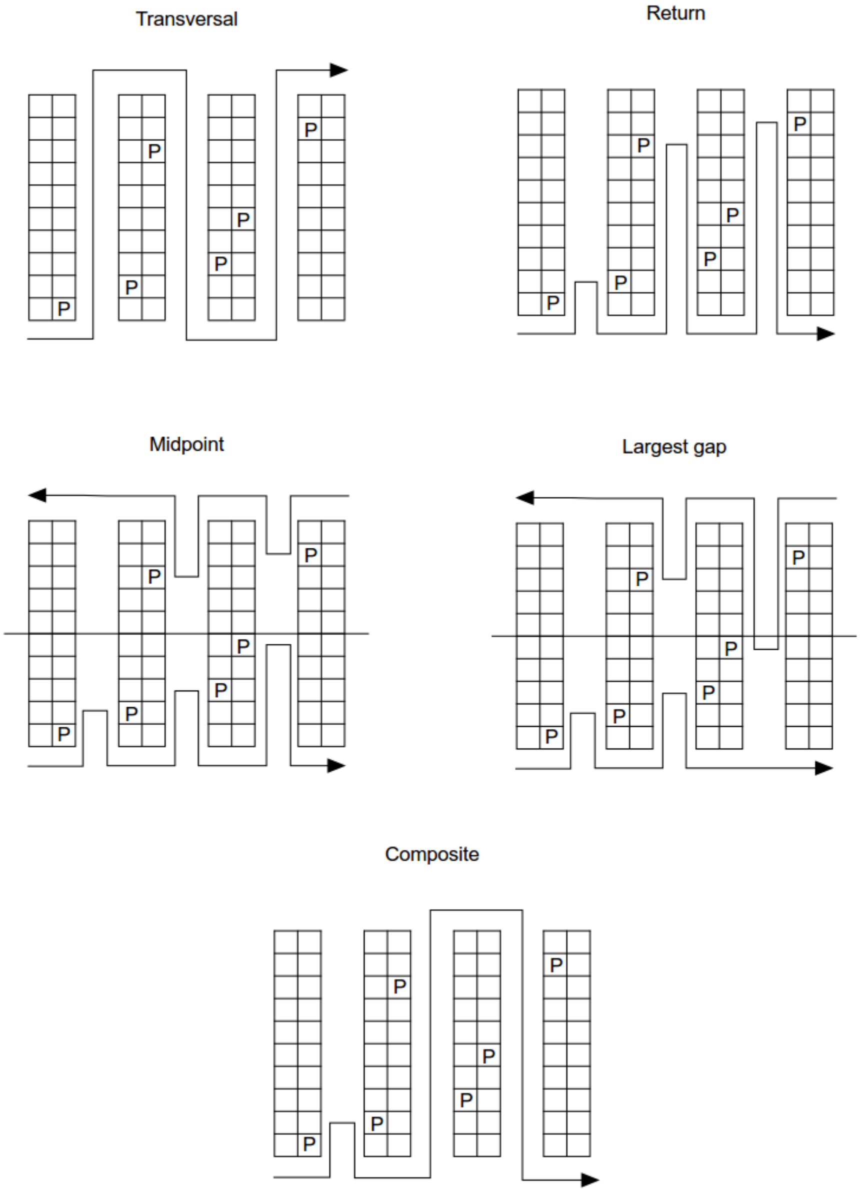

Figure 3.2: Routing strategies.
(source: Petersen, 1997)

Considering random storage allocation, distinct depot locations and distinct picking list sizes, Petersen (1997) conducted experiments with several heuristics. The best performing heuristics (composite and largest gap) produced routes that are, on average, 9% to 10% longer than optimal routes. The reason for this is that the heuristics only consider one type of routes, whereas optimal routes seem to correspond to hibrids of several routing policies. Although these heuristics perform worse than the optimal strategy, Petersen (1997) notes that its routes are easier for the pickers to understand

and are more consistent, which explains its greater real-world adoption. Another important conclusion is that the best heuristic for a given situation will depend on factors such as the warehouse layout or the number of items in a pick list.

De Koster and van der Poort (1998) studied efficient order picking routes for both traditional warehouses with a central depot and warehouses with decentralized depositing. Both optimal and heuristic solutions are studied for order picking routes in these warehouses. Extending the work of Ratliff and Rosenthal (1983), they propose a polynomial algorithm that reduces travel time, although this reduction strongly depends on the layout and operation of the warehouse. Roodbergen and De Koster (2001), on the other hand, extend the previous works to warehouses with multiple blocks (additional cross-aisles) and a central depot.

Theys et al. (2007) argue that exact order picking routing algorithms only exist for warehouses with at most 3 cross aisles. Increasing the number of cross-aisles will result in a growing number of routing possibilities, which, depending on the number of items to pick, may not be efficiently solved through optimal methods. Thus, they develop a meta-heuristic approach to solve the problem of routing and sequencing order pickers in multiple-block warehouses. They reformulate the order-picking routing problem as a classical travelling salesman problem (TSP), in order to adapt a TSP-based heuristic.

Nikolakopoulos (2013) proposes a Threshold Accepting meta-heuristic algorithm for the order picking routing problem, the results of which are compared to optimal picking routes. He uses randomly generated instances with several combinations of warehouse layouts, storage locations, and order sizes, all consistent with small real-world warehouses. The proposed meta-heuristic was also tested on a real case, for a Greek major electric retailer’s warehouse. His method was able to find solutions that were, on average, only 0.03% longer then the optimal routes for the generated instances. On the real case, the algorithm produced routes with approximately 17% to 23% shorter distances compared to the method used in practice by the company.

3.2.2. Order-picking systems

De Koster et al. (2007) identifies a number of variants for order picking systems. Considering order picking performed by humans, we can have *picker-to-parts* or *parts-to-picker*. *Picker-to parts*, the most commonly used system, consists of having the order picker travel the isles to collect the items. With *parts-to-picker*, unit loads are retrieved from their storage location by automated systems that bring them to the picking position. There, the picker takes the required quantity and the remaining loads are automatically stored again. The latter is not appropriate for supermarket stores, since these will also be receiving customers. Additionally, there can be *put systems*, in which the items, after picked in batch, are distributed over the customers' orders by order pickers. These are particularly useful in picking for orders with a large number of different items in a short time window.

Another distinction concerns the decision on which items to pick on a route. Here, we can have *discrete picking*, in which the picker collects all the items for a single order in each route, or *batch picking*, in which the picker receives a list of items from multiple customers orders. In the case of batch picking, two main alternatives exist to allocate the items to the customers' orders: *sort-while-pick* (when done immediately in the cart by the picker) or *pick-and-sort* (if done after picking, in a *put system* or *sorting system*).

Furthermore, the picking may be organized into zones. In *zone picking*, the warehouse is split into several areas, each with different pickers. These will only pick the portions of the orders assigned to their zone, after which the items are passed to the following zone or delivered to the depot.

Petersen and Aase (2004) study several combinations of storage, picking and routing policies for different order sizes, warehouse shapes, depot locations, and demand distributions. They conclude that *order batching* has the largest impact on reducing total order picking time. For their analysis, they assume that the picking cart has separate compartments to maintain order integrity and eliminate the need for downstream sorting. This constitutes a sort-while-pick process, which assures that the benefits from batch picking are not wasted on sorting operations. De Koster et al

(1999) also study the effect of batching in order picking efficiency, assuming the same sort-while-pick process. The simple and straightforward batching method *first-come-first-served* (FCFS), which combines the orders as they arrive until reaching the picking cart's capacity or the maximum batch size, led to major improvements over the discrete order picking method. Still, seed algorithms and saving algorithms performed better than FCFS in most cases: at best, they achieved 24% time savings; at worst, they caused a 3% increase in picking time. Nevertheless, Petersen and Aase (2004) advocate the use of FCFS, since more complex batching techniques are difficult to implement in real-world cases and have been dismissed by practitioners.

In their simulations, Petersen and Aase (2004) consider a picker walking speed of 150 feet per minute and picking time of 0.30 minutes per item. They argue that these estimates are consistent with observations and with the literature. However, Gray et al (1992) use a walking speed of 50 feet per minute and picking time of 0.29 minutes per item. In their simulation model, Gue et al (2006) use 20:1, 10:1 and 5:1 pick time to walk time ratios, stating that these are tipically 20:1 or less.

3.2.3. Warehouse layout design and storage assignment

As mentioned, Ratliff and Rosenthal (1983) address the problem of order-picking routing in a rectangular warehouse with crossovers only at the end of the aisles. Subsequent efforts insert additional complexity by considering decentralized depositing (De Koster and Van Der Poort, 1998), additional cross-aisles (Roodbergen and De Koster, 2001; Theys et al., 2007), and irregular shape warehouses (Zhang and Liu, 2009), which aim to reflect a larger number of real cases. These decisions report to the *internal layout design* or *aisle configuration problem*, which include the determination of the number of blocks and aisles, as well as the length and width of these aisles (De Koster et al., 2007).

Gue, et al. (2006) study the effect on order picking efficiency of performing this task in warehouses with narrow aisles, under different levels of activity. The narrow aisles do not allow workers to cross other workers in the same aisle, forcing them to stop and to be less eficient in their task. As expected, they find that, in general, congestion increases with the number of workers. Also, larger picking areas are less likely to be blocked. Congestion led to blocking time, ranging from 2% to 11% for the more reasonable scenarios. One less expected conclusion is that this is less of a problem when picking density is higher and pickers must stop more often. The reason for this is that a worker can only be blocked while travelling, so blocking is less of a concern when a worker spends more time picking and less time traveling. Although in a supermarket store we do not expect to see narrow aisles, these findings may relate to in-store order picking when the supermarket is crowded and the pickers must share the aisles with a large number of customers.

Figures 3.3, 3.4 and 3.5 illustrate the layouts of real supermarket stores in Portugal. From these examples we can understand that, although the stores' layout may vary in shape and size, they will usually consist mainly of parallel aisles with at least one cross-aisle at the middle, in addition to the front aisle and the back aisle. In many cases they will not be exactly rectangular, but will, instead, have an irregular shape. Contrary to what would be expected in a traditional warehouse, the aisles' length, their width and the number of shelves will not always be the same within a supermarket store: in the produce and the frozen products departments, for instance, most central aisles are frequently replaced with waist-level islands.

Figure 3.3: Supermarket store layout (15 000 sq ft).

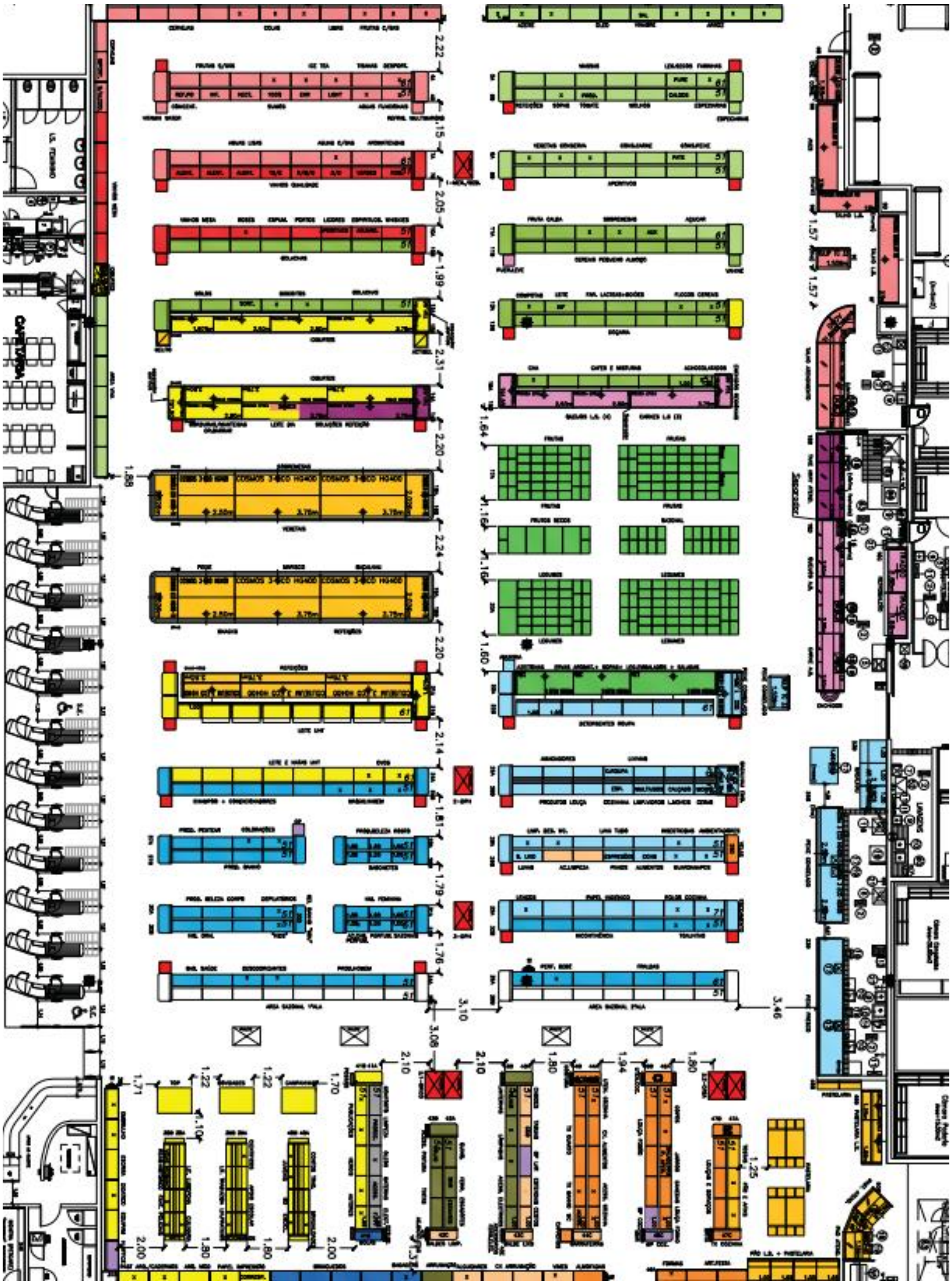

Figure 3.4: Supermarket store layout (25 000 sq ft).

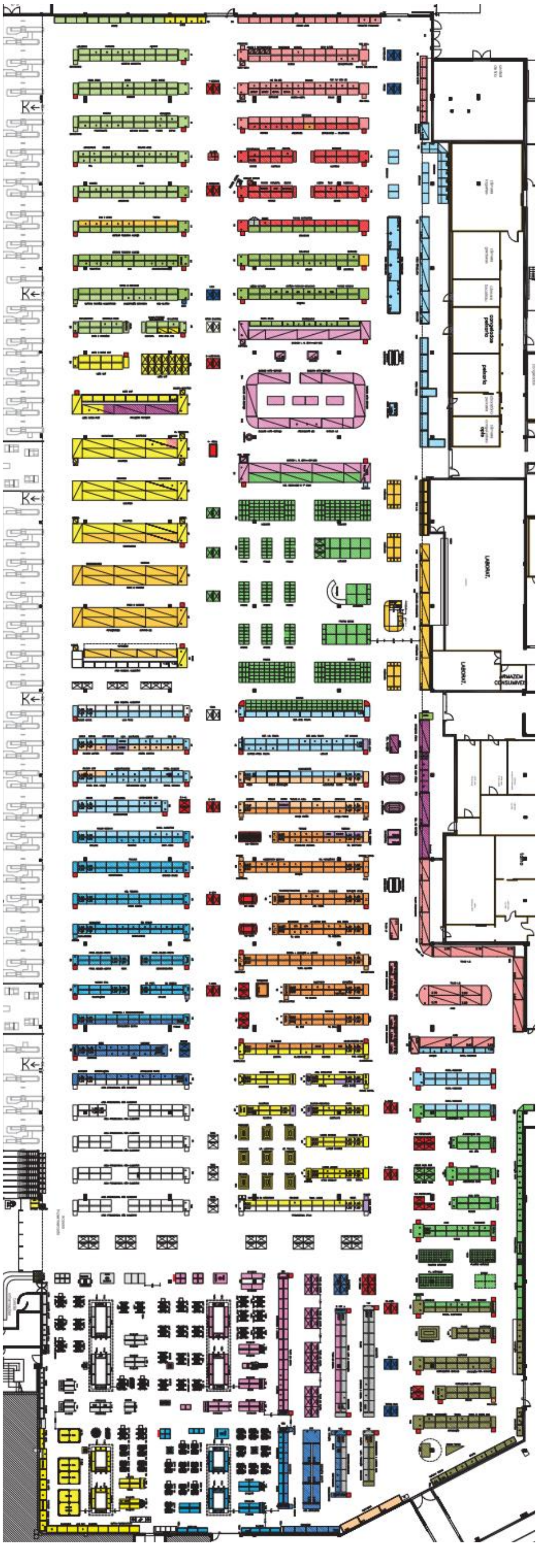

Figure 3.5: Supermarket store layout (100 000 sq ft).

Decisions about warehouse layout are closely related to the ones concerning storage assignment. Storage assignment refers to the decisions about the location of each product in the warehouse or store. Considering the numerous ways to assign products to storage locations, De Koster et al. (2007) mentions the most common policies:

i. *Random storage*, in which the items' location is randomly chosen between all available locations. This results in high space utilization (or low space requirements), but with increased travel distance.
ii. *Closest open location storage*, in which the items are stored in the first empty location found by the workers.
iii. *Dedicated storage*, in which each item has a specific and fixed location where to be stored. This results in low space utilization, but pickers become more familiar with the items' locations.
iv. *Full-turnover storage*, in which items are stored according to their turnover. Products with higher sales are located in the more accessible locations, usually near the depot.
v. *Class-based storage* borrows features from the other policies. Items are divided into classes (usually around three) according to their turnover. Each class has a dedicated area of the warehouse and within this area items are usually randomly stored. Fastest-moving items are generally referred to as A-items, the next class being B-items and so on. Following Pareto's 80/20 rule, A-items represent around 20% of the items and account for approximately 80% of the turnover. These items should be nearer the depot.
vi. *Family grouping storage*, in which the location of a group of items takes into account the associations between them. This means that items regularly ordered together should be close to each other.

Petersen and Aase (2004) find full-turnover and class-based storage to be more efficient than random storage, although these may increase picker congestion.

From the previous supermarket layouts, we conclude that in a supermarket store we will find a dedicated storage policy. Usually the store is divided into departments, each including a certain type of products: meat, fish, produce, frozen food, bakery, dairy, beverages, health & beauty, household, pets & plants, and other. Occasionally, we may find a few examples of what could be considered as family grouping, usually for promoting sales of a certain product in combination with another. Although supermarkets do not adopt class-based storage, congestion may be likely to occur in certain aisles, since a few products and departments are more likely to be visited by any given customer.

3.3. Capacitated Vehicle Routing Problem

The order-picking problem may be considered a variant of the popular Travelling Salesman Problem (TSP), in that it fundamentally consists of a routing problem: defining the best route for a picker to collect a set of items stored in distinct places within a warehouse and returning the starting point. We recall that the TSP consists of finding the shortest route for a salesman to visit a predetermined group of cities, where each of them is visited once and only once, and starting from and returning to a specified city or place (Dantzig et al, 1954).

However, when we take into consideration the limited capacity of the picking vehicle, it seems more adequate to frame it as a Vehicle Routing Problem (VRP), more specifically a Capacitated Vehicle Routing Problem (CVRP).

The problem refered to as the Vehicle Routing Problem may have been first described by Dantzig and Ramser (1959) as a generalization of the TSP. Their objective was to find a routing solution for gasoline delivery to multiple service stations supplied by a fleet of trucks, all departing from the same terminal, that fulfilled all clients (service stations) with minimum distance travelled by the trucks. In this paper, the authors propose the first mathematical programming formulation and algorithmic approach for the VRP. Eksioglu et al (2009) argue that, by introducing more than one vehicle for the first time, Dantzig and Ramser’s study may be considered to be the first VRP approach.

Clarke and Wright (1964) start from Dantzig and Ramser's formulation to arrive at a modified model that is applicable to fleets of trucks with varying capacities.

The VRP models developed since have become increasingly complex, to mirror real-life problems. Since the basic VRP (the Capacitated VRP) is known to be an NP-hard problem (Toth and Vigo, 2002), the additional complexities make them more and more impracticable to be solved through exact algorithms, except for small sized problem instances. As real-life applications are usually very large in scale, VRPs are usually dealt with through heuristics or metaheuristics (Jaegere, 2014). These complexities include distance constrained VRP, VRP with time windows, VRP with backhauls, VRP with pickup and delivery.

For the distance-constrained VRP, the capacity constraint of the CVRP is replaced by a length or time constraint, such that each route must not exceed a certain travelled distance or time. Distance constraints may be combined with capacity constraints, in the case of the distance-constrained CVRP. The VRP with time windows considers that each customer must be served within a given time interval. The vehicle must arrive at the customer and perform the service for a certain duration, all within that time interval, before moving on to the next customer (Toth and Vigo, 2002).

For VRP with backhauls (Goetschalckx and Jacobs-Blecha, 1989), customers are divided into two subsets: linehaul customers, each requiring the delivery of certain products, and backhaul customers, where a given quantity of products must be picked up. There are precedence constraints, imposing that backhaul customers can be served only when all linehaul customers have received their products. So, the vehicle leaves the depot and will first deliver all the products it carries to the linehaul customers of that route. Only then may he start collecting products from backhaul customers. The VRP with pickup and delivery (VRPPD) also includes delivery and pickup in the same route, but in this case, every customer or location has an associated delivery quantity and another pickup quantity. The vehicle must satisfy a set of transportation requests, each with a corresponding pickup point, a delivery point and a demand to be transported between the two. Thus, the vehicle will leave the depot and visit each location, where it may deliver a given quantity of a product (or people) it is carrying and it may then pickup another specified quantity (Desrosiers et al., 1986; Toth and

Vigo, 2002). The VRP with simultaneous pickup and delivery (VRPSPD) differs only in the fact that either the destination or the origin of all items are common (for instance, a central depot). A number of other variations may be constructed from these versions of VRPs, such as the VRP with pickup and deliveries and time windows (VRPPDTW) or VRP with backhauls and time windows (VRPBTW).

Another of such complexities refers to the consideration of varying travel times. This may occur, for instance, when the time required to go from one client to another is dependant on the time of day, e.g., during rush hour. This is refered to as the time dependent vehicle routing problem (TDVRP) and is an extension of the VRP to account for the effects of congestion. In reality, congestion may be caused by a large number of unpredictable events (accidents, weather conditions or other random events) or it may be recurrent (greater traffic in rush hours or other periodic hourly, daily, weekly or seasonal cycles). This second type may be more predictable and, as such, can be represented by a deterministic function of the distance between the points and the current time (Malandraki, 1992).

Over the years, a large number of methods have been proposed and applied to the several variations of the VRP, including CVRP. These include branch-and-bound algorithms, branch-and-cut algorithms, set-covering-based algorithms, constructive heuristics, two-method heuristics, improvement heuristics, and metaheuristics (Toth and Vigo, 2002). Toth and Vigo (2002) discuss six main types of metaheuristics that have been used to solve vehicle routing problems: Simulated Annealing, Deterministic Annealing, Tabu Search, Genetic Algorithms, Ant Systems, and Neural Networks. They conclude that tabu search seems to be the most effective approach for the problem being tested, including granular tabu search, a variation developed by Toth and Vigo. However, they also argue that the further development of genetic algorithms and hybrid ant systems may make them able to match tabu search's effectiveness.

This seems to align with Schmitt (1994), according to whom initial poor results of the basic genetic algorithms to solve order based problems such as TSP or VRP led to a late realization of the potential for this method. Thus, very little research was conducted initially, which led to slow advances for genetic algorithms on TSP and

VRP. The development of new encoding schemes and operators, however, made it a viable heuristic to apply to these problems.

Chapter 4: Genetic Algorithms

We begin this chapter with an overview of Genetic Algorithms (GAs), the solution methodology we propose for the problem in question. This is followed by a discussion of GA approaches to similar problems, namely the TSP, the VRP and the order picking routing problem. The chapter is concluded with a detailed description of the particular GA approach we adopt throughout this work, the Biased Random-key Genetic Algorithm (BRKGA).

4.1. GA overview

Genetic algorithms (GAs) are a form of search heuristics inspired by the Darwinian evolutionary process, which determines that the predominance of certain biological traits will occur by action of natural selection. This brings an explanation for the phenomena of adaptation and speciation. Charles Darwin (On the Origin of Species by Means of Natural Selection, 1859) described the concepts of natural selection and survival of the fittest, which hold that fitter individuals are more capable to be successful in the tasks they perform and to prevail over their environment. This means they will be more likely to survive and to reproduce at higher rates. Less fit individuals, on the other hand, will have lower chances of survival and reproduction. As generations ensue, it is expected that the population will be composed of individuals who are, on average, fitter and more capable to succeed. This illustrates that the population has evolved conditioned by fitness.

The concept of genetic algorithms was initially described by John H. Holland in the 1960s and developed throughout the following decades. Holland (1975) tries to understand the biological process of adaptation and to find new applications for that process. His book "Adaptation in Natural and Artificial Systems" sets the stage for the development of GA as a field of study, defining its theoretical foundations. Holland considers the universality of the mechanism, exploring diverse applications that extend beyond biology to fields such as economics, psychology, game theory or artificial intelligence. He demonstrates that computer programs can evolve, in similar ways as species do.

Other early contributions from different scientists were also important to the development of the topic, namely Ingo Rechenberg, Hans-Paul Schwefel, Bremermann, Fogel and also a few of John H. Holland's students, like Ken DeJong and David Goldberg (Reeves, 2003; Mitchel, 1995).

According to Mitchell (1995), the basic elements of most GAs are: a population of individuals, a selection process based on fitness, a crossover operator and a mutation operator.

Each individual is a potential solution for the problem and is evaluated according to its fitness. For this we need a "fitness function", an objective function that will measure the performance of the individual according to a set of parameters, thus estimating how well it solves the problem (Michalewicz, 1995; Mitchel, 1995; Busetti, 2007).

Inspired by the evolutionary process, for each iteration the GA holds a population of individuals. Each individual, as a potential solution for the problem, is evaluated by the fitness function and rated accordingly. From here, a new population will be formed, with a number of individuals being selected to move on to the next generation and other individuals being recombined through *crossover*. The probability of a given individual to be selected will be dependent on its fitness value: better solutions will have a higher probability to be selected. Recombination simulates the reproduction process, as a new offspring individual will be formed from the features of two parent solutions. Additionally, some members of this new population will undergo modifications that will result in completely new solutions. This modification process, known as *mutation*, arbitrarily modifies one or more features of the individual, to arrive at a solution with completely new information. Thus, it introduces some variability into the population.

Similarly to the effects of natural evolution, it is expected that with the succession of generations, the process of evolution conditioned by fitness will result in progressively better results until the population contains solutions that are optimal or near optimal.

Genetic algorithms are a generic problem-solving method that is independent of domain. Thus, it can be applied to virtually any real-world problem. According to Michalewicz (1995), GAs have been successfully applied to a number of optimization problems, including wire routing, scheduling, transportation problems, and traveling salesman problems. Also, although they are not always able to find a global optimum solution to a problem, they reach good solutions relatively quickly. On the other hand, GAs are susceptible to be have their populations dominated by genes from a few relatively high fit solutions, converging on local optima (Busetti, 2007).

GAs can differ in many ways. According to Reeves (2003), there are many possible variations in population size, initialization methods, fitness definition, selection and replacement strategies, crossover operators, and mutation operators. Some methods add new information to the chromosomes, such as age or tags, and some allow for varying population size or multiple populations.

Next, we mention genetic algorithm approaches to TSP, VRP, and order picking routing problems. This is followed by a detailed explanation of the GA approach we adopt for this work, Biased Random-key Genetic Algorithms.

4.2. GA approaches to related problems

Larrañaga et al. (1999) review the different representations of GA developed to solve the TSP by then:

i. *Binary representation*: each solution is represented by a sequence of locations, with each location being encoded as a string of bits. The tour 1-2-3-4-5-6 for a TSP with 6 cities would be represented by (000 001 010 011 100 101), as each city would be represented by 3-bit strings (the required number of bits for a binary representation of 6 numbers).

ii. *Path representation*: this is the most natural tour representation, in which each city is represented by its own number in the tour sequence. A tour with the sequence 3-2-4-1-7-5-8-6 would simply be represented by (3 2 4 1 7 5 8 6).

iii. *Adjacency representation*: here the tour is also represented by a list of cities, but each number in the list represents the following city to visit after the one corresponding to its position. For instance, city 7 in position 3 would mean that from city 3 the route would go to city 7. Thus, the tour 1-3-7-2-5-4-6-8 would be represented by (3 5 7 6 4 8 2 1): from city 1 to 3, from 2 to 5, from 3 to 7, from 4 to 6, and so on.

iv. *Ordinal representation*: each solution is represented by a set of coordinates that refer to an ordered list L of cities still to be visited. After selecting a city, say i, to be included in the solution, the corresponding list L is updated by removing city i from it. This way, repetitions, i.e., visiting cities more than once, are not allowed.

v. *Matrix representation*: each solution is a binary matrix in which the element in row i and column j is a 1 if, and only if, in the tour city i is visited before city j. In another variation, that element would be 1 if, and only if, city i is visited immediately before city j.

Binary and matrix representation use GAs' standard binary representation, which recurrently originates infeasible solutions with the use of classical crossover and mutation operators. Also, this faces some additional difficulties with large problem instances, as problem representations can grow unmanageable. Repair algorithms may need to be used to correct the resulting solutions. With path representation, classical crossover and mutation operators cannot be used, but a large number of new operators have been developed that are capable of providing good results (e.g. partially-mapped crossover, cycle crossover, order crossover, heuristic crossover, edge recombination crossover, displacement mutation, exchange mutation, inversion mutation). Ordinal representation allows for the use of classic operators, but both this and the adjacency representation have been known to lead to poor results (Larrañaga et al., 1999).

Schmitt (1994) studies the development of GAs to solve order-based problems, focusing on TSP and VRP with Time Constraints (VRPTC). Considering that very little research had been conducted on the application of GA specifically to the VRP (unlike the TSP), he tries to develop a GA that achieves results comparable to other heuristics for VRPTC. First, several distinct combinations of GA parameters are tested on the TSP, in order to understand the most efficient procedures and to gain insights in order to develop a GA approach for the VRPTC. The best performing configuration and design parameters on the TSP were then used in the first experiments with the VRPTC, however with disappointing results. The study was expanded through the hybridization of the GA with tour improvement functions. This new version compared favorably to traditional approaches for VRPTC in terms of solution quality. However, it performed substantially worse in terms of computational time. In the course of this study, Schmitt addressed several configuration issues which he defines as areas of conflict in the existing literature: population generation, population size, evolutionary strategy, population replacement strategy and crossover strategy. The findings of the study were that, in general, the best results were obtained by initializing the population with 50% potentially good solutions and 50% random solutions, population sizes larger than 200, elitist population replacement strategy and Asexual or Order 1 crossover operators. In order to reduce the computation effort, population size should be under 50 and a Cycle crossover operator should be used.

Baker and Ayechew (2003) try to demonstrate that GAs are an effective method to solve the basic VRP, a problem for which GAs were not as popular as other metaheuristics. They introduce a hybrid of GA with neighborhood search methods to solve the basic VRP, and test it against reported benchmark results using *tabu search*, *simulated annealing*, and *ant colony optimization*. They achieved objective function values only 0.5% above the best known results, on average, with good computational times. They use a non-binary representation, in which each solution is a set of index numbers that specifies the vehicle (and the related tour) to which each customer/location is assigned. For n customers and m vehicles, the chromosome for an individual solution will be a string of length n, with each gene value in the range $[1,m]$. This means that position i of the solution will contain the number of the vehicle that will visit customer i. Thus, the representation does not explicitly specify the exact route for each vehicle and only allocates customers to vehicles. A nearest-neighbor solution for the TSP is required for the sequencing of the customers for each route. In order to accelerate the convergence of the GA, a neighborhood search is performed on each individual as it is generated, either for the initial population or by the reproductive process.

Snyder and Daskin (2006) present a random-key genetic algorithm incorporating a local tour improvement heuristic to solve the generalized traveling salesman problem (GTSP). They consider the symmetric GTSP, in which the distance matrix is symmetrical, meaning that the distance from node i to node j is the same as the distance from node j to node i. Deriving from the suggested encoding of Bean (1994) for the GTSP, each gene has an integer part, to identify each node included on the tour, and a fractional part with the random keys to establish the sequencing. For each new generation, the elite individuals will be the top 20%, and these will be copied directly from the previous population. For the remainder, 70% will be obtained from crossover and 10% via immigration (randomly generated). They apply a local improvement heuristic to each new individual created by crossover or random generation, using 2-opt and “swap” operations to attempt to improve the solutions. The algorithm was tested on a set of 41 standard problems with known optimal solutions, and it was able to find an optimal solution for the majority of the problems considered. The authors believe that this GA compares quite favorably to other tested heuristics, since it is simple to implement and can easily be extended to incorporate alternative objective functions and constraints.

Zhang and Liu (2009) propose a modified GA to solve the order-picking routing problem in an irregular warehouse. They use a *path representation*, tracking not only the items' locations, but also the items' quantity. A data structure is used in which the orders do not have to be split into different batches prior to the development of the sequence. They assume a single batch with all the elements to pick and let the algorithm determine a sequence for these. However, prior to the evaluation of the original sequence, the model forces a trip to the depot whenever the vehicle exceeds its capacity. The model also inserts the depot's position at the beginning and after the last storage point of the sequence, meaning that the vehicle will necessarily start at the depot and return to it at the end. This creates the final version of the solution, which will then be evaluated. The GA operators will be applied to the original sequence, without the depot insertions. The authors argue this algorithm is preferred to other versions in which the trips to the depot are randomly inserted into the original sequence, generating infeasible solutions that exceed the cart's capacity and allowing for inefficient routes that do not utilize it completely. Although some of these inconveniences can be overcome with penalties for infeasible solutions, this approach makes for a more efficient algorithm that still arrives at better solutions, since the searching space is smaller.

Tasan and Gen (2012) strive to find an efficient solution for the VRP with simultaneous pick-up and delivery (VRP-SPD) with a genetic algorithm based approach. In order to test their algorithm, they transformed 24 known CVRP test problems into VRP-SPD by adding random pickup data. They find their approach to perform well and efficiently, since it achieves better solutions than the ones provided by CPLEX 10.0 solver in 1000 seconds, requiring significantly shorter computational times (only between 13% to 33% of the CPLEX time). The VRP-SPD is represented by *path notation* with all the nodes at once, which are then split into routes by the decoder. Similarly to Zhang and Liu (2009), this is accomplished by considering the vehicle's capacity, and splitting the routes whenever this capacity would be exceeded. For the VRP-SPD, the vehicle's capacity must not be violated by the items to deliver nor by the items to pick-up for each set of customers included in a route. However, this does not guarantee that the vehicle's capacity will not be exceeded after visiting a particular customer. Thus, the concepts of weak and strong feasibility are used: weak feasibility guarantees the total pick-up and the total delivery loads do not exceed the vehicle's capacity; strong feasibility assures that the load does not exceed the vehicle's capacity after visiting each customer. Weakly infeasible routes are prevented by the decoder, as described above, whereas strongly infeasible routes are allowed, but avoided by the use of high penalty costs. The initial population is generated by random permutations of the customers' nodes. For the following generations, individuals are picked by roulette wheel selection, with probability directly proportional to each individual's fitness value. Partial-map crossover and swap mutation are the genetic operators used in this study.

4.3. Biased Random-Key Genetic Algorithms

GAs differ from sequential search methods, among other elements, in the way new solutions are generated, and this comes with its drawbacks. Depending on the encoding of the solution and the crossover operator, crossing two feasible solutions does not guarantee feasible offspring (Reid, 2000). These infeasible candidate solutions are then penalized by the objective function or modified by repair algorithms.

A number of methods have been developed to overcome this issue, in order to ensure the feasibility of the solutions after all transformations that occur in the process. Highly specialized representations of the problem (decoder algorithms) may minimize the occurrence of infeasible solutions. However, both these decoders and the repair algorithms are usually computationally intensive, difficult to design, and highly problem-specific (Reid, 2000).

Random-key genetic algorithms (RKGAs), introduced by Bean (1994) for sequencing and optimization problems, are a method for representing solutions that guarantees the feasibility of any offspring if the parents are feasible. This is achieved with computational efficiency and without the need for specialized representations. RKGAs use random numbers, usually real numbers in the interval [0, 1], as tags representing the solutions. Thus, the chromosomes are built as strings or vectors of n real numbers, where n is the number of parameters representing the problem instance being solved.

If a feasible solution can be represented by a random key vector then any mutation and any crossover will inevitably also be feasible. The GA will search within the *n*-dimensional unit hypercube and the random keys will subsequently be mapped to the literal space for evaluation, through a decoder. This decoder will associate those chromosomes with a solution of the combinatorial optimization problem, which can be evaluated by the fitness function (Bean, 1994). This process is depicted in Figure 4.1.

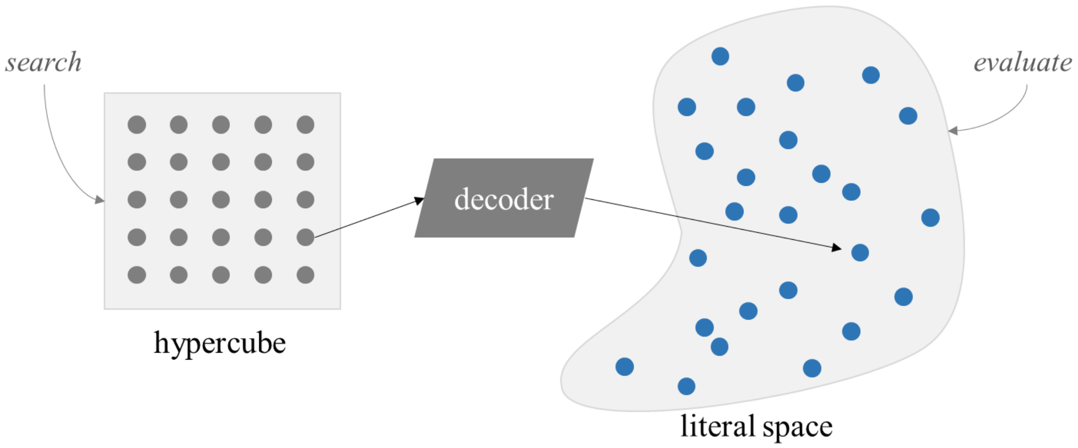

Figure 4.1: Random keys mapping process.

The initial population will be composed by a number of randomly generated random-key vectors. These will then be decoded, evaluated, ranked according to their fitness, and partitioned into two groups. The best individuals of the population (the elite individuals) will be copied unchanged to the next generation. Non-elite individuals will only contribute to the crossover procedure. Thus, RKGAs adopt an elitist strategy (Goldberg, 1989) to guarantee that the best solutions of the population are kept from one generation to the next. The algorithm randomly chooses two parents from the full population and through *parametrized uniform crossover* (Spears and DeJong, 1991), an offspring will be created composed of genes from both parents. Each individual gene will be randomly selected from one parent or the other, although at all times with higher probability towards one of them. Instead of mutation of existing individuals, RKGAs adopt the concept of immigration, where a number of completely new individuals is generated and added to the population. This way, through high immigration (or mutation) rates, RKGAs can overcome the tendency for rapid convergence of its elitist strategy (Bean, 1994). This transitional process is depicted in Figure 4.2.

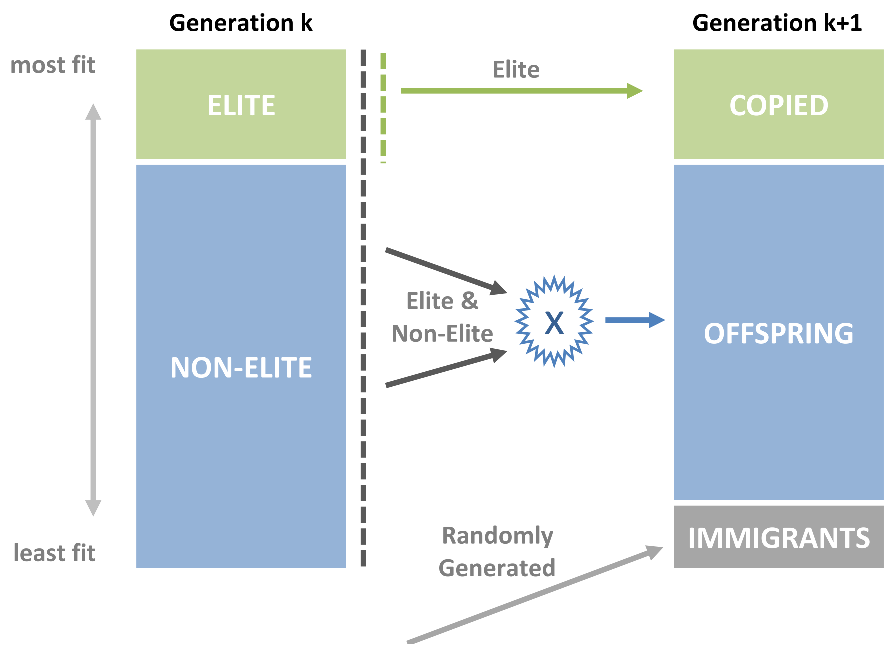

Figure 4.2: RKGA generational transition.

In this work, we propose a *biased random-key genetic algorithm* (BRKGA), based on the framework provided by Gonçalves and Resende (2011). BRKGAs are a variation of random key genetic algorithms that include bias towards the elite individuals in the crossover procedure. First, for the selection of parents, the BRKGA stipulates that one of them must necessarily come from the elite group. Thus, each offspring will have at least one parent randomly chosen among elite individuals only. The second parent is randomly selected from the entire population (elite and non-elite). In addition to this bias in the selection of parents, the parametrized uniform crossover will attribute to the best parent a higher probability in the selection of each gene of the offspring. Thus, BRKGA differs from RKGA in the fact that each offspring will have at least one elite parent and will probably inherit more genes from the fittest parent. The transitional process for BRKGA is depicted in Figure 4.3. Throughout the rest of this work, we will use the terms “chromosome”, “individual” and “solution” interchangeably to refer to either an order-picking tour or its representation in the BRKGA.

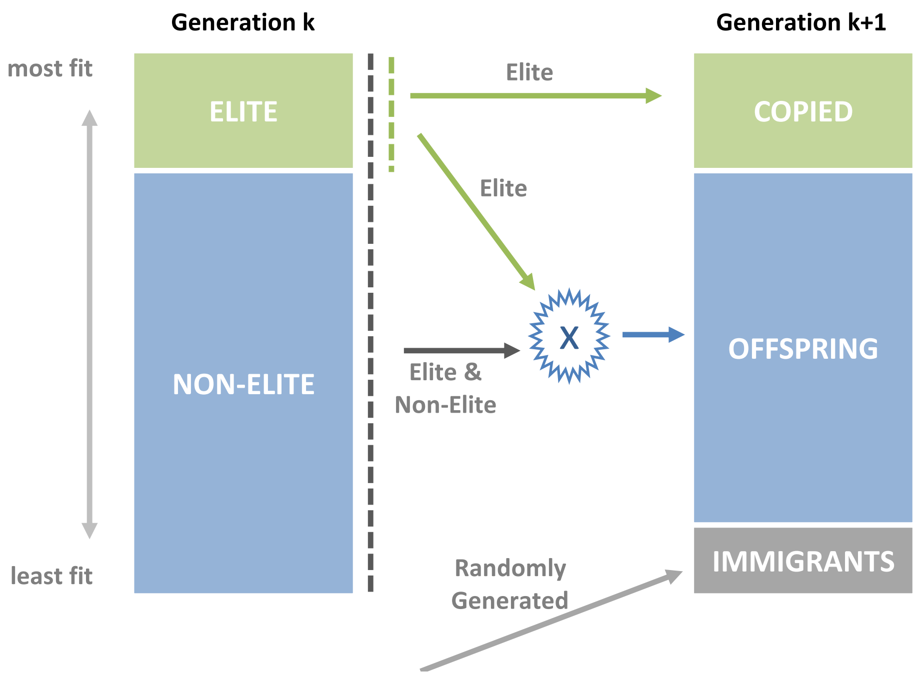

Figure 4.3: BRKGA generational transition.

Chapter 5: Problem description and formulation

This chapter starts by listing and analyzing the actions related to order placement and fulfillment, as well as the actions required to solve the order picking problem. Then, we provide a detailed description of the problem addressed, followed by its mathematical formulation.

5.1. Order placement and fulfillment

Analyzing the tasks involved with placing and fulfilling customers' orders will help understanding the problems that have to be addressed in the in-store order picking process.

i. The customer places the order, usually online or via a mobile device, by selecting a set of products and the corresponding quantities to be acquired. Customer choices take into account the items available in the specific store or warehouse that will be serving the order. The customer also defines the delivery time window and method. Based on the options provided by the grocer, the customer may choose home delivery, store pick-up or other.

ii. The grocer determines when each order has to be fulfilled. This is subject to the shipping time or the delivery time, depending on the method selected by the customer.

iii. The orders are grouped based on their delivery time and the amount of time needed to perform the picking.

iv. The picking team receives instructions regarding the next orders to be fulfilled. These instructions include the picking lists, specifying items and quantities. The items may be organized into discrete orders, batches or picking zones, depending on the picking strategy. Along with the items to collect, the team may also receive directions on the recommended picking itinerary.

v. The items are prepared for shipping or pick-up. After picking, the items need to be organized, which (depending on the picking strategy) may include sorting the items by order or placing a few items in separate containers (particularly, fresh and frozen products).

A number of problems arise from this process, which are somehow connected and mutually dependent. Two problems of particular relevance are discussed below:

1. Scheduling the orders, a case of the Job Shop Scheduling Problem (McMahon and Florian, 1974) in which each picking team is a machine and each picking itinerary is a job. Here, we need to minimize the total length of the schedule, assuring that the items' waiting time after picking is minimum.
2. Determining the best itinerary, or routing order pickers, a case of the Vehicle Routing Problem, or more specifically the Capacitated Vehicle Routing Problem (Dantzig and Ramser, 1959), considering the orders to be fulfilled and

their schedule. This problem may be addressed in several ways, depending on the picking policy:

a. To determine the best itinerary for each order, in a single-order picking policy.
b. To determine the best itinerary for the next scheduled orders with an order batching policy.
c. To determine the best itinerary for each class of items in a picking by zone policy.

This work is dedicated to the latter problem, the order picking routing problem, so only this one will be addressed in the following pages. This problem consists in finding the optimal routes for pickers to collect a number of items. Given a specified number of products and their corresponding quantities, the goal is to establish a sequence for the picker to visit the product's locations in order to minimize the distance traveled.

Single-order picking and *picking by zone* policies have their own advantages and disadvantages, but it seems evident that both are particularly susceptible to great inefficiencies in some cases. A main advantage of *single-order picking* is its practical simplicity, since it is a very straightforward process, easy for the workers to comprehend and adopt. However, with *single-order picking*, it is plausible that the pickers will have to make similar tours repeatedly to retrieve similar items for different orders. If one item is near another item from a different order, the picker will not collect it, even though little time and effort would be required to do it. On the other hand, with *zone picking*, items on distinct sides of zone borders would also not be picked on the same tour, despite being near. One of the main advantages of *zone picking*, avoiding picker congestion, is less relevant in the case of in-store order picking, since the number of order pickers in a given store is expected to be low. Congestion should be normal (particularly on rush hour) due to a greater customer affluence, which should not be significantly affected by order picking policies.

Furthermore, depending on several variables, such as the store layout, the number of orders to pick, store congestion and others, it is plausible that single-order or zone picking policies may be preferable at different moments. Thus, a more flexible picking policy could reduce the disadvantages of these policies and maybe retain some of their advantages. This may be achieved through *batch picking*, which is indeed the policy we study in this work.

A *batch picking policy* usually requires a batching method: a procedure to determine which orders should be grouped together for a picking tour. This also means that, although each tour typically includes more than one order, each order may not be entirely collected by a single tour.

As a requirement of the structure of the data available to test our BRKGA, we are forced to assume that the process of order batching is dealt with separately, and thus the batching information is considered as an input. This means that we receive a picking list, in the form of a pool of items to be picked, regardless of their corresponding order. All of the items to be picked are considered as a whole, and the algorithm will look for the most efficient set of itineraries to pick all items. It should be noticed that, this way, orders may be served by any number of routes and routes can include items from any number of orders. Such a strategy implies the need for a process of order accumulation, with the purpose of sorting the items by order prior to their shipment. Therefore, a *pick-and-sort* strategy is considered. In a pick-and-sort strategy, the picker collects a number of items, irrespective of their orders. After the picking process is concluded, a sorting process has to take place to associate each item with its respective order before shipment. The order accumulation process, i.e. sorting, is out of the scope of the problem regarded here, since we address the picking stage of the problem only.

5.2. Detailed Description

The present work aims to determine the most efficient picking routes that allow the supermarket/grocery to collect the next orders to be dispatched or picked-up. Considering their resources (particularly the number of pickers), the fastest the grocer can collect all the orders and have them ready for shipping, the more efficient the process will be considered.

After receiving customer orders, the shipping schedule will firstly determine the orders which must be picked next. The items from these orders will form a picking list, which may result in several routes. This picking list results from the previously performed batching of customers' orders, as illustrated by the example in Figure 5.1. In this example, there are 10 items, resulting from three distinct orders. The first order includes items 37, 02, 71 and 52; the second order consists of items 71, 48 and 92; the third order contains items 44, 16 and 29. It should be noted that item 71 is included in two different orders, so it appears twice in the picking list.

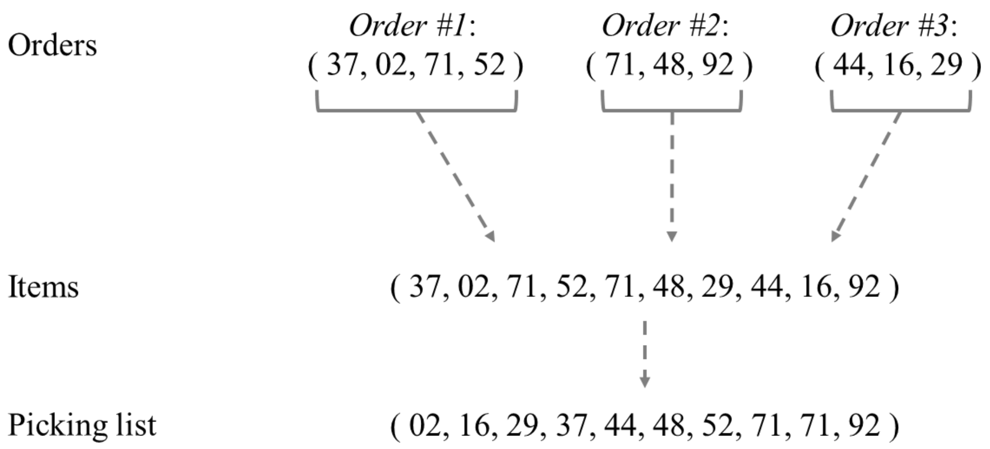

Figure 5.1: Example of order batching into picking lists.

After the picking, the items will be put into separate containers, each corresponding to a specific customer order. As said before, here we assume that the previous batching process and the subsequent sorting process are dealt with separately and out of the scope of this work.

Although we are trying to minimize the time required to pick all the items for a set of orders, our efficiency criteria will not be time, but travelled distance. The picker's walking speed, or the time it takes to stop the cart, pick the item, and resume walking, are parameters which are independent of the route itself. These depend only on the picker and the quantity, type, and placement of the items to collect. This means that it ultimately depends on the composition of the orders and the picker's walking and working speed, which is the same regardless of the elected picking route. Thus, the only variable that influences the final picking time, considering the identified pickers and orders, is the overall distance travelled by the pickers.

The first step for determining the most efficient picking routes is to transpose the store's layout and storage points (where each item is stored) into a matrix that can easily be read and interpreted by the model. This matrix represents the minimum distance between each pair of locations in the supermarket. These locations (nodes) are the depot (starting point and unloading dock), the storage points, and the endpoints of the aisles (which may or may not also be storage points).

Then, it is necessary to determine the sequences that allow the pickers to collect all the items in the orders, while minimizing the total distance they travel. For each sequence, the picker leaves the depot with an empty cart, travels to each item location, collects the item in, at most, the specified quantity, and moves on to the next item location. The picker keeps collecting items until the cart available capacity is reached. After the last item, the picker returns to the depot to deliver the items so they can be ready for shipment.

5.3. Mathematical Formulation

A solution for this problem consists of a set of routes such that every item is picked on one of the routes, in the desired quantity, without violating the picking cart's capacity. Amongst all possible solutions, we are interested in finding one that minimizes the sum of the total distance over all routes. The number of routes has no limitation as we consider an unlimited number of pickers.

Before providing the mathematical model for this problem, we introduce some notation.

Sets and Indices:

I: set of locations for the depot and the *n* items, n+1 = |I|;

i: locations index, including items and depot, i Є I;

k: route index.

Parameters:

$d_{ij} - \ distance\ between\ the\ location\ i\ and\ the\ location\ j, with$ i, j Є I

$D_i - \ required\ quantity\ of\ item\ in\ location\ i, with\ D_{depot} = 0\ and$ i Є I

$C_i - \ volume\ per\ unit\ of\ item\ in\ location\ i, with\ C_{depot} = 0\ and$ i Є I

$C - cart\ capacity\ (volume).$

Decision variables:

$$y_{ij}^k = \begin{cases} 1, & if\ in\ route\ k\ the\ picker\ goes\ from\ location\ i\ to\ location\ j; \\ 0, & otherwise; \end{cases},$$

$with$ i, j Є I

$x_{ik} = quantity\ of\ item\ in\ location\ i\ picked\ by\ route\ k,$

$with\ x_{depot,k} = 0, \forall\ k\ and$ i Є I

Model:

Minimize $$\sum_{\boldsymbol{i}\in\boldsymbol{I}}\sum_{\boldsymbol{j}\in\boldsymbol{I}}\sum_{\boldsymbol{k}}\boldsymbol{d_{ij}}.\boldsymbol{y_{ij}^{k}} \qquad (\mathbf{1})$$

Subject to:

$$\sum_{k} x_{ik} = Di, \qquad \forall\, i \in I; \qquad (2)$$

$$\sum_{i} x_{ik}\, .\, C_i \leq C, \qquad \forall\, k; \qquad (3)$$

$$\begin{cases} \sum_{i\in I} y_{ij}^{k} \leq 1, & \forall\, j \in I, k; \\ \sum_{i\in I} y_{ij}^{k} - \sum_{i\in I} y_{ji}^{k} = 0, & \forall\, j \in I, k; \end{cases} \qquad (4)$$

$$x_{ik} \leq M \sum_{j\in I} y_{ji}^{k}, \quad \forall\, i \in I, k; \text{for M sufficiently large} \quad (5)$$

$$x_{ik} \geq 0,\ y_{ji}^{k} \in \{0,1\}, \qquad \forall\, i \in I, k. \qquad (6)$$

The objective function, which is given by Equation (1), states the minimization nature of the problem and represents the sum of total distance traveled along all routes. Equation (2) imposes that every item I s collected and in the required quantities, while Equation (3) forces each route to comply with the cart’s capacity. Equations (4) force routes to visit at most once each item location and to be a closed tour. They also impose that routes must start and end at the depot. Constraint (5) ensure that if items are picked on a certain route then that route must visit the items' location. Finally, Equation (6) defines the nature of the variables.

Chapter 6: Solution Methodology

In this chapter we present the Biased Random-key Genetic Algorithm approach (Gonçalves and Resende, 2011) developed for the order picking problem we are trying to solve. We begin with an overview of the proposed solution methodology. The overview is followed by a discussion of the biased random key genetic algorithm, including the detailed description of the solution encoding and decoding, as well as of the evolutionary process and the solution generation procedure.

6.1. Overview

Given the combinatorial nature of the order picking problem and its inherent complexity, we chose to model the problem using genetic algorithms based on random keys, since it has already been applied successfully to several complex problems, see, e.g., Fontes and Gonçalves (2015, 2013), Gonçalves and Resende (2013), and Gonçalves, Costa, Resende (2014). This approach combines a genetic algorithm with a solution generation procedure.

In general terms, the approach consists of the following three phases:

i. *Decoding the parameters generated by the BRKGA*. This first phase uses the BRKGA to set the sequence in which the items to be picked are assigned to the routes;

ii. *Solution generation and evaluation*. This phase uses a procedure which builds a solution based on parameters generated by the BRKGA in the previous phase. After being constructed, the solutions are evaluated according to the objective function defined in equation (9);

iii. *Evolution of the parameters used by the solution generation procedure*. This phase uses the genetic algorithm to evolve (improve) the parameters used in the previous phases.

Figure 6.1 illustrates the sequence of steps applied to each chromosome generated by the BRKGA.

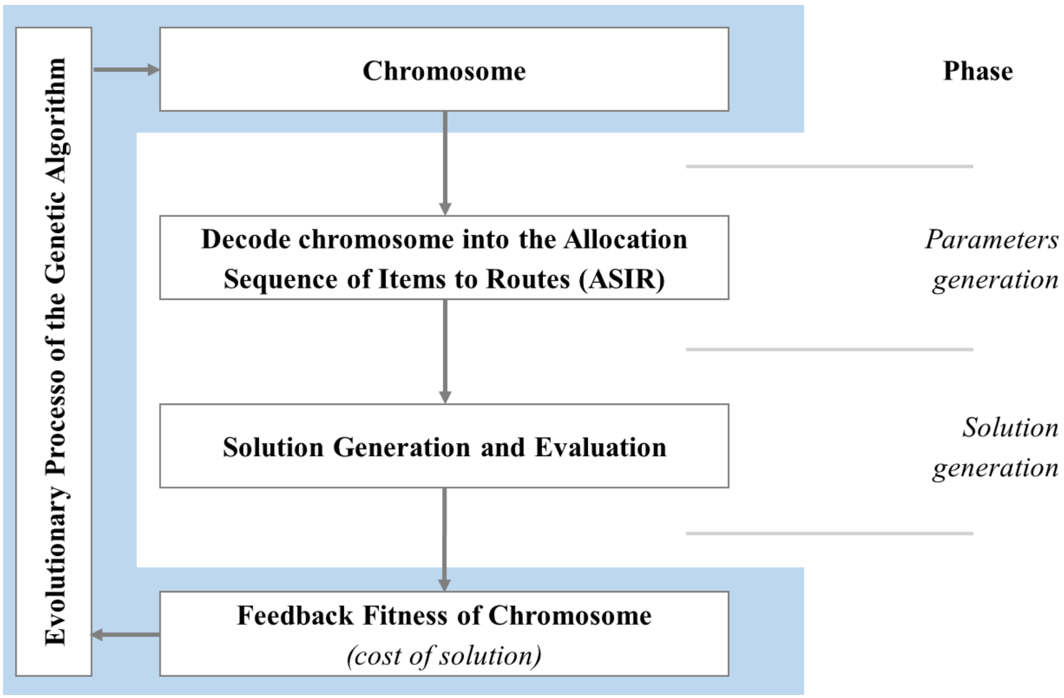

Figure 6.1: Architecture of the algorithm.

6.2. Biased Random-key Genetic Algorithm

This section describes the genetic algorithm and the evolution strategy in detail.

6.2.1. Chromosome representation

At the core of the model is an evolutionary program (a genetic algorithm using random keys), whose first step is to generate a set of solutions. Each solution is represented by a chromosome made of all the items to be collected (the picking list) in a particular sequence. Thus, a chromosome is made up of N genes, where N is the number of items in the picking list. Each gene has a random key number in the real interval [0, 1]. Figure 6.2 depicts the structure of the chromosome.

$$\textbf{CHROMOSOME} = (\underbrace{\textbf{gene 1, gene 2, ..., gene N}}_{\substack{\textbf{Allocation Sequence} \\ \textbf{of Items to Routes (ASIR)}}})$$

Figure 6.2: Solution encoding.

6.2.2. Allocation sequence of items to routes decoding

The *N* genes in the chromosome are used to determine the vector containing the allocation sequence of items to routes (ASIR). The decoding (mapping) of the *N* genes into a sequence is accomplished by sorting the values of the genes in ascending order. The indices of the sorted gene values correspond to the sequence in which the items will be assigned to the routes. Figure 6.3 presents an example of the decoding process for the ASIR. As illustrated, for each item in the picking list a gene will be generated, with a random key. By ordering the chromosome genes in increasing order of the gene values we obtain the following ASIR = (29, 52, 48, 02, 44, 71, 92, 71, 16, 37). This means that item 29 will be the first to be allocated to a route, followed by item 52, then item 48 and so on, ending with item 37.

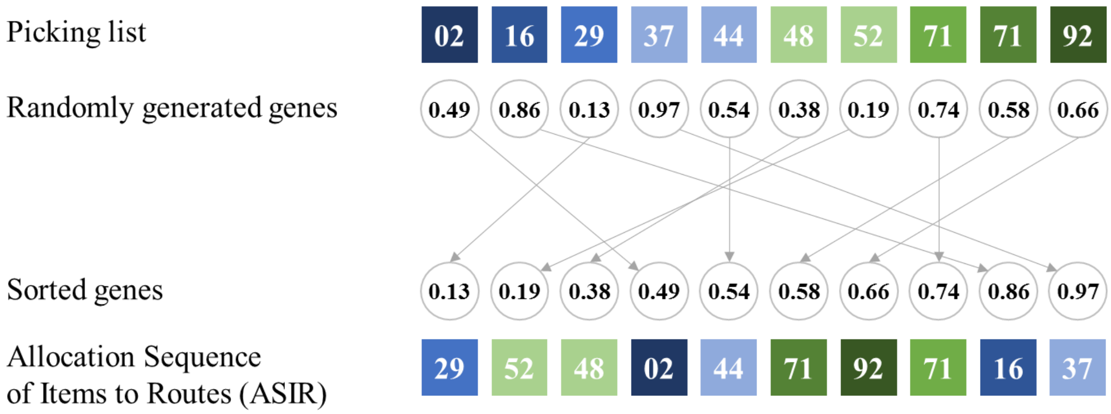

Figure 6.3: Decoding of the allocation sequence of items to routes (ASIR).

6.2.3. Solution generation procedure

The construction of the solution is done by allocating the N items to the routes, according to the ASIR vector. Each route *k* has a maximum capacity *C*, corresponding to the capacity of the picking cart.

At any given moment only one route (cart) is active. An item is allocated to the route which is active at the time of its allocation. Items are inserted into the active route as long as the route has available capacity. Items are added to the active route while the cumulative volume is smaller than the capacity of the cart. When the item being allocated exceeds the capacity of the cart the active route is closed and a new route is opened with the capacity *C*. The current item is allocated to the new route. The process goes on until all items are allocated to a route. Figure 6.4 depicts an example in which the maximum cart capacity is C = 15 000.

		Cumulative Volume		
ASIR	**Volume**	**Route 1**	**Route 2**	**Route 3**
29	2 495	2 495		
52	1 112	3 607		
48	6 779	**10 386**		
02	5 616	16 002	5 616	
44	846		6 462	
71	1 944		**8 406**	
92	9 099		17 505	9 099
71	1 944			11 043
16	1 098			12 141
37	131			**12 272**

Figure 6.4: Allocation of items to routes.

The first item to be allocated to route 1 is item 29. This item's volume is 2495, so the cart´s available capacity is reduced from 15 000 to 12 505. The next items to be allocated are item 52 (with volume 1112) and item 48 (with volume 6779). At this point, we have already occupied 10 386 of the cart's capacity, so there is 4614 left. The following one is item 02, with volume 5616. Since the cart can no longer contain this item, route 1 is closed and route 2 becomes active, receiving item 02. Items 44 and 71 complete this route, since by adding item 92, the cart capacity would be exceeded. Route 3 is created and this route will accommodate the remaining items 92, 71, 16 and 37.

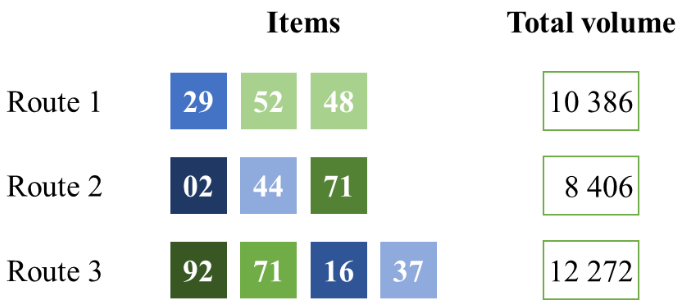

Figure 6.5: Routes' items and volumes.

Figure 6.5 illustrates this solution, in which route 1 and route 2 pick 3 items each and route 3 picks 4 items. This is the solution that materializes the ASIR example presented in Figure 6.3.

The final step, before the feedback of the fitness of the chromosome to the evolutionary process, is to calculate the total distance travelled to pick all the items. Figure 6.6 presents the distances between every pair of items locations (including the depot) for the routes constructed by the ASIR solution generation procedure.

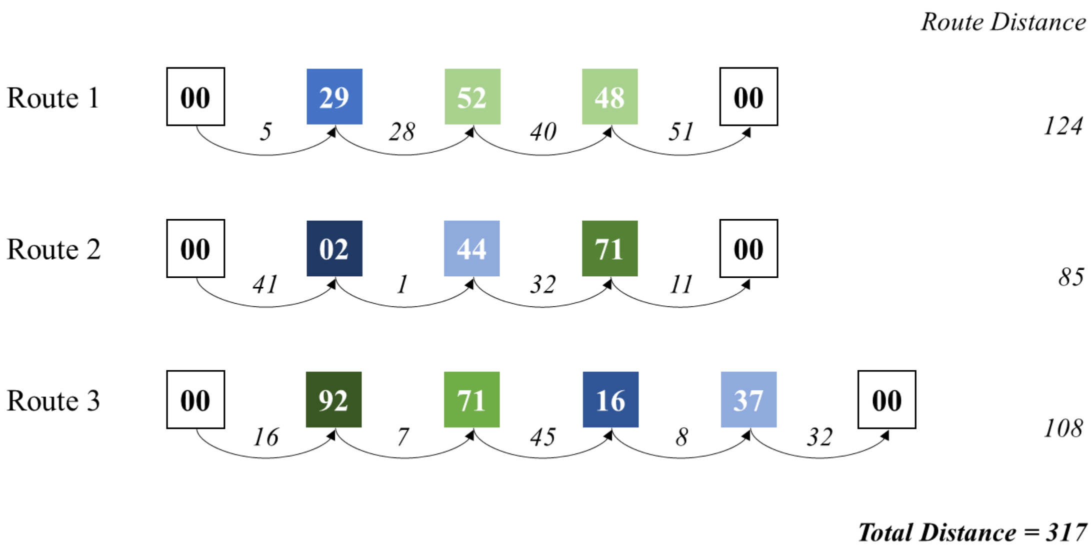

Figure 6.6: Total travelled distances for the solution.

As it can be seen, the total travelled distances for routes 1, 2, and 3 are 124, 85, and 108, respectively. Therefore, the total travelled distance, for this example, is 317.

6.2.4. Fitness measure

Genetic algorithms are based on the notion of survival of the fittest. This implies the need to evaluate each individual within a population in order to determine its fitness. The concept of fitness, here, relates to how suitable each individual is for solving the problem at hand.

The evaluation is performed through the use of a fitness function, a measure of the quality of the solution. For our order picking routing problem the fitness of the solution is given by the total distance travelled by the picker, as defined by Equation (1) in Chapter 5:

$$\text{Cost} = \sum_{\boldsymbol{i,j,k}} \boldsymbol{d_{ij} . y_{ij}^{k}} \quad ,$$

where, as previously defined, d_{ij} is the distance between location i and location j and y^k_{ij} is a binary variable assuming the value one if route k includes the trip from location i to location j and zero otherwise.

After generating a population, the model then evaluates each individual with the help of this fitness function. The lower the cost is, the better the solution is in solving the problem and the fitter that individual will be considered, when compared to the rest of the population. This information is determinant for the subsequent steps of the evolutionary process, by allowing the model to sort the population and identify the elite individuals (TOP) for the reproduction procedure.

6.2.5. Evolutionary process

The initial population of chromosomes is randomly generated. Each individual from this population will be evaluated and subsequently a new generation will be formed. Each new generation will result from the transformation of the previous one, through the application of genetic operators.

Reproduction is primarily achieved using an elitist strategy (Goldberg, 1989) in which the best individuals of a population (TOP) are copied into the next generation. For this, the population is previously sorted according to how efficient each individual solution is. The advantage of this strategy is to ensure that the best solution will be continuously improving from generation to generation.

The crossover operator adopted for this approach is a *parameterized uniform operator* (Spears and DeJong, 1991), instead of the traditional single, double or multi-point crossover. In this form of crossingover, chromosomes (parents) are randomly chosen to produce an offspring chromosome. One parent is randomly selected from within the best individuals of the population (TOP), while the other is randomly selected from the whole population, including the best individuals (Gonçalves and Resende, 2011). Then, for each gene of the new chromosome, one of the parents is selected to contribute with the value of its corresponding gene. A probability parameter is associated with the best parent, typically favoring it. For each gene a random number between 0 and 1 is generated. If the number generated is less than or equal to the probability parameter, the gene value is taken form the best parent; otherwise, it is taken from the other parent. Assuming the probability parameter to be 0.7, Figure 6.7 exemplifies the crossover process of two chromosomes with 4 genes and the resulting offspring.

	gene 1	*gene 2*	*gene 3*	*gene 4*
Parent Chromosome #1	**0.30**	**0.55**	**0.79**	**0.87**
Parent Chromosome #2	**0.26**	**0.91**	**0.12**	**0.44**
Randomly generated values	*0.58*	*0.36*	*0.89*	*0.25*
Probability parameter = 0.7	<0.7	<0.7	>0.7	<0.7
Offspring Chromosome	**0.30**	**0.55**	**0.12**	**0.87**

Figure 6.7: Parameterized uniform crossover example.

If we look at the example in the previous figure we understand that the first gene of the offspring (0.30) was acquired from the first parent. This is because the random value generated for the first gene (0.58) is lower than that of the designated probability parameter (0.7). The same happens for genes 2 and 4. Gene 3 (with the value 0.12) is inherited from the second parent, since the random value (0.89) is above 0.7.

The mutation operator employs the concept of immigration in place of the low probability gene to gene mutation. At every generation new members (MUT) are generated from the same distribution used to generate the initial population. This process aims to prevent premature convergence of the population, by inserting completely new solutions in each generation. This new genetic material ensures the diversity of each population.

Figure 6.8 illustrates the evolutionary process for the transition from one generation to the next.

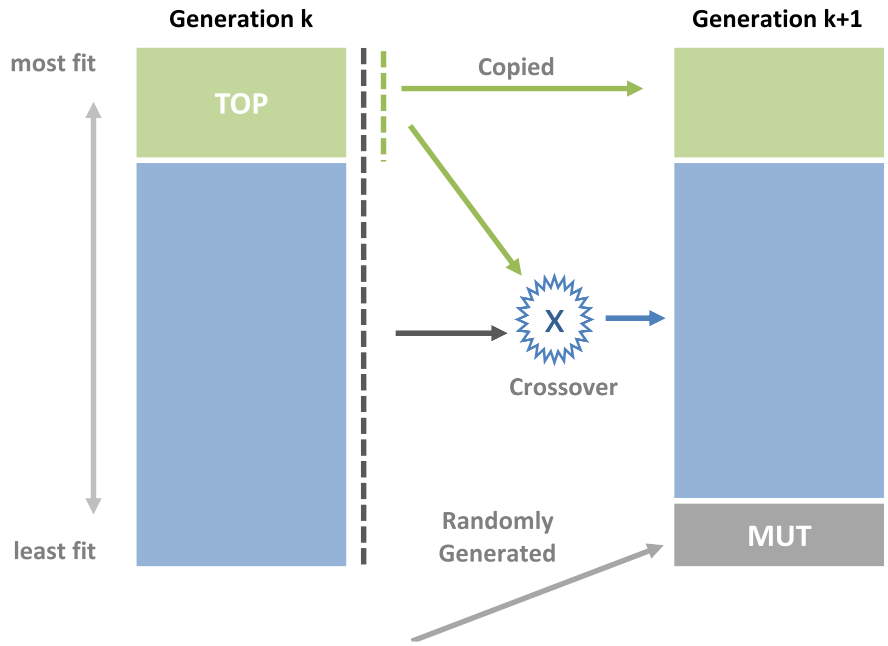

Figure 6.8: Generational transition.

Each new generation consists of the TOP individuals from the previous generation, the offspring chromosomes resulting from crossover, and the new randomly generated MUT chromosomes. This new population of chromosomes is used to obtain new solutions, which are then evaluated through the fitness function. These chromosomes are once again ranked according to its fitness and the transition process restarts. This will take place until the stopping criteria is met, which in the present case happens when 100 populations have been generated.

Chapter 7: Numerical experiments

This chapter begins with a description of the benchmark instances used to evaluate our model's performance. Then, the parameters we use for our BRKGA are detailed and, lastly, the numerical results are presented and compared to benchmark results.

7.1. Problem instances

7.1.1. Real-world data

Official and real-world data from traditional supermarkets, online supermarkets and e-groceries shown to be extremely difficult to obtain. No publicly available data was found, so several attempts were made at the most relevant supermarket chains operating in Portugal:

- Continente
- Jumbo
- Pingo Doce

- Lidl
- Minipreço

Requests for collaboration were submitted to these companies, in order to obtain generic and unidentified store layouts and transactional data. With this information we expected to be able to find efficient order picking routes for representative customers' orders. This could also help us understand the efficiency gains that would arise from doing order picking with the assistance of an efficient heuristic:

i. When compared to the methods currently employed by the companies that offer online grocery shopping;
ii. When compared to order picking with no assistance, in which the picker would have to rely merely on the store layout and the product's location;
iii. When compared to the estimated time an average customer would spend collecting all the items for the same orders.

Unfortunately, cooperation with any of the aforementioned supermarkets was not possible. All feedback pointed in the same direction: no data disclosure was authorized, not even data that was considered not to be sensible nor confidential in nature. Currently, there seems to be a generalized resistance to disclosing information within this industry, even when the source of the information does not need to be identified. This resistance is all the more evident on data regarding the online business.

7.1.2. Benchmark instances

With no real-world data available, the remaining option was to turn to literature data that would display similar features, thus allowing us to arrive at relevant conclusions about our model.

Despite this data not being specific to the e-groceries order picking problem when performed in a supermarket store, using existing and already tested data had a significant advantage: since our results could be compared to previous results for the same instances, it allowed us to understand how efficient our approach would be in finding good solutions for these order picking problems.

The instances used in this work to test our approach are those generated by Chabot et al. (2015).

Chabot et al. (2015) try to find an efficient solution for order picking in a particular case: the warehouse activity of a Canadian company that delivers furniture and electronic equipment. For this, they develop an adaptive large neighborhood search (ALNS) heuristic and compare it to an exact branch-and-cut algorithm and to the company's current decision-making procedure. The picking method currently being used by the company consists of always choosing the product in the farthest section at the highest level. Then, the next closest item is selected sequentially until the capacity constraint is reached. The ALNS is a method composed of a set of simple destruction and reconstruction heuristics that aims to find better solutions at each iteration. To establish the initial solution, the authors implement a sequential insertion heuristic that performs a greedy search for the best insertion for one product at a time. Then, for each iteration, the ALNS applies a *destroy operator* and a *repair operator*. This new solution will be exposed to a simulated annealing acceptance criterion in order to establish if it should be adopted or discarded in favor of the existing solution.

The instances were generated considering the features of real warehouse configurations. Five different types of aisles were created with 12 to 60 sections and 96 to 480 locations, as presented in Table 7.1. All configurations include 4 levels, which may not always correspond to the number of shelves in a supermarket aisle.

Type of aisle	Sections	Levels	Locations by side	Total locations
1	12	4	48	96
2	24	4	96	192
3	36	4	144	288
4	48	4	192	384
5	60	4	240	480

Table 7.1: Description of aisle types.

For each type of aisle, 10 sets of instances were generated, each having 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 picking items. Each of these 10 sets contains 5 instances, with a total of 250 instances for the five types of aisles, as shown in Table 7.2. Each instance corresponds to a picking list: a list of items to be picked by a specified due date or at a given moment.

Types of aisles	Number of sets per type of aisle	Number of instances per set	Total instances
5	10	5	250

Table 7.2: Number of instances.

7.2. BRKGA configuration

The population is composed of a number of chromosomes corresponding to 15 times the number of items, each chromosome having as many genes as the number of items to be picked. The parameter TOP is 15%, corresponding to the percentage of the number of chromosomes in a population that is selected as the Elite and will be directly copied into the next generation. The parameter MUT is set as 20%, meaning that each new generation will have 20% of completely newly generated chromosomes. For the crossover procedure, the probability of choosing a gene from the best parent is 0.7. The stopping criteria is 100 generations, meaning the BRKGA runs until 100 populations were created.

A summary of the BRKGA configuration parameters is provided in Table 7.3.

Parameter	Description	Value
p	population size	15 x *N*
TOP	size of elite population	15%
MUT	size of mutant population	20%
$\boldsymbol{\rho_e}$	probability of selecting the gene from the TOP chromosome	0.7
Fitness function	Equation (1)	$\text{Cost} = \sum_{\boldsymbol{i,j,k}} \boldsymbol{d_{ij}} . \boldsymbol{y_{ij}^k}$
Stopping criterion		100 generations

Table 7.3: Configuration parameters for the BRKGA.

7.3. Experimental results

To evaluate the performance and capabilities of our BRKGA algorithm and compare it to the results achieved by Chabot et al. (2015) and their adaptive large neighborhood search (ANLS) heuristic, we performed a series of computational experiments. The

BRKGA described in Chapter 6 was implemented in C++ and ran 10 times for each instance. The numerical experiments were conducted on a computer with an Intel Core i7-2630QM @2.0GHZ CPU running the Linux operating system with Fedora release 21.

The results of the BRKGA are presented in Table 7.4, along with the results for the same instances published by Chabot et al. (2015) using the ALNS.

For 5 of the 10 problem sets considered, the BRKGA matches the results achieved by the ALNS. This is the case for sets 10, 20, 30, 40, and 60. For sets 50 and 80, the ALNS obtained better quality solutions. However, for sets 70, 90, and 100 the BRKGA outperforms the ALNS.

	BRKGA		**ALNS**	
	Best Value	**Avg Time**	**Value**	**Time**
10	282.08	0.56	282.08	0.24
20	509.84	0.61	509.84	0.28
30	674.48	1.07	674.48	0.60
40	861.60	1.98	861.60	1.20
50	980.56	3.15	969.76	1.76
60	1 089.68	5.28	1 089.68	2.76
70	1 320.15	8.14	1 322.40	3.84
80	1 452.01	9.17	1 451.36	5.00
90	1 628.17	13.89	1 639.92	6.80
100	1 770.54	15.77	1 779.28	8.56

Table 7.4: Compared results of the experiments.

Regarding the required computational time to achieve the solution, the ALNS appears to be faster, based merely on the times presented in Table 7.4. However, since the experiments were performed on different computers with distinct characteristics, the results are not directly comparable. It should be noticed that the ALNS was implemented on machines equipped with two Intel Westmere EP X5650 six-core processors running at 2.67 GHz (Chabot et al., 2015). Thus, if we discount the fact that the ALNS computations were performed on a Dual CPU, each CPU having higher clock speed (2.67 GHz each on the Intel Westmere EP X5650 versus 2.00 GHz on the Intel Core i7-2630QM) and a larger number of cores (6 cores each on the Intel Westmere EP X5650 versus 4 cores on the Intel Core i7-2630QM), it seems that the BRKGA requires a smaller number of calculations for each solution (for useful information regarding relative speeds of different CPU processor, we refer the reader to the PassMark Software[1] CPU benchmark site).

Thus, the RBKGA was able to find the best results more often than the ALNS. In addition, it seems to be faster, considering the computational times required, as well as the characteristics of the computers used in each experiment.

However, we have well-grounded reasons to assume that even better results could be achieved through the hybridization of the BRKGA with a local improvement heuristic. The use of local search within global search algorithms, as is the case of GAs, allows for the intensification of the search around good quality solutions. This reasoning is substantiated by findings such as those of Baker and Ayechew (2003), Buriol et al. (2005), or Snyder and Daskin (2006). Such a local search procedure could be applied, for instance, to each element of the population or to the best element of each generation. Unfortunately, for the present work, there was no available time to test this hypothesis further.

In conclusion, the BRKGA approach we present proves to be an effective and efficient method to solve the order picking problem, as well as other routing problems, since it is capable of finding high quality solutions with small computational effort.

[1] *www.cpubenchmark.net*

Consequently, it seems clear that its application to the order picking process of online grocers would lead to significant time savings and greatly reduce the cost of order picking for these businesses.

Chapter 8: Conclusions

In this chapter we present a summary of the developed work and highlight its main contributions. Then, we expose the limitations faced during the development of the work and close with an outlook of potential developments.

8.1. Summary

In this work, we developed a *biased random-key genetic algorithm* (BRKGA) approach to the problem of in-store order picking routing, a problem of particular importance for the online grocery industry.

As concluded in Chapter 2, online grocery shopping is a growing market, with very positive outlook. Moreover, any grocer wishing to become an important player in the future must arrive early to this market and gain traction before its competition. However, the history of this business has made it clear that there are significant obstacles to be considered. Assuring customer delivery and order picking means that these inefficient processes that are performed by customers in the offline business will now be provided by the grocer. Thus, logistics play a decisive role in this industry, and failing to find efficient solutions will prevent any chance of achieving a profitable operation.

Although home delivery cannot be ruled out, customer pickup at the store may be an interesting delivery solution for many customers and for the online grocers as well. On the other hand, as concluded in Chapter 3, if we consider the most prevailing *bricks-and-clicks* business model, we will have to perform order picking in-store. Though necessary, this is a highly inefficient process, since the store's layout and product assignment are not designed for efficient order picking, but for maximizing offline shoppers' spending instead.

Chapters 5 and 6 demonstrate the process leading to the design of a methodological approach that would prove to be effective in finding efficient routes for in-store order picking. We developed a biased random key genetic algorithm (BRKGA) which is capable of finding good results against benchmark solutions. This genetic algorithm evolves a population of solutions for a number of generations, with a bias towards the fittest individuals in each generation. This allows us to reach good quality solutions in a short time. With such a model in use, bricks-and-clicks grocers would be capable of reducing the time and resources it takes to fulfill customers' orders from the stores, being one step closer to a profitable and sustainable business model.

8.2. Revisiting the research questions

In the introductory chapter of this thesis, we stated three research questions that guided our work:

Question 1: *Can a Genetic Algorithm approach help improve in-store order picking efficiency?*

Question 2: *Are Biased Random-key Genetic Algorithms competent for solving order picking routing problems, when compared to benchmark solutions?*

Question 3: *Will the adoption of an efficient in-store order picking system result in effective financial gains for online grocers?*

As mentioned, the early versions of Genetic Algorithms were ineffective for order-based problems such as TSP and VRP, which led to these being overlooked in favor of other heuristics (Schmitt, 1994; Toth and Vigo, 2002). However, the development of new GAs consistently improved on the results of the initial approaches. Our work enhances the understanding that, just as other competing heuristics, GAs should be regarded as viable options for order-based problems, especially if combined with heuristics that are more effective in exploiting promising regions of the search space. We present a mathematical model for one such order-based problem, the in-store order picking routing problem, and then successfully develop a Biased Random-key Genetic Algorithm approach for this problem. On this matter, and regarding *Question 1*, we conclude that Genetic Algorithms can be effective and, thus, help improve the efficiency of in-store order picking.

In addition, it was also important to understand how well our approach (Biased Random-key Genetic Algorithms) would perform against benchmark solutions of order picking routing problems, hence *Question 2*. The results showed that the BRKGA is capable of improving on benchmark solutions, allowing us to conclude that this approach is indeed both efficient and effective for order picking routing problems.

Lacking specific in-store order picking data, we are unable to precisely quantify the financial impact of an efficient order picking process. However, several evidences described throughout this work give clear indication that the benefits will be substantial:

i. Customer delivery and in-store order picking are the most significant logistic costs for an online grocer;
ii. Order picking is a labor-intensive process (in a typical warehouse, it accounts for up to 60% of all labor activities);
iii. Travel time stands for at least 50% of the order-picking processing time;

iv. Unlike picking in a warehouse, a number of issues that influence picking efficiency (store layout, storage assignment, depot location, and others) may not be optimized in a store, thus, only decisions concerning order picking can be made;

v. Efficient in-store order picking solutions will allow the picking to be performed three times as fast;

vi. Picking and handling can absorb 8.6% of an online grocer's online sales (Galante et al., 2013), in a market known for its low margins.

Taking into account these findings in order to answer to *Question 3*, we conclude that implementing an efficient in-store order picking system, based on a model such as the one developed in this work, will in fact result in effective financial gains for online grocers, improving their economic viability.

8.3. Main contributions

This work attempts to contribute to the study of online grocers' logistic challenges, both in terms of scientific expansion and economic viability for the industry.

Online grocery is a business that still faces a number of important problems. Some relate to marketing, like the conflict between the customers' added convenience and the greater cost, or their preference for choosing fresh products themselves. Other problems involve logistics, in which customer delivery (*the last mile*) and order picking are among the primary concerns.

This work demonstrates how in-store order picking for online grocers can be modeled and solved through an efficient algorithm, one that arrives at good quality solutions in little time. This can be a step towards lowering logistic costs, therefore reducing barriers to a wider adoption of the online grocery business model. Considering the operations of an existing traditional (*bricks-and-mortar*) grocer, an efficient in-store order picking system would not significantly increase its costs. Hence, combining efficient order picking with the adoption of a *click-and-collect* service (instead of home delivery, for instance), would allow the company to enter the online business with low upfront investment and low operating costs.

By applying developments of logistics and heuristics to a problem that bares some similarities to other order picking problems, but has a particular setting, this work attempts to further the scientific knowledge. Considering the specific process of in-store order picking for the online grocery industry, no previous work was found on this subject, although its economic relevance seems evident, as illustrated in Chapter 2.

Additionally, our work applied the *biased random-key genetic algorithm* (BRKGA) framework provided by Gonçalves and Resende (2011) to the order picking routing problem, with promising results. Thus, this is an extension of the efforts to develop the adoption of genetic algorithms for routing problems, namely the CVRP.

The results obtained proved that our BRKGA is an effective and efficient heuristic. Even though enhancements can be made to our approach, the solutions provided by our model already resulted in improvements to some of the results achieved by Chabot et al. (2015) for the same benchmark instances.

8.4. Limitations and suggestions for future research

Proceeding from the little attention in-store order picking has received in the existing literature, one of the main challenges for this work was the lack of specific information about the industry and its processes. In particular, no data was found related specifically to in-store order picking. Also, no data was supplied by the industry. Consequently, our work was developed for a set of instances with similar characteristics, but which may not be fully compatible with the online grocers' reality. One obvious extension of this work is to gather real online grocery shopping data, analyze it (particularly in the way it compares to the instances used for this work and other order picking data), and then apply this BRKGA approach to it.

At the same time, this could overcome a few limitations of the instances used in this work. With real data, information on composition of orders and their due dates could be considered. This means that the decisions regarding *scheduling*, *order batching*, and *order accumulation* could be incorporated in the model. Thus, all tasks concerning order picking could be considered conjointly and further conclusions about the picking strategy and its global efficiency would be drawn. Gaining access to real-world data, it would also be possible to quantify the financial gains that derive from efficient in-store order picking. Conducting a number of case studies to determine the practical efficiency gains that will be provided by a BRKGA model against the methods currently employed by online grocers to pick in-store would be an interesting line of research. These studies could quantify the cost savings for online grocers and the effective expenses that any grocer will have to bear in order to enter the online business, enabling a serious cost-benefit analysis and informed decision-making.

With regard to our BRKGA model, and as detailed in Chapter 7, the results from this approach are positive and very encouraging. However, some developments could be tested, in order to improve the algorithm's results and effectiveness. The most obvious of these developments would be to add a local improvement method to the BRKGA, something akin to what has been tried in other cases with positive outcomes (Baker and Ayechew, 2003; Buriol et al., 2005; Snyder and Daskin, 2006).

When time is taken into account, we can no longer consider an unlimited number of pickers (or routes). For any given period of time, only a certain number of routes can be executed. Also, if there are due dates associated to the picking items and a limited number of pickers, the resulting routes may differ. In addition, grocery orders will include fresh and frozen products, which may not be collected long in advance. Therefore, these items will have to be picked within a time frame, and this problem may be modeled as a VRP with time-windows.

Another important development would be to model this problem as a time dependent vehicle routing problem (TDVRP), since grocery stores are particularly sensitive to rush hours. Thus, although distances between locations remain constant, the picker's walking speed, and maybe even the picking time, will be aggravated, due to store congestion. Studying the impact of this on pickers' performance and picking itineraries may trigger new scheduling policies and marketing strategies.

Other VRP variants could also be studied to infer the savings obtained by looking at the picking problem in different ways. Considering that store replenishment is a necessary process for any offline grocery store, VRP with backhauls (VRPB) and VRP with simultaneous pickup and delivery (VRPSPD) could be interesting options to reconcile store replenishment with order picking. Thus, in-store order picking would not add too much time and resources to the already existing processes of a grocery store.

These suggestions for future research which conclude this work make two things very clear. On the one hand, the absence of specific online grocery logistics data is an obstacle to the development of scientific studies that could address specific issues of this industry, and help these companies find adequate and profitable business models. On the other, since relatively little attention has been given to the logistic issues of online grocers, as opposed to their marketing issues, there is still much room for improvement. This also means that the online grocery industry will be able to rapidly benefit from the developments made on logistics for other businesses in the last decades, thus skipping a few steps and quickly reaching its profitability tipping point.

References

Baker, B.M. and Ayechew, M.A. (2003), “A genetic algorithm for the vehicle routing problem”, *Computers & Operations Research*, Vol. 30, N. 5, pp.787-800.

Bakos, J. Y. (1991), “A strategic analysis of electronic marketplaces”, *MIS quarterly*, Vol. 15, No 3, pp. 295-310.

Bartholdi III, J.J., and Hackman, S.T. (2014), “Warehouse & distribution science: release 0.96”, Warehouse Science: The Supply Chain and Logistics Institute, School of Industrial and Systems Engineering, Atlanta: Georgia Institute of Technology.

Bean, J.C. (1994), “Genetics and random keys for sequencing and optimization”, ORSA Journal on Computing, Vol. 6, N. 2, pp. 154-180.

Bell, R. and Cuthbertson, R. (2004), “Collaboration in the retail supply chain”, in *Retail Strategy: the view from the bridge*, Reynolds, J., Cuthbertson, C. (editors), Oxford, pp. 52-77, ISBN 9780750656962

Bevan, J. and Murphy, R. (2001), “The nature of value created by UK online grocery retailers”, *International Journal of Consumer Studies*, Vol. 25, No 4, pp. 279-289.

Boyer, K. K., and Hult, G. T. M. (2006), “Customer behavioral intentions for online purchases: An examination of fulfillment method and customer experience level”, *Journal of Operations Management*, Vol. 24, No 2, pp.124-147.

Boyer, K.K. and Hult, G.T.M. (2005), “Extending the supply chain: Integrating operations and marketing in the online grocery industry”. *Journal of Operations Management.* Vol. 23, N. 6, pp. 642-661.

Breuer, P., Forina, L., Moulton, J. (2013), “Beyond the hype: Capturing value from big data and advanced analytics”, *Perspectives on retail and consumer goods*, Vol. 1, N. 1, pp. 4-9.

Buriol, L.S., Resende, M.G.C., Ribeiro, C.C. and Thorup, M. (2005), “A Hybrid Genetic Algorithm for the Weight Setting Problem in OSPF/IS-IS Routing”, *Journal of Combinatorial Optimization*, Vol. 46, N. 1, pp. 36-56.

Busetti, F. (2007), Genetic algorithms overview.

Cattani, K., Perdikaki, O., and Marucheck, A. (2007), “The Perishability of Online Grocers”, *Decision Sciences*, Vol. 38, No 2, pp. 329-355.

Chabot, T., Coelho, L.C., Renaud, J. and Côté, J-F. (2015), “Mathematical models, heuristic and exact method for order picking in 3d-narrow aisles”, Working document, Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT).

Clarke, G. and Wright, J.W. (1964), “Scheduling of vehicles from a central depot to a number of delivery points”, *Operations Research*, Vol. 12, N. 4, pp. 568-581.

Coopey, R., O'Connell, S., & Porter, D. (1999), “Mail order in the United Kingdom c. 1880-1960: how mail order competed with other forms of retailing”, *The International Review of Retail, Distribution and Consumer Research*, Vol. 9, No 3, pp. 261-273.

Dantzig, G., Fulkerson, R. and Johnson, S. (1954), “Solution of a large-scale traveling-salesman problem”, *Journal of the Operations Research Society of America*, Vol. 2, N. 4, pp. 393-410.

Dantzig, G.B. and Ramser, J.H. (1959), “The truck dispatching Problem”, *Management Science*, Vol. 6, N. 1, pp.80-91.

Darwin, C. (1859). *On the origins of species by means of natural selection*, London: Murray, 247.

De Koster, R. and Van der Poort, E. (1998), “Routing orderpickers in a warehouse: a comparison between optimal and heuristic solutions.”, *IIE Transactions*, Vol.30, N. 5, pp.469-480.

De Koster, R., Le-Duc, T. and Roodbergen, K.J. (2007), “Design and control of warehouse order picking: a literature review”, *European Journal of Operational Research*, Vol. 182, N. 2, pp.481-501.

De Koster, R., Van der Poort, E. and Wolters, M. (1999), “Efficient orderbatching methods in warehouses”, *International Journal of Production Research*, Vol. 37, N. 7, pp. 1479-1504.

De Vuist, C., Kesteloo, M., and Hoogenberg, M. (2014), “Click and cash: How to profit in online grocery shopping” in *Strategy&*, PwC.

Delaney-Klinger, K., Boyer, K. K., and Frohlich, M. (2003), “The return of online grocery shopping: a comparative analysis of Webvan and Tesco's operational methods”, *The TQM Magazine*, Vol. 15, No 3, pp. 187-196.

Desrosiers, J., Dumas, Y., and Soumis, F. (1986), "A dynamic programming solution of the large-scale single-vehicle dial-a-ride problem with time windows", *American Journal of Mathematical and Management Sciences*, Vol. *6, N.* 3-4, pp. 301-325.

Drury, J. (1988), "Towards more efficient order picking", *IMM monograph N. 1*, Report, The Institute of Materials Management, Cranfield, U.K.

Eksioglu, B., Vural, A.V. and Reisman, A. (2009), "The vehicle routing problem: A taxonomic review", *Computers & Industrial Engineering*, Vol. 57, N. 4, pp. 1472-1483.

Fontes, D.B.M.M. and Gonçalves, J.F. (2013), "A multi-population hybrid biased random key genetic algorithm for hop-constrained trees in nonlinear cost flow networks", *Optimization Letters,* Vol. 7, N. 6, pp. 1303-1324.

Fontes, D.B.M.M. and Gonçalves, J.F. (2015), "A Genetic Algorithm for Scheduling Alternative Tasks Subject to Technical Failure.", in *Optimization, Control, and Applications in the Information Age*, Migdalas, A. and Karakitsiou, A. (editors), Springer Proceedures in Mathematics and Statistics, pp. 139-152. Springer International Publishing.

Galante, N., López, E.G., Monroe, S. (2013), "The future of online grocery in Europe", *Perspectives on retail and consumer goods,* Vol. 1, N. 1, pp. 22-31.

Goetschalckx, M., and Jacobs-Blecha, C. (1989), "The vehicle routing problem with backhauls", *European Journal of Operational Research*, Vol. *42, N.* 1, pp. 39-51.

Golberg, D.E. (1989), *Genetic algorithms in search, optimization, and machine learning,* Boston: Addison-Wesley Longman Publishing Co., Inc.

Gonçalves, J.F. and Resende, M.G.C. (2011), "Biased random-key genetic algorithms for combinatorial optimization*", Journal of Heuristics*, Vol. 17, N. 5, pp. 487-525.

Gonçalves, J.F. and Resende, M.G.C. (2013), "A biased random key genetic algorithm for 2D and 3D bin packing problems*", International Journal of Production Economics, Vol. 145, N. 1, pp.500-510.*

Gonçalves, J.F., Costa, M.D., Resende, M.G.C. (2014), "A biased random-key genetic algorithm for the minimization of open stacks problem", International Transactions in Operational Research, Vol. 1, N. 22.

Goodman, J. (2008), *Who Does the Grocery Shopping, and When Do They Do It?*, The Time Use Institute.

Gray, A.E., Karmarkar, U.S. and Seidmann, A. (1992), "Design and operation of an order consolidation warehouse: Models and application", *European Journal of Operational Research*, Vol. 58, pp. 14-36.

Gue, K.R., Meller, R.D. and Skufca, J.D. (2006), "The effects of pick density on order picking areas with narrow aisles", *IIE Transactions*, Vol. 38, N.10, pp. 859-868.

Hays, T., Keskinocak, P., and De López, V.M. (2005), "Strategies and challenges of internet grocery retailing logistics" in *Applications of Supply Chain Management and E-Commerce Research*, Geunes, J., Akçali, E., Pardalos, P.M., Romejin, H.E. and Shen, Z.J.M. (editors), Vol. 92, N. 2, pp. 217-252, Springer US. ISBN: 978-0-387-23391-8.

Henn, S. and Schmid, V. (2013), "Metaheuristics for order batching and sequencing in manual order picking systems", *Computers & Industrial Engineering*, Vol. 66, N. 2, pp.338-351.

Holland, J. H. (1975), *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: University of Michigan Press.

Jaegere, N., Defraeye, M. and Van Nieuwenhuyse I. (2014), "The vehicle routing problem: state of the art classification and review", Faculty of Economics and Business Research Report, Faculty of Economics and Business Leuven, Belgium.

Janson, M. A., Cecez-Kecmanovic, D., and Mäkeläinen, V. P. (2007), "E-Grocer Strategies: A Case Study", in *ECIS*, pp. 58-69.

Kantar Worldpanel (2015), A*ccelerating the growth of e-commerce: 2015 Edition.*

Larrañaga, P., Kuijpers, C.M.H., Murga, R.H., Inza, I. and Dizdarevic, S. (1999), "Genetic algorithms for the travelling salesman problem: a review of representations and operators", *Artificial Intelligence Review*, Vol. 13, N. 2, pp. 129-170.

Larson, J.S., Bradlow, E.T., Fader, P.S. (2005), "An exploratory look at supermarket shopping paths", International Journal of Research in Marketing, Vol. 22, N. 4, pp. 395-414.

López, E.G., Said, R., Westphely, K. (2014), "How to win in online grocery: Advice from a pioneer", *Retail Practice, Consumer Packaged Goods Practice,* McKinsey & Company, Retrieved September 17, 2015, from http://www.mckinsey.com/insights/consumer_and_retail/how_to_win_in_online_grocery_advice_from_a_pioneer.

Malandraki, C. and Daskin, M.S. (1992), "Time dependent vehicle routing problems: formulations, properties and heuristic algorithms", *Transportation Science*, Vol. 26, N. 3, pp.185-200.

Michalewicz, Z. (1995*), Genetic Algorithms + Data Structures = Evolution Programs*, 3rd Edition, New York: Springer-Verlag.

Mitchel, M. (1995), "Genetic Algorithms: an overview", *Complexity*, Vol. 1, N. 1, pp. 31-39.

Morganosky, M. A. and Cude, B. J. (2000), "Consumer response to online grocery shopping", *International Journal of Retail & Distribution Management*, Vol. 28, No 1, pp.17-26.

Müller-Lankenau, C., Wehmeyer, K., and Klein, S. (2005), "Multi-channel strategies: capturing and exploring diversity in the European retail grocery industry", *International Journal of Electronic Commerce*, Vol. 10, No 2, pp. 85-122.

Murphy, A. J. (2003), "(Re) solving space and time: fulfilment issues in online grocery retailing", *Environment and Planning A*, Vol. 35, No 7, pp. 1173-1200.

Nikolakopoulos, A. (2013), "Army rapid fielding by optimizing order picking routes in warehouses with parallel aisles – Implementation in a real case study", *Journal of Computations & Modelling*, Vol.3, N. 4, pp. 137-163.

Perea y Monsuwé, T., Dellaert, B. G., and De Ruyter, K. (2004), "What drives consumers to shop online? A literature review", *International journal of service industry management*, Vol. 15, No 1, pp. 102-121.

Petersen, C.G. (1997), "An evaluation of order picking routeing policies", International Journal of Operations & Production Management, Vol. 17, N. 11, pp. 1098-1111.

Petersen, C.G. and Aase, G. (2004), "A comparison of picking, storage, and routing policies in manual order picking", *International Journal of Production Economics*, Vol. 92, N. 1, pp.11-19.

Ratliff, H.D. and Rosenthal, A.S. (1983), "Order-Picking in a Rectangular Warehouse: A Solvable Case of the Traveling Salesman Problem", *Operations Research.* Vol. 31, N. 3, pp.507-521.

Reeves, C. (2003), “Genetic algorithms”, in *International Series in Operations Research and Management Science*, Glover, F. and Kochenberger, G.A. (editors), Handbook of Metaheuristics, Vol. 57, pp. 55-82, Springer US.

Reid, D.J. (2000), Feasibility and Genetic Algorithms: the behavior of crossover and mutation, Salisbury, South Australia: DSTO Electronics and Surveillance Research Laboratory.

Roodbergen, K.J. and De Koster, R. (2001), “Routing order pickers in a warehouse with a middle aisle”, *European Journal of Operational Research*, Vol. 133, N. 1, pp.32-43.

Schmitt, L.J. (1994), “An empirical computational study of Genetic algorithms to solve order based problems: An emphasis on TSP and VRPTC”, Ph.D. Dissertation, Fogelman College of Business and Economics, University of Memphis.

Smith, C. (2014), “How E-Commerce Is Finally Disrupting the $600 Billion-a-Year Grocery Industry”, Business Insider, Retrieved September 21, 2015, from http://www.businessinsider.com/e-commerce-disrupting-huge-grocery-market-2014-12

Snyder, L.V. and Daskin, M.S. (2006), “A random-key genetic algorithm for the generalized traveling salesman problem”, *European Journal of Operational Research*, Vol. 174, N. 1, pp. 38-53.

Spears, W.M. and Dejong, K.A., (1991), On the virtues of parameterized uniform crossover, in: Proceedings of the Fourth International Conference on genetic Algorithms, pp.230-236.

Tansan, A.S. and Gen, M. (2012), “A genetic algorithm based approach to vehicle routing problem with simultaneous pick-up and deliveries”, *Computers & Industrial Engineering*, Vol. 62, pp. 755-761.

Theys, C., Dullaert, W., Braysy, O. and Raa, B. (2007), "Towards a metaheuristic for routing order pickers in a warehouse", in *Evolutionary methods for design, optimization and control*, Neittaanmaki, P., Periaux, J. and Tuovinen, T. (editors), pp. 385-390, Barcelona: CIM NZ.

Toth, P. and Vigo, D. (2002), *The Vehicle routing Problem*, Bologne: Society for Industrial and Applied Mathematics.

Vanelslander, T., Deketele, L. and Van Hove, D. (2013), "Commonly used e-commerce supply chains for fast moving consumer goods: comparison and suggestions for improvement", *International Journal of Logistics: Research and Applications*, Vol. 16, N. 3, pp. 243–256.

Yrjola, H. (2001), "Physical distribution considerations for electronic grocery shopping", *International Journal of Physical Distribution & Logistics Management*, Vol. 31, No 10, pp. 746-761.

Yrjola, H., and Tanskanen, K. (2005). "8. Effective E-Grocery Logistics", in Kornum, N. and Bjerre, M. (editors), *Grocery E-commerce: Consumer Behaviour and Business Strategies*, pp. 160-182, Cheltenham: Edward Elgar Publishing Limited.

Zhang, H. and Liu, B. (2009), "A New Genetic Algorithm for Order-Picking of Irregular Warehouse". *International Conference on Environmental Science and Information Application Technology*, Vol. 1, pp. 121-124. ISBN: 978-0-7695-3682-8.