Public Transport-based Crowd-shipping with Backup Transfers

by

Kerim Uygur Kızıl

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Industrial Engineering

KOÇ ÜNİVERSİTESİ

June, 2021

Public Transport-based Crowd-shipping with Backup Transfers

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Kerim Uygur Kızıl

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Assist. Prof. Barış Yıldız

Prof. Fikri Karaesmen

Assist. Prof. Vedat Bayram

Date:

To my dear sister Rana, who is destined for great things.

ABSTRACT

With the rising urbanization and booming e-commerce, traditional last-mile delivery systems fail to keep up with the exploding demand and to satisfy the need for faster, cheaper, and more environmentally friendly deliveries. Several new approaches are put forward as an alternative to classical delivery systems in this regard, yet none of them offers the same level of flexibility, capacity, reliability, and managerial control by itself. This thesis proposes a new last-mile delivery model that combines several new approaches and technologies to address this issue. More precisely, we suggest using public transit as a backbone network completed by automated service points, crowd-shipping, and backup transfers with zero-emission vehicles to provide a low-cost and environmentally friendly express delivery service. Offering cheap, reliable, and fast transfers between any two locations in an urban area, such a delivery network can provide unique opportunities for small and medium-sized businesses to participate in the e-commerce ecosystem and help large e-tailers to lower their costs and improve service quality. The design problem for the envisioned system is formulated as a two-stage stochastic program, and a branch-and-price (BP) algorithm is devised to solve it. Taking advantage of the nearly decomposable structure that would emerge in possible real-world applications, our study presents the first example of using decomposition branching in a BP framework to enhance computational efficiency. Extensive computational studies on realistic problem instances reveal valuable managerial insights for the proposed system and attest to the efficacy of the suggested methodology.

ÖZETÇE

Artan kentleşme ve e-ticaret hacmiyle birlikte, geleneksel son kilometre teslimat sistemleri, büyüyen talebe ayak uydurmakta ve daha hızlı, daha ucuz ve daha çevre dostu teslimat ihtiyacını karşılamakta yetersiz kalmaktadırlar. Bu bağlamda klasik dağıtım sistemlerine alternatif olarak birkaç yeni yaklaşım öne sürülmüştür, ancak bunların hiçbiri tek başlarına klasik sistemler ile aynı düzeyde esneklik, kapasite, güvenilirlik ve yönetimsel kontrol sunamamaktadırlar. Bu tez, bu sorunu ele almak için birkaç yeni yaklaşımı ve alandaki son teknolojileri birleştiren bir son kilometre teslim modeli önermektedir. Önerilen model, toplu taşımanın düşük maliyetli ve çevre dostu teslimat hizmetleri sağlamak amacıyla şehir içinde bir omurga sistem olarak kullanılmasını ve bu sistemin otomatik dolaplar barındıran teslimat noktaları, kitle kaynaklı taşımacılık ve sıfır emisyonlu araçlar kullanılarak yapılan yedek transferler ile desteklenmesini öngörmektedir. Bir şehrin herhangi iki noktası arasında ucuz, güvenilir ve hızlı transferler sunan böyle bir dağıtım ağı, küçük ve orta ölçekli işletmelerin e-ticaret ekosistemine katılması için benzersiz fırsatlar sağlamasının yanında büyük e-perakendecilere maliyetleri düşürme ve hizmet kalitelerini artırma konularında yardımcı olacaktır. Öngörülen sistemle bağdaşan tasarım problemi bu tezde iki aşamalı bir stokastik program olarak formüle edilmiş ve problemin çözümü için bir dal-fiyat algoritması geliştirilmiştir. Problemin gerçek hayat uygulamalarında karşılaşılabilecek "neredeyse ayrıştırılabilir" yapıdan hareketle, bu çalışma, çözüm verimliliğini artırma amacıyla dal-fiyat algoritması içinde ayrıştırma dallanması yönteminin kullanımının ilk örneğini sunmaktadır. Gerçekçi problem örnekleri üzerine yapılan kapsamlı sayısal deneyler, önerilen sistem için anlamlı yönetimsel içgörüler ortaya koymakta ve geliştirilen metodolojinin etkinliğini doğrulamaktadır.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and appreciation to the following people for making this thesis possible.

First of all, I would like to thank my supervisor Dr. Barış Yıldız for guiding and supporting me in every step of my master's studies. I truly appreciate not only his immense scientific knowledge and expertise but also his commendable patience and optimism. This thesis would not have been possible without his invaluable encouragement, advice, and guidance.

I also would like to thank my thesis committee members Dr. Fikri Karaesmen and Dr. Vedat Bayram, for reviewing my work and offering their valuable feedback.

I wish to acknowledge the continuous support of my family. They have always been there for me when I needed them. I especially thank them for being genuinely interested in my work and for facilitating a peaceful home-office environment in a time that most of the world had no choice but to work remotely. I especially thank my little sister Rana, whom I dedicate this thesis, for being my best friend at home.

I want to mention many great friends I feel very lucky to have, especially for their support during my master's studies. Very special thanks go to Çelen Naz and Buse for always accompanying me through my studies. I have been truly motivated by their diligence, their work ethic, and of course, their invaluable friendship. I would like to thank Umut and Baturhan for the great times we had on the campus, whether it was spent studying together or just hanging out. I also acknowledge the great friendships I have with Bengi, Yasin, Karim, Gökçe, and Şule.

Finally, I would like to acknowledge the Scientific and Technological Research Council of Turkey (TUBITAK) for financially supporting this thesis and the related project (grant number 218M605).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ASP	Automated Service Point
BC	Base Case
BP	Branch-and-Price Algorithm
BRT	Bus Rapid Transit
BT	Backup Transfer
CG	Column Generation
CS	Crowd-shipping
DB	Decomposition Branching
DC	Distribution Cycle
IP	Integer Program
MIP	Mixed Integer Program
PCBP	Public Transport Based Crowd-shipping Problem with Backup Transfers
PP	Pricing Problem
PT	Public Transport
PTV	Public Transport Vehicle
RBPTS	Railway-based Public Transport Services
SA	Exhaustive Search Algorithm
SLDFC	Scheduled Lines with Dedicated Freight Capacities
SLV	Scheduled Line Vehicle
SR	Service Region
TRY	Turkish Liras — The official currency of Turkey

Chapter 1

INTRODUCTION

As disruptions brought by the information revolution are reshaping the commercial landscape, e-commerce establishes itself as the new normal in the retail business. For e-tailers vying for a competitive advantage in this world of new opportunities and challenges, the name of the game is last-mile logistics, which is by far the most expensive, time-consuming, and operationally challenging part of the whole delivery process (Vanelslander et al., 2013). It is not only the businesses that are looking for innovations in last-mile logistics. The rising urbanization and exploding e-commerce add tremendous pressure on city logistics systems that are already stretched to their limits or beyond. This, in turn, leads to several negative externalities, such as increased air pollution, noise, accidents, and congestion (Savelsbergh and Van Woensel, 2016). Alleviating these externalities of urban transportation to the extent possible whilst retaining the positive economic impact it has on the economy is high on the agenda of city residents and governments worldwide.

Besides failing to keep up with the rising expectations for faster, cheaper, and more environmentally friendly deliveries, classical last-mile delivery systems cannot answer the needs of small and medium-sized businesses to become a part of the e-commerce ecosystem. Not having enough resources to build their own delivery capacity and the lack of cheap and fast delivery options raise a significant barrier for those businesses to offer their products and services through e-commerce channels. Rapidly growing meal and grocery delivery start-ups such as Grubhub, UberEats, Instacart, and Glovo and the lack of such examples in other retail businesses indicate the critical importance and the urgent need for the low price many-to-many express delivery services.

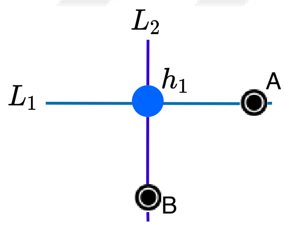
Several new approaches in last-mile delivery have recently been tried and put forward by the industry and researchers with these strong motivations. Most of those approaches essentially share a common theme of reducing the need and usage of company-owned delivery fleets that are typically composed of combustion engine delivery trucks or vans. One can achieve this by assigning some proportion of the demand entirely or partially to more environmentally friendly and/or cheaper transport alternatives. For example, in crowd-shipping (CS), it is the crowd that provides such an alternative delivery capacity (Punel and Stathopoulos, 2017). Models that put forward the idea of using public transport (PT) as an alternative envision completing some legs of package journey by PT vehicles instead of company-owned delivery fleets (Ghilas et al., 2016, Mourad et al., 2019). Similarly, the deployment of automated service points (ASPs) at strategic locations is another novel idea aiming to reduce the distance traveled by traditional trucks or vans by letting the end customer complete the last leg of the journey (Iwan et al., 2016, van Duin et al., 2020). Taking advantage of the recent advances in new vehicle technologies by replacing the combustion engine delivery trucks with less costly (environmentally and economically) alternatives (i.e., electric vehicles, autonomous delivery robots/drones, e-bikes, or e-motorcycles) is also considered to achieve virtually zero-emission logistics (Jaller et al., 2018, Murray and Chu, 2015, Quak et al., 2016).

Despite the tremendous effort to realize and develop these novel ideas, classical delivery practices still dominate the last-mile delivery landscape, especially in express small package deliveries. That is mainly because, despite their distinctive advantages, none of the aforementioned approaches by itself can offer the same level of flexibility, capacity, reliability, and managerial control as company-owned fleets provide. CS uses the already-existing trips of the public to meet small package transfer demands with less environmental and economic costs. On the other hand, CS also introduces uncertainties and a lack of control in the delivery capacity, as the number of crowd-shippers willing to participate in deliveries and their offer acceptance behavior cannot be fully controlled by the network management (Yildiz and Savelsbergh, 2019). Using PT vehicles for some part of the journey can eliminate the use of trucks in some main roads of the city, yet it cannot fully support door-to-door deliveries completely as main PT lines can be rather far from the origin and destination points of packages. For example, if railway lines are to be chosen as the main PT lines of such a system, the walking distance between the nearest station and an individual's home can be as long as 3.5 kilometers, (Hasiak, 2019), much larger than the mean distance of 670 meters people usually walk for active transportation purposes (Millward et al., 2013). Furthermore, too many loading-unloading stops for PT vehicles can lengthen the duration of trips, which may negatively affect the satisfaction of the PT riders (Ghilas et al., 2013). Similarly, while the deployment of ASPs helps alleviate challenges of home delivery services by enabling unattended deliveries and offering better consolidation and bundling (Cardenas et al., 2017), it does not obviate the need for a delivery fleet by itself. New vehicle technologies (i.e., drones, delivery robots, e-bikes, and e-motorcycles) have the potential of completing transfers with virtually no emissions, yet they have a limited capacity, range, or speed to replace the delivery trucks/vans fully. Building on these observations, this thesis proposes a new last-mile delivery model in which several new delivery models and technologies are utilized together to make up for each other's shortcomings and boost their advantages to offer cheap and environmentally friendly express small package transfers between any two points in a given urban area. In the envisioned system, the transfer demands are fulfilled by the collaborative efforts of PT vehicles, crowd, ASPs, and the back-up delivery capacity offered by near-zero emission vehicles.

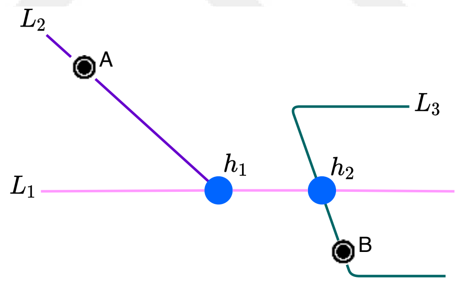
The envisioned system specifically works according to the following principles:

- The main objective of the system is to fulfill many-to-many express delivery demands (i.e., two or three-hour delivery requests) within a city in a cost-efficient way. For that purpose, ASPs are installed in selected PT stations and several other locations in the city. ASPs in PT stations are referred to as *PT connections* since they primarily serve as gateways for the packages to enter/exit the PT network and enable transshipments within it. PT connections are equipped with the necessary machinery and personnel to seamlessly process the packages (i.e., to collect, load, unload, and temporarily store). The remaining ASPs function as origin and destination points for transfer demands and are called *satellites* of the system.

- Transfers between PT connections are completed with PT vehicles. To establish all pairs connectivity between the satellites, or in other words, to serve every possible origin-destination pair for package transfers, PT connections are required to be interconnected. This implies that at least one interchange station must be designated as a PT connection to serve as a transshipment point to join any two PT lines in the system. Illustrating the necessity of such transshipment points, we present two small examples in Figure 1.1. In Figure 1.1a, we see that without h_1 chosen as a PT connection (to serve as a transshipment point), PT connections A and B cannot be linked with PT lines L_1 and L_2. Similarly, in Figure 1.1b, h_1 and h_2 both must be opened as PT connections to link A and B using lines L_1, L_2 and L_3.



(a)



(b)

Figure 1.1: Examples for connectivity requirements.

- Crowd-shippers, i.e., travelers who are willing to insert package transfers into their already planned trips to earn a small compensation or to do good for the society/environment, take part in package transfers between the satellites and the PT connections. Note that we do not consider using CS for long-distance point-to-point transfers different from many existing CS approaches. In the envisioned system, crowd-shippers are only responsible for a small part of the whole delivery process, requiring transferring packages over relatively short distances. By doing that, we aim to make it easier for the system to capture a large enough CS capacity without a need to offer large compensations, which is critical to provide a low-cost express delivery alternative for the system users.

- Pick-up and deliveries by crowd-shippers and the use of backup transfers (BT) are organized in periodic *distribution cycles* (DC) to ensure timely deliveries. Crowd-shippers who are willing to be assigned package(s) in a certain DC must register their trips before the beginning of the DC to enhance system efficiency (Yıldız, 2021). At the beginning of a DC, the management checks all active packages (waiting to be transferred to their destination satellite) and the registered CS trips and decides package assignments to the crowd-shippers and determines which satellites will be visited by BT vehicles. After the assignment decision is made, the BT tours are planned to visit the required satellites with the minimum total cost. If needed, packages are repositioned on the PT network so that all the incoming packages of the satellites that are to be visited by a BT tour are collected in its starting PT connection. Once the repositioning phase is over, BT tours start as planned and conclude within the given touring time limit. Here we consider repositioning operations assuming that the PT transfers cost much less than the use of BT vehicles. However, the system can also be operated without any repositioning before BT tours if one would like to relax this assumption. We also want to note that DC frequencies and the allowed tour times for the BT vehicles are critical to limit the excessive delays in delivery operations and ensure customer satisfaction. In this initial study, we assume that the network management (also referred to as the platform throughout the thesis) has already made these strategic decisions and focus on the network design problem that aims to minimize the expected operating costs of a DC.

- Backup transfers are conducted with zero-emission vehicles (i.e., electric vehicles) rented by the network management on demand. Although not a modeling necessity, we envision such a renting mechanism (instead of using company-owned vehicles), considering that in a well-designed network with a robust CS capacity, the use of backup transfers would be exceptions rather than the primary medium for package transfers between the satellites and the PT connections. BT vehicles start their tour from PT connections to pick up incoming packages for the satellite(s) they will visit and end their tour in a PT connection (possibly a different PT connection than they have started) to drop the outgoing packages they have collected from the visited ASPs.

The envisioned system offers a seamless connection between any two points in the covered areas of the city and allows anyone to be a sender (retailer) and/or a receiver (consumer), allowing small and medium-sized businesses to become a part of the e-commerce ecosystem and individuals to exchange goods with each other in line with the idea of *sharing economy*. We can describe the operational aspect of the system with a practical B2C example, depicted in Figure 1.2.

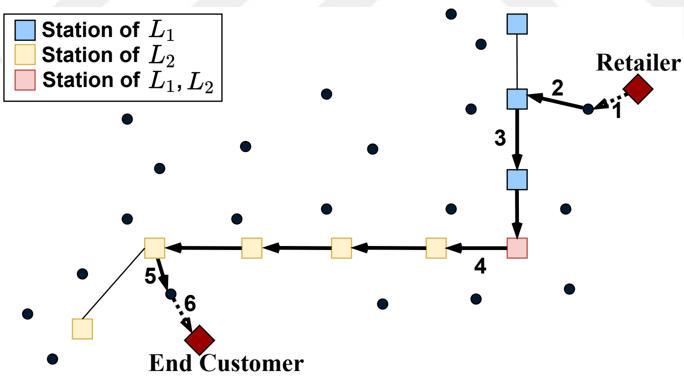


Figure 1.2: An illustrative journey.

Suppose an individual orders herself a new pair of headphones from an online retailer. Say the retailer is located in a busy area in the same city as the individual, yet is far away from the individual's home. More specifically, there is no one direct PT line between the locations of the individual and the retailer. The package's journey consists of five stops: origin satellite, origin PT connection, transshipment/transfer PT connection, destination PT connection, and destination satellite, respectively. The retailer puts the headphones in a close-by parcel locker (origin satellite) to be picked up by a crowd-shipper. Crowdshipper transfers the package to the origin PT connection, where the package enters the PT network. The transfer between origin and destination PT connections is carried out by PT vehicles (PTVs), for example, in a special compartment for buses or in a dedicated car attached at the end of passenger trains, with no help from crowd-shippers. Since one transfer is necessary between two PT lines, the package is temporarily stored in the transshipment PT connection then starts its journey in the final PT line to reach the destination PT connection. Finally, a crowd-shipper picks up the package from the destination PT connection and transfers it to the destination satellite, where the individual picks it up. The transfers between stops are shown in Figure 1.2 with arcs 1, 2, 3, 4, 5, and 6. Transfers 1 and 6 are out of the system's scope as they are completed by the retailer and end customer, respectively. Transfers 2 and 5 are carried out by a crowd-shipper or BTs, depending on the platform's decision about which satellites will be served by CS and which satellites will be served by BT vehicles in the respective distribution cycles. Note that, for the case of CS, we consider inserting package transfers into the already planned trips of the crowd and do not require dedicated trips. PTVs handle transfers 3 and 4. These transfers can be made seamlessly for a small cost since these trips are already scheduled and no additional costs will be incurred associated with the carrying of the packages unless the overall passenger/freight capacity of the PT lines is exceptionally limited. Observe that the package requires a transfer at an interchange station between lines L_1 and L_2.

This thesis introduces a strategic level design problem for the envisioned delivery network, namely Public Transport Based Crowd-shipping Problem with Back-up Transfers (PCBP). To model this problem, we build an integer program (IP) with a two-stage stochastic programming structure, where the locations of the PT connections are determined in the first stage, and the routing decisions for the BTs are considered in the second stage, after observing the actual transfer demand and CS delivery capacity. Our formulation uses a scenario-based approach to consider the uncertainties in the transfer demands and CS delivery capacity levels. Therefore, decisions regarding where to use BT versus CS are specific to each scenario, and the objective is to minimize the respective expected costs. We propose a branch-and-price (BP) algorithm to solve realistic instances of this problem. Furthermore, making use of the "nearly decomposable" structure of the suggested formulation, we use a new branching rule, namely, decomposition branching (DB) proposed by Yildiz et al. (2018). We then generate and solve realistic problem instances from the city of Istanbul to gain meaningful insights into such a system. More specifically, we investigate how the number of PT connections affects the number of satellites that CS can serve, how the system responds to different levels of CS capacity and transfer demands, the effects of the BT tour time limits on the system performance, and the algorithmic advantages of DB over plain BP.

Chapter 2

LITERATURE REVIEW

Externalities associated with the rapidly expanding demand in urban logistics systems motivate researchers in many different domains to search for more environmental-friendly and less costly delivery alternatives. Today's consumers are also remarkably sensitive to the environmental impact of the last mile delivery operations they cause by purchasing goods (Caspersen and Navrud, 2021). The increasing interest and recent developments in vehicles with almost zero emissions (e.g., electric motorbikes, drones) can be interpreted as one output of this search for alternative delivery options to reduce the use of traditional diesel trucks.

Apart from the ever-developing automotive technologies that aim to substitute diesel trucks, there are several more holistic approaches in logistics literature that intend to change the way packages are transferred within cities. Those holistic approaches generally aim to eliminate or significantly reduce the use of trucks altogether, whether they are diesel or not, by not using dedicated trips for freight whenever possible. CS, as introduced in the previous chapter, is one of the leading examples of such approaches. CS aims to use already existing private trips to transfer packages, ultimately eliminating any need for separate trips for freight, per the idea that "the least polluting trip is the one not performed." CS is generally accepted as an environmentally friendly delivery option by all stakeholders. For example, CS is shown to reduce carbon emissions up to 17.19% in a recent case study (Bi et al., 2020). Another innovative idea is the use of scheduled lines with dedicated freight capacities (SLDFC). SLDFC addresses the underused capacity in scheduled line vehicles (SLVs, which can be viewed as a generalization of PTVs) and suggests that the efficiency of public transport systems can be significantly improved if SLVs are to transfer freight along with passengers.

From an application perspective, our work builds upon the concepts of CS and

SLDFC. Referring back to Figure 1.2, transfers 5 and 6 relate to the CS concept, while transfers 3 and 4 are conducted by PTVs (i.e., SLVs) and are directly related to the idea of SLDFC. It is important to note here that both CS and SLDFC characterize the integration of passenger and freight transport in city logistics systems so that the transport demand can be met with less emission, congestion, and fewer vehicles involved in the processes.

In this thesis, we investigate the state of the art on CS and SLDFC in three main groups, each forming a section. The first group consists of the studies that focus on CS and SLDFC themselves as concepts, along with their possible variants and extensions, from a theoretical standpoint as enablers of sustainable, integrated logistics systems. These studies often discuss the possible advantages and drawbacks of such new systems, introduce current business models based on these systems, or position such systems as an integral part of environmental-friendly and low-cost alternatives in last-mile logistics. The second group involves empirical studies investigating the viability and applicability of potential or current integrated systems, analyzing the critical factors for their success, and proving some prescriptive conclusions for decision-makers. This group can be further divided into two subcategories for studies that primarily focus on CS systems and ones that mainly address SLDFC-based systems. Lastly, the third group deals with the decision-making problems regarding CS and SLDFC network design and management and proposes mathematical models along with solution approaches to address them. Our study can be classified within this last group.

The problem introduced in this thesis, namely, PCBP, can also be broadly classified as a hub network design problem. Consequently, as the fourth section of this chapter, we briefly discuss the similarities and differences between PCBP and a *standard* hub network problem. We conclude this chapter by listing our contributions to the existing literature. Subsequently, we list our contributions to the literature, outline the organization for the rest of the thesis and end this chapter.

2.1 CS and SLDFC as Concepts

Many papers that discuss the CS concept do it by examining several operations of logistics companies (many of them start-ups) and how these operations include public and/or occasional drivers (i.e., crowd). Rougè and Montreuil (2014) use the term *crowdsourced delivery* for CS and say that the industry focusing on the CS concept emerged in the early 2010s. Their study focuses on the business models of eighteen CS start-ups. The authors examine some advantages of CS systems with respect to customers, retailers, and society. High speed, more personalization, access to new products, and low prices are the points they discuss as benefits to CS customers. Similar to our discussion about the concept of sharing economy in the previous chapter, the authors suggest that the click-and-mortar stores can benefit greatly from a CS system to compete with e-tailers, as the companies will be able to use their physical stores as "mini distribution centers". As the gains for society, the authors list minimizing the environmental footprint and reducing the congestion in urban areas.

Some authors adopt a wider definition of logistics to investigate crowdsourced information logistics (IL) separately in their studies. For example, Mladenow et al. (2016) investigates the logistics innovations regarding crowdsourcing in two main categories. The first category focuses on innovations targeting classic logistics operations. The term CS, as used in this thesis, would fall into this category. The authors give examples of *Flexe*, *MyWays*, and *CheckRobin* as applications used for crowdsourcing for classic logistics. The second category regards IL. The example given for this category is Waze, a traffic and navigation app that allows users to submit information about possible traffic jams, accidents, and road closures. Using the user-submitted information, the app offers the best route to other users. We would like to note here that IL operations are inseparable parts of any CS system. That is, online platforms are used in any CS system to integrate the users' efforts and match the demand with supply (Ermagun et al., 2019). In this thesis, we assume that a platform is already deployed for the IL side of the system (e.g., sharing information about the locations of the packages and the assignment decisions) and focus on the transfer of packages exclusively.

CS is also often seen as an integral part of an *ideal city* that has minimal impact on the environment and enables a sustainable life for its dwellers. The study by Munoz and Cohen (2015) adopts this wider perspective and considers sharing cities as enablers of sustainable consumption and production. They list five sharing economy categories: energy, food, goods, mobility & transport, and space sharing. CS is given as a sub-component for the mobility & transport category along with ridesharing and carsharing. The authors underline that CS targets the excess capacity in vehicles for package transfers and leads to more sustainable consumption compared to traditional delivery alternatives.

The idea of integrating the journeys of freight and passengers is already implemented for long-distance transfers. For example, in addition to dedicated freight aircraft, the airlines use passenger flights as a means to transfer cargo (Feng et al., 2015). In late 2000s, the majority of cargo carried by the top eight Asian airlines were transferred in the belly section of passenger planes (Wong et al., 2009). There are similar examples for such integrations in long-distance rail and sea transport as well. On the other hand, last mile operations for passengers and packages usually are separate from each other, as Arvidsson et al. (2016) discuss in their paper. In this conceptual paper, the authors suggest that since the last mile delivery processes in freight and passenger transport mostly use the same resources of a city (e.g., infrastructure such as roads), it may be a good practice to share these resources in a more efficient way. Three perspectives are considered to assess the impact of increased sharing between freight and passenger transport: space, vehicle and time. The authors suggest that an increased sharing can be achieved by using crowdsourced delivery solutions (i.e., CS), and/or by using public transport vehicles for freights as well as passengers (i.e., SLDFC). For SLDFC, the Bussgods example is given. Bussgods is a Swedish company that uses passenger buses to transfer goods within the country. We would like to note that this example demonstrates the use of scheduled line vehicles for *long-distance* freight transfers, as opposed to our proposed system in which SLVs are used for last mile deliveries.

Another paper investigating the freight and passenger first/last mile (FLM) operations in a holistic approach is a recent one by Nocera et al. (2020). The paper identifies five phases for FLM operations planning and gives some possible strategies that can be used to reduce the number of vehicles needed with the ultimate goal of reducing overall costs. Combining freight and passenger transport is given as a possible strategy in FLM in this regard.

We would like to conclude the discussion on conceptual studies by mentioning a comprehensive paper by Savelsbergh and Van Woensel (2016) that focuses on the challenges and opportunities for urban logistics. The authors list five trends that are the main drivers for the developments in urban logistics: increasing population and urbanization, e-commerce growth, growing need for speed and timely deliveries, the sharing economy, and sustainability. The authors also mention several current advances in the technology domain that can alleviate the negative externalities caused by the increased pressure that the listed trends put on urban logistics systems. The paper underlines that offering a good-quality logistics service in an urban area and, furthermore, doing it at a low cost requires coordinating the flows of packages, consolidating the goods as much as possible, and co-operating among multiple organizations. There are several opportunities given in the paper that can enable low-cost but high-quality logistics service offerings. These opportunities include both CS and SLDFC, as we discuss in this thesis. The authors suggest that CS is a dynamic delivery system by itself, as both the demand and resources (i.e., delivery capacity) arrivals are realized over time, so the concept leads to new optimization challenges. Also, SLDFC is mentioned in the paper as a way to integrate public and freight transportation networks. The authors say that excess capacity in PTVs can be used to replenish stores, and this approach would be very effective as most public transport stations are located in “time- and vehicle-restricted [...] areas”.

2.2 Empirical Studies Regarding CS and SLDFC

2.2.1 CS-based Empirical Studies

The empirical studies focusing on CS operations often aim to pinpoint main factors that play important roles in the success of CS-based transport systems (i.e., drivers or encouraging factors) and try to identify (e.g., via surveys or interviews) any barriers to consider in the implementation step of such systems. A work by Briffaz and Darvey (2016) builds a framework for a descriptive study to investigate how relevant CS is for the stakeholders of Geneva, Switzerland. Surveying the potential customers of a CS system in Geneva, the authors found out that 74% of the respondents would be willing to use CS services. Survey results also show that low prices and increased flexibility are the two most motivating factors for using the services. Another interesting result of the study is that among 64% who say they would be willing to carry packages in a CS system, 44% state that they would do it for the ecological interest, while only 28% say they would be interested in the system for their own financial gains.

A study by Paloheimo et al. (2016) conducted in the city of Jyväskylä in Finland involves an actual CS trial in which 140 crowdshippers (i.e., drivers) and 85 customers are registered for home-deliveries from/to the city library. The results of this 6-week trial present important insights on how the drivers and customers perceive such as system. For example, before the trial, both the drivers' and customers' main concern was about the scheduling of the deliveries. This result suggests that any CS system should aim for high flexibility. The use of a pick-up point network to allow unattended deliveries would be a good solution in this regard, as we also propose in this thesis. As for the encouraging factors, the trial results show that most customers choose to participate in such a system to make life easier, while most drivers say they are in the system because they want to try something new. Potential financial benefits are found to be less important of an encouraging factor as only around 30% of the drivers and less than 5% of the customers cited this as the reason for their participation.

As with the last two studies mentioned, we see that the willingness to act as a crowdshipper is of particular interest and investigated by many researchers. Another study regarding this is the paper by Serafini et al. (2018). This paper is especially relevant to our thesis as it considers public transport as the main medium used for CS operations, as we similarly do in our proposed system. The authors use stated preference (SP) exercises to pinpoint the encouraging factors for the crowd to act as crowdshippers. Also, the discrete choice models are utilized to investigate underlying behaviors. The authors collect the study's data by holding 240 interviews with the Rome residents who use the metro system. A section of the survey includes questions about the four predefined features that describe the CS system's characteristics: location of automated parcel lockers (APLs) (inside or outside of metro stations), remuneration (3 or 1 Euros per delivery), delivery booking (real-time or off-line), and bank crediting modes (after each delivery or every 5 deliveries). Then, considering these features, the authors introduce four possible scenarios that are most likely to be encountered under actual CS platform conditions. The results suggest that the best one of these four scenarios leads to an 86.4% probability of acting as a crowdshipper, while the same probability for the worst scenario is only 12.8%. This percentage of 12.8% is due to the fact that APLs are outside of stations, and enumeration is low (1 Euro/delivery) in this scenario. On the other hand, if the APLs were inside the stations while the enumeration is still kept as low, this percentage is shown to be increased to 54.8%. This is an important insight as it shows having some pick-up and drop-off points inside the metro stations increases the efficacy and reliability of the system by attracting more crowdshippers. Later on, a follow-up study to Serafini et al. (2018) is conducted. This more recent study (Gatta et al., 2019) also includes a demand-side survey in addition to a supply-side survey. This new survey includes four dimensions of CS systems from the perspective of service users: shipping fees (typical or lower than currently available), shipping time (typical or lower than currently available), parcel tracking (available or not), and time flexibility (yes or no). Again, the authors select four scenarios that are deemed to be possible under typical CS conditions and observe that willingness to use CS services change between 12.4% and 66.1% among those scenarios.

Another study exploring the factors that influence the crowd's willingness to act as crowdshippers is a recent one by Dietmann (2020). The study's methodology includes a research model that aims to assess the encouraging factors under four main categories: intrinsic motivation, extrinsic motivation, shipping conditions, and habits and lifestyle. The author builds a survey for assessment purposes and conducts this survey online by reaching out to potential participants by using social media platforms such as LinkedIn and Xing. The survey results show that the willingness to participate depends on different factors such as the individual's experience carrying parcels, their free capacity, their detour acceptance threshold, and the size of the package to be transported.

2.2.2 SLDFC-based Empirical Studies

Similar to their CS-based counterparts, SLDFC-based empirical studies mainly aim to identify the challenges and barriers in the process of combining passenger and freight transport by conducting surveys or interviews and investigate some pilot projects/cases to estimate the benefits of fully functioning SLDFC systems.

The paper by Cochrane et al. (2016), for example, presents the results of a three-round Delphi study involving 34 transportation experts. The paper's main objective is to explore challenges and opportunities related to possible *freight on transit* (FOT) operations and build a conceptual framework to assess potential FOT operations in Toronto, Canada. The survey results show that the transportation experts think that the two most positive impacts of FOT operations are reduced energy use (i.e., savings in fuel consumption) and the creation of a new domain for revenue for public transit agencies. The two most notable negative impacts, according to the experts, are high initial investment costs and an increase in overall costs due to additional handling and transfers. As for the challenges in implementation, the experts cite financing issues and the perceived scarce of capacity on crowded public transit networks. As their discussion on the feasibility and implementability, the authors suggest that FOT should be combined with low-impact delivery modes (e.g., bikes, freight tricycles) to avoid the creation of any freight-related traffic near transit stations. The system proposed in this thesis, in which the public transit is combined with CS and/or environmental-friendly backup transfer vehicles (i.e., electric motorbikes), is in line with this discussion of the paper's authors.

The early idea, and to some extent, the practice of transporting freight in PTVs, date back to the 20th and even 19th century when mails were transferred via buses or trams in some cities of Belgium and Germany (De Langhe et al., 2019). In the early 2010s, this concept is rethought and brought back, mainly to build sustainable business models that involve carrying both packages and passengers in the same PTV in an urban area. One pilot study, called Cargo Hitching, was started in the Netherlands in 2013. There is a case study by Van Duin et al. (2019) that focuses particularly on this pilot project where two intracity passenger bus lines are utilized to transfer freight as well. Another pilot project was conducted in the city of Sapporo, Japan, by Kikuta et al. (2012). This project combined the passenger and freight transfers in the city's subway system by using a novel delivery cart prototype for packages. According to the questionnaire conducted by the authors, 88% of the respondents approved this project. The most eminent reason for approving the project is the perceived possibility of a decrease in inner-city congestion that will happen due to a decrease in the number of on-street parking freight vehicles. Also, it is worthwhile to note that 76.5% of the survey participants state that they find carriers with freight goods on board the subway vehicles “nothing annoying”. This implies that there is no significant barrier in using PTVs to carry freight in terms of passenger acceptance. If adapted for subway vehicles, this approach is projected to lead to a 47.6% reduction in carbon dioxide emissions (Zhou and Zhang, 2020).

Another study that investigates the practical use of PTVs for freight transportation is a recent paper by Bruzzone et al. (2021). In this paper, the authors identify several key performance indicators (KPIs) to assess the system before and after the integration of passenger and freight transport (IPFL) and consider the cases of Venice, Italy, and Velenje, Slovenia. The developed KPIs include average daily traffic, variation of traveled distances, service frequency, and energy requirements. A detailed analysis of the Venice case in accordance with these KPIs reveals that the implementation of IPFL reduces the total external costs from emissions around

15%, while the daily total energy use of PTVs decreases approximately 25%.

2.3 Decision Problems on the Design and Management of CS or SLDFC Systems

One of the earliest studies on CS modeling is a study by Archetti et al. (2016), in which the authors consider a system that allows a company to utilize the services of occasional drivers (ODs) (i.e., crowd-shippers) alongside the use of a self-owned fleet and dedicated drivers. Introducing "the vehicle routing problem with occasional drivers" (VRPOD) to the literature, the authors devise a multi-start heuristic to produce solutions for realistic problem instances and share their insights on the cost-reducing potential of the occasional drivers have in the system. In a follow-up study, Macrina et al. (2017) propose an extension to VRPOD, allowing ODs to make more than one delivery and considering a split delivery policy for ODs. The authors underline that the use of CS systems can reduce emissions and congestion mainly because the ODs perform journeys that are already planned to take place, which leads to a reduction in the number of vehicles used along with the total traveled distance.

In a more recent study by Dahle et al. (2019), authors extend the state-of-the-art to a "pickup and delivery problem with time windows and capacities", in which the system in consideration allows ODs to serve more than one customer and ODs can deny serving a request if the compensation offered by the company is lower than a personal threshold. Such a setting allows for studying the effects of different compensation schemes. The authors present three schemes in this regard and show that their model can lead up to 10-15% cost savings. They also suggest that more complex compensation schemes may result in larger cost savings.

In another study, Arslan et al. (2019) investigate a dynamic version of VRPOD and analyze CS's potential benefits in one-to-many, few-to-many, and many-to-many deliveries. Our work differs from the aforementioned studies in several aspects. First, the system we investigate varies notably from the CS settings described in these previous studies, as our proposed concept allows —and in most cases, requires— a coordinated, scheduled line journey for packages, along with CS transfers. In other words, the discussed studies do not entail transshipment for packages, yet in our work, a package is carried with the effort of possibly more than one SLVs, and crowdshippers. Second, our work takes a strategic approach as to, for example, where to employ CS or back-up fleet operations and what stations of the scheduled line to use as PT connections for these operations. In addition, our proposed model also deals with the routing decisions for the BT tours and incorporates the uncertainties in demand and delivery (CS) capacity that are closely related to the location decisions.

Macrina et al. (2020) considers a CS setting in which crowdshippers (i.e., ODs) can choose where they would like to pick up the packages. To make the system more appealing to ODs, the CS service network includes transshipment nodes as well as the main depots. In fact, the authors suggest that the implementation of transshipment nodes can increase the number of available ODs in the system as they serve as intermediate depots closer to the destinations. On the other hand, traditional vehicles still exist in this setting as they are to be used to serve intermediate nodes to make parcels available there. The paper defines the related problem as "the vehicle routing problem with occasional drivers, time windows, and transshipment nodes". An MIP is suggested for modeling the problem to minimize the total cost, which consists of costs associated with traditional vehicles, plus the compensation paid to ODs. The authors propose a metaheuristic algorithm (i.e., variable neighborhood search) for solving the problem, and they test the algorithm on large-size problem instances.

Ghilas et al. (2013) is one of the pioneering studies on decision aspects of SLDFC. The authors introduce the Pickup and Delivery Problem with Fixed Scheduled Lines (PDP-FL) to the literature, pointing to the increased capacity utilization potential of integrating public and freight transport. Authors propose a delivery system in which delivery requests are picked up by a pickup and delivery (PD) vehicle and transferred to a "station-hub", which is a part of a fixed scheduled line. Then the request is transferred to its destination station-hub by an SLV. From there, the request is again picked up by a PD vehicle and delivered to its final destination.

A conceptually similar study is by Masson et al. (2017), which focuses on a twotiered transportation problem. City buses are used to transfer parcels to a set of bus stops in the first tier, and city freighters take the parcels to their final destinations in the second tier. The authors introduce a mathematical model to minimize the number of necessary city freighters and travel costs. They devise an adaptive large neighborhood search to solve the problem. This system is actually very similar to ours, with the distinction that our system allows efforts of both PD vehicles (BT vehicles in our case) and crowd-shippers for the transfers to/from hubs. Moreover, our problem also deals with the optimal deployment of the PT connections, along with the routes of BTs, as opposed to PDP-FL, whose solution offers only route plans for PDs.

In a more recent study by Li et al. (2021), authors focus on an urban rail service design problem that would allow collaborative transport of passengers and freight. They mention two ways of integrating package transfers into the existing passenger rail network. Either dedicated freight trains can be used, or –more in line with the idea of SLDFC–, packages can be carried inside passenger trains (i.e., shared vehicles). The paper introduces an optimization problem to maximize the profit of the designed freight transport service. The problem is formulated as an MIP, and two instances derived from two real-life cases are solved with two heuristic iterative algorithms.

2.4 PCBP and Hub Network Design Problems

Considering PT connections as hubs, our problem falls into the broad class of hub network design problems (O'Kelly and Miller, 1994). For the vast literature on the topic, we refer the interested reader to the detailed analyses by Alumur and Kara (2008), Campbell and O'Kelly (2012), Farahani et al. (2013), and recently by Alumur et al. (2020). Our study differs from *standard* hub network problems in five dimensions within this literature. First, we do not assume a complete hub network and explicitly consider the connectivity requirements for PT connections to reach out to each other using transshipments. Second, our model does not assume direct transfers between the hubs (PT connections) and spokes (satellites) and allows

BT vehicles to visit several satellites. Third, in our model, the need for visiting a satellite with a BT vehicle is determined by the choice of the PT connections to open. Forth, focusing on an urban transportation setting, we allow BTs to start and end their tours from different PT connections, which is not considered in classical hub network design problems that typically focus on long-distance transfers. Finally, we consider dynamic routes to connect satellites to PT connections as opposed to static routing decisions that disregard uncertainties in demand and capacity. Some of these extensions are individually investigated in the hub network design literature (Alumur et al., 2009, Contreras et al., 2011, Yaman et al., 2007). However, to the best of our knowledge, our study is the first to consider all these extensions together.

2.5 Contributions and organization of the thesis

The contributions of this thesis are three-fold.

- We propose a new last-mile delivery system that can offer low-cost and environment-friendly express deliveries between any two points in the service region. The suggested approach combines several emerging technologies/business models in a way to make up for each other's shortcomings and boost their advantages. Focusing on the strategic level design problem of this novel last-mile delivery system, we introduce PCBP to the literature.

- We propose a path-based two-stage stochastic programming formulation to model PCBP and develop a branch-and-price algorithm for its solution. To the best of our knowledge, this study is the first one to use decomposition branching in a branch-and-price framework to exploit a specific structure that emerges in many real-world settings.

- We conduct an extensive computational study that attests to the efficacy of our solution approach and reveals critical managerial insights. In particular, our results indicate that:

    - The suggested BP approach can effectively solve large size realistic problem instances. Taking advantage of the nearly decomposable structure of the problem we study in our numerical experiments, extending the basic BP algorithm with decomposition branching can provide substantial computational performance improvements. According to our results, DB can provide up to an 87.3% decrease in run time, where the average run time reduction for all the problem instances is 56.3%.

– With a reasonable amount of crowd provided delivery capacity, only to execute short distance transfers between the satellites and the PT connections, it does not require deploying a large number of PT connections and incurring high operational costs for backup transfers to offer an environmentally friendly express delivery service over a large urban area.

– Any measures to increase CS participation rates can significantly affect the system's overall efficiency since even a moderate increase in CS supply can substantially decrease the required number of PT connections as well as the need for BT operations. Along with the CS participation rates, consideration of uncertainties on crowd-provided delivery capacity has a critical impact on the optimal network design.

2.6 Organization of the Thesis

The rest of the paper is organized as follows. Chapter 3 presents a thorough description of our proposed system, the formal definition of the related design problem (i.e., PCBP), and introduces the formulated two stage stochastic program. Chapter 4 clarifies the details of the solution approach we suggest to solve PCBP. In Chapter 5, we detail our approach to create problem instances, present the results of our numerical experiments and discuss the computational and managerial insights we derive from them. Finally, we conclude with some final remarks in Chapter 6.

Chapter 3

PROBLEM DEFINITION AND MODEL FORMULATION

In this chapter, we first provide the definitions and notation we use to define PCBP and present its mathematical model. Then we explain the details of the solution approaches we develop to solve it.

3.1 Notation and problem definition

Let S be the set of pick-up and delivery points (satellites). Similarly, let H be the set of candidate locations for the PT connections. We consider a scenario-based approach to incorporate the uncertainties in delivery demand and supply (i.e., crowd provided delivery capacity) in our model. The set Ω contains the scenarios we consider in our model. A scenario $\omega \in \Omega$ is characterized by the sets $S_h(\omega)$, $h \in H$. For a scenario $\omega \in \Omega$, the set $S_h(\omega)$ contains the satellites whose all inbound and outbound packages can be transferred to PT connections by the "crowd" if a PT connection is opened in h. Here we would like to note that, although outbound packages of a satellite can be dropped off by crowd-shippers at any PT connection, for incoming packages, the platform needs to designate one PT connection per satellite to offer corresponding packages to the crowd-shippers. That is, a satellite's incoming packages are to be distributed by the crowd-shippers from a designated PT connection, which we allow in our model to be determined dynamically for each distribution cycle, considering the realized CS delivery capacity in it. However, without any significant change in our model, it is also possible to consider operationally easier policies, such as assigning satellites to the closest open PT connection. Thinking that it would be easier to find crowd-shippers to transfer packages between two close-by locations, such an assignment may indeed provide the best one, as we observe in our numerical experiments. The probability associated with a scenario

$\omega \in \Omega$ is denoted with p^ω, and we have $\sum_{\omega \in \Omega} p^\omega = 1$.

Let $\bar{H} \subseteq H$ be the set of PT connections deployed by the platform. We refer to the satellites in $S_{\bar{H}}(\omega) = \bigcup_{h \in \bar{H}} S_h(\omega)$ as *satellites covered by CS*, in scenario ω, for the network design $\bar{H}$. In a given design $\bar{H}$, if CS cannot cover a satellite $s \in S$ in a scenario ω, then s needs to be serviced by a backup transfer to fulfill the related demand in case ω is realized. A backup transfer tour (also referred to as route or path) starts from a PT connection $h_1 \in \bar{H}$, visits one or more satellites to fulfill their corresponding package drop-off and pick-up demands and returns to $h_2 \in \bar{H}$ within the touring time limit $t_{\max}$ and the tour length limit $d_{\max}$. As mentioned before, to ensure a certain service level (ensure short delivery lead times), it is imperative to have frequent distribution cycles and short BT time limits (t_{max}), combined with fast and high-capacity PT backbones such as mass rapid transit systems. This initial study assumes that network management has already made such strategic decisions and focuses on the resulting network design problem. We leave it to future studies to develop models that specifically focus on these critical aspects of the problem.

A backup transfer tour is denoted with π, which is characterized by a starting PT connection o_π, an ending PT connection d_π, and a sequence of satellites ℓ_π to visit in between. For the sake of simplicity, we assume that BT vehicles follow the shortest path (which may be different from the fastest path) to move between any two points in their itineraries. The shortest path distance and travel time between two locations $i, j \in S \cup H$ is denoted by d_{ij} and t_{ij} respectively. We say a backup transfer tour π is feasible if the total length of the route (d_π) is no more than the vehicle range (i.e., battery range of the e-motorcycle), and the tour time t_π is less than $t_{\max}$. The total cost associated with a BT tour π is given as c_π, which consists of the fixed cost $\bar{c}$ and the per-kilometer cost $\bar{\bar{c}}$.

The set of all the feasible BT tours is denoted by P. For a satellite $s \in S$, we define $P(s)$ as the set of BT tours that visit the satellite, i.e., $P(s)$ includes all π such that $s \in \ell_\pi$. Similarly, $\hat{P}(s,h)$ and $\hat{\hat{P}}(s,h)$ are the sets of BT tours that include satellite s and start or end in PT connection h, respectively.

Before discussing how to incorporate the connectivity requirements into the model, as discussed in Figure 1.1, we want to underline that the PT network to consider in PCBP does not need to include all the PT lines in the urban area. A suitable subset of the existing PT lines would suffice to form the "backbone" of the delivery system we suggest in this study. In that regard, with high load capacities and frequent trips, metro lines, bus rapid transit (BRT) systems, and light rail systems can be considered as good candidates to form the backbone. We also want to note that most real-world PT networks have cycle free hub-and-spoke structures (Stiglic et al., 2018), which is especially true for metro, BRT, and light rail systems. From the modeling perspective, assuming cycle-free PT networks can simplify the imposition of the connectivity requirements. Note that, if the considered PT network has a tree structure (i.e., does not have any cycles), then for any pair of PT connections $h \in H$ and $m \in H$, there exists a unique set of PT connections $O(h, m)$ which needs to be active (open) to ensure connectivity of h and m. The suggested methodology to find the sets $O(h, m), h \in H, m \in H \setminus \{h\}$, is given in Appendix A. To simplify the exposition and considering that it would not be unrealistic for most real-world applications, we introduce our formulation and solution approaches assuming a cycle free PT network. However, in Appendix B, we present a straightforward modification for our formulation and respective changes in the solution methodology to relax this assumption, if needed.

As the main motivation behind the delivery model we propose in this study, we assume that transfers within the PT lines and deliveries made by the crowd-shippers cost much less than using BTs. Clearly, for small package transfers, this is not an unrealistic assumption, especially for the high-capacity PT lines we envision to use. As mentioned before, for the CS, we consider inserting package transfers into the already planned trips of the crowd, instead of requiring dedicated trips. Since the burden on the crowd-shippers is minimal in such a setting, the compensations paid for those deliveries, consistent with the existing studies, we think it is reasonable to assume that CS deliveries would cost less than the deliveries by BTs (Arslan et al., 2019, Yildiz, 2021). Consequently, we define PCBP as to find optimal locations for a given number of PT connections among a given set of alternatives H, such that the expected costs due to use of BTs are minimized to carry out express shipments between a given set of satellites S, considering a given set of scenarios Ω. For easy reference, we present the outline of the basic notation in Table 3.1.

Indices	**Description**
h or m	Index denoting a PT connection candidate
s	Index denoting a satellite
ω	Index denoting a scenario
π	Index denoting a BT tour
Sets	**Description**
S	Set of pick-up and delivery points, i.e., satellites
H	Set of candidate locations for PT connections
Ω	Set of CS capacity scenarios
$S_h(\omega)$	Set that contains the satellites whose all inbound and outbound packages can be handled by the crowd in scenario ω if a PT connection is opened in h
P	Set of all feasible BT tours
$P(s)$	Set of feasible BT tours that visit satellite s
$\hat{P}(s, h)$	Set of feasible BT tours that visit satellite s and start in h
$\hat{\hat{P}}(s, h)$	Set of feasible BT tours that visit satellite s and end in h
$O(h, m)$	Set of other PT connections that must be opened for PT connections h and m to be interconnected
$P_h(\omega)$	Set of feasible BT routes that start from h and visit at least one satellite from $S_h(\omega)$
$H(\omega, s)$	Set of PT connections that enable crowd-shipping for satellite s in scenario ω
Params	**Description**
k	Number of PT connections to be opened in the city
p^ω	Probability of realizing scenario ω
d_{ij}	Shortest path distance between the locations $i, j \in S \cup H$ in kilometers
t_{ij}	Shortest path travel time between the locations $i, j \in S \cup H$
d_π	Total length of the tour $\pi \in P$ in kilometers
t_π	Total duration of the tour $\pi \in P$
$t_{\max}$	Maximum allowed tour duration for a BT tour
c_π	Total cost of tour $\pi \in P$
$\bar{c}$	Fixed cost associated with utilizing a BT tour
$\bar{\bar{c}}$	Per-kilometer cost of a BT tour
o_π	Origin PT connection of a tour $\pi \in P$
d_π	Destination PT connection of a tour $\pi \in P$
ℓ_π	Sequence of visited satellites in tour $\pi \in P$
$\hat{c}^\omega_\pi$	The reduced cost of a routing variable $\pi \in P$ in scenario $\omega \in \Omega$
Variables	**Description**
x_h	Binary variable that indicates whether a PT connection is deployed in candidate location $h \in H$
v^ω_π	Binary variable that indicates whether the tour $\pi \in P$ is used in scenario $\omega \in \Omega$

Table 3.1: Outline of the notation.

3.2 Two-stage stochastic programming formulation

We formulate PCBP as an integer (binary) program with two sets of decision variables. The *design variables* x_h, $h \in H$, indicate whether to open a PT connection in location $h \in H$ or not. The *routing variables* v_π^ω, $\pi \in P$, $\omega \in \Omega$ indicate whether the backup transfer tour π is used in scenario ω or not. The design variables represent the first stage decisions, whereas the routing variables represent the second stage decisions in the two-stage stochastic program $\mathcal{P}$, which we formally define as follows.

$$
\begin{aligned}
(\mathcal{P}) \ \min \ & z = \sum_{\omega \in \Omega} \sum_{\pi \in P} p^\omega c_\pi v_\pi^\omega && && (3.1)\\
\text{s.t.} \ & \sum_{\pi \in P(s)} v_\pi^\omega + \sum_{h \in H(\omega,s)} x_h \geq 1 && \forall \omega \in \Omega, \ \forall s \in S && (3.2)\\
& \sum_{\pi \in \hat{P}(s,h)} v_\pi^\omega \leq x_h && \forall \omega \in \Omega, \ \forall h \in H, \ \forall s \in S && (3.3)\\
& \sum_{\pi \in \hat{\hat{P}}(s,h)} v_\pi^\omega \leq x_h && \forall \omega \in \Omega, \ \forall h \in H, \ \forall s \in S && (3.4)\\
& \sum_{h \in H} x_h = k && && (3.5)\\
& x_h + x_m - x_n \leq 1 && \forall h, m \in H, h \neq m, \forall n \in O(h,m) && (3.6)\\
& x_h \in \{0,1\} && \forall h \in H && (3.7)\\
& v_\pi^\omega \in \{0,1\} && \forall \omega \in \Omega, \ \forall \pi \in P && (3.8)
\end{aligned}
$$

The objective is to minimize the expected cost of BT tours. Constraints (3.2) ensure that each satellite is covered, either by backup transfers or by CS. Constraints (3.3) and (3.4) are to avoid the use of any BT tours unless their starting and ending points are designated as PT connections. Note that, given that the problem is feasible, there is always an optimal solution that satisfies (3.3) and (3.4), simply because it is unnecessary to visit a satellite more than once (with more than one BT vehicle). Taking advantage of this observation we avoid the following big-M type constraint that might be considered as “more intuitive” but result in a much weaker formulation.

$$\sum_{p \in P: o_p = h} v_\pi^\omega \leq |S| x_{o_\pi} \qquad \forall \pi \in P, \omega \in \Omega \tag{3.9}$$

$$\sum_{p \in P: d_p = h} v_\pi^\omega \leq |S| x_{o_\pi} \qquad \forall \pi \in P, \omega \in \Omega \tag{3.10}$$

Alternatively, one could consider replacing (3.3) and (3.4) with the inequalities (3.11) and (3.12) to avoid the weak linear relaxation bounds due to big-M constraints. However, in that case, the resulting formulation would include a large number of constraints (one for each routing variable), and present additional difficulties to implement the column generation based solution approach we develop in the following subsections.

$$v_\pi^\omega \leq x_{o_\pi} \qquad \forall \pi \in P,\ \omega \in \Omega \tag{3.11}$$

$$v_\pi^\omega \leq x_{d_\pi} \qquad \forall \pi \in P,\ \omega \in \Omega \tag{3.12}$$

The number of PT connections to be opened is limited by the Constraint (3.5). Assuming a tree structure for the PT network, as mentioned before, Constraints (3.6) ensure interconnectivity. Finally, the Constraints (3.7) and (3.8) define the domain restrictions for the decision variables.

Since the number of the routing variables grows exponentially with the number of satellites and PT connections, it is not possible to solve $\mathcal{P}$ directly for the problem instances one needs to solve in real-world applications. To overcome this difficulty, we develop a Branch-and-Price algorithm as we describe in the next chapter.

Chapter 4

BRANCH-AND-PRICE AND DECOMPOSITION BRANCHING ALGORITHMS

We have defined the problem PCBP and given the relevant notation in the previous chapter. In this chapter, we outline our suggested methodology to efficiently solve realistic PCBP instances.

4.1 Solving PCBP with Branch-and-Price

Although the number of feasible routes becomes too large in realistic size problem instances, only a few of them will be used in the optimal PCBP solution. Therefore we propose a branch-and-price approach to solve $\mathcal{P}$, which essentially combines the branch-and-bound framework with a column generation procedure to add routing variables iteratively as needed (Barnhart et al., 1998). In the following sections, we present the details of the column generation procedure we use to solve linear relaxations at the branch-and-bound nodes and describe the branching rules we use to develop the search tree.

4.1.1 LP Solution (Column Generation)

Let $\mathcal{P}_{LP}$ denote the linear relaxation of $\mathcal{P}$. Furthermore, let $\mathcal{P}_{LP}^k$ be the restricted $\mathcal{P}_{LP}$ formulation in k_{th} column generation (CG) iteration, which includes all of the design variables but only a subset of the routing variables that are included in the model in the previous CG iterations. At each CG iteration, we solve a *pricing problem* to identify the routing variables that are not in the restricted formulation but have a negative reduced cost or conclude that there are no such variables and the algorithm can be terminated.

Pricing Problem

Let δ, ϕ and ψ represent the dual variables associated with the constraints (3.2), (3.3) and (3.4) in $\mathcal{P}^k_{LP}$, for some $k \geq 0$. Then the reduced cost $\hat{c}^{\omega}_{\pi}$ of a routing variable v_{π} in scenario ω can be calculated as follows.

$$\hat{c}^{\omega}_{\pi} = p^{\omega} c_{\pi} - \left(\sum_{s \in \ell_{\pi}} [\delta^{\omega}_s + \phi^{\omega}_{o_{\pi} s} + \psi^{\omega}_{d_{\pi} s}] \right) \tag{4.1}$$

To look for a path with negative reduced cost in a given scenario $\omega \in \Omega$, we use an exhaustive search algorithm (SA) that is equipped with several enhancement strategies to improve computational efficiency. Before proceeding with the details, we first provide some new notations and definitions we use to describe SA and formally state a few straightforward results we use to improve its computational efficiency.

Definition 1. *A partial path γ is composed of a starting PT connection h_{γ} and a sequence of satellites to visit ℓ_{γ}. If $\ell_{\gamma} = \emptyset$, we say γ is empty. A non-empty partial path γ has a length l_{γ} (the total distance to cover starting from h_{γ} and visiting the satellites ℓ_{γ} in the given order), elapsed time t_{γ} (the time it takes to complete the partial tour), cost $c_{\gamma} = \bar{c} + \bar{\bar{c}} l_{\gamma}$, and the reduced cost $\hat{c}_{\gamma} = p^{\omega} c_{\gamma} - \left(\sum_{s \in \ell_{\gamma}} \delta^{\omega}_s + \phi^{\omega}_{h_{\gamma} s} \right)$.*

Definition 2. *A partial path γ_s is called an augmentation of the partial path γ, if γ_s starts from the same PT connection as γ, and visits the same sequence of satellites before ending at s (i.e., $h_{\gamma_s} = h_{\gamma}$ and $\ell_{\gamma_s} = \{\ell_{\gamma}, s\}$).*

For notational convenience, for a partial path γ we also define $d_{\gamma i}$ and $t_{\gamma i}$ as the distance and the travel time between the last satellite in the sequence ℓ_{γ} and a point $i \in S \cup H$. For a partial path γ, the last satellite in the sequence ℓ_{γ} is denoted by s_{γ}.

Proposition 1. *For a scenario $\omega \in \Omega$, and a PT connection $h \in H$, let $P_h(\omega)$ be the set of feasible BT routes that start from h and visit at least one satellite from $S_h(\omega)$. There is always an optimal solution of PCBP that does not use any of the BT tours in $P_h(\omega)$, for all $\omega \in \Omega$, and $h \in H$.*

Proof. The result follows from observing that, feasible paths that start from a PT connection h can only be used if h is active (open). But when a PT connection h is active satellites in $S_h(\omega)$ are already covered in the scenario $\omega \in \Omega$ and there is no need to visit them by BT vehicles. □

Proposition 2. *Let π be a BT tour that is used in an optimal solution of PCBP for some scenario $\omega \in \Omega$. Then for any satellite $s \in \ell_\pi$, there cannot be a reordering of the satellites that come before it such that both the total distance and elapsed time to reach the satellite s is less than what is observed following the sequence ℓ_π up to s.*

Proof. The result simply follows from the fact that visiting the same PT connections and satellites the feasible path with the shortest distance has a smaller cost, assuming a positive per-kilometer cost for BT vehicles. □

Proposition 3. *Let π be a BT tour with the minimum reduced cost at column generation iteration k, for a scenario $\omega \in \Omega$. Assume π visits at least two satellites and let $\{s_1, s_2\}$ be a sub-sequence in ℓ_π such that $\bar{\bar{c}} d_{is_1} - \delta^\omega_{s_1} - \phi_{o_\pi s_1} > 0$, where i is the last point that π visits before s_1. Then it must be the case that $\bar{\bar{c}}\big(d_{is_1} + d_{s_1,s_2} - d_{is_2}\big) - \delta^\omega_{s_1} - \phi^\omega_{o_\pi s_1} \leq 0$.*

Proof. Assume that the claim is false, and there exists a sub-sequence $\{s_1, s_2\}$ in ℓ_π such that $\bar{\bar{c}} d_{is_1} - \delta^\omega_{s_1} - \phi_{o_\pi s_1} > 0$, but $A = \bar{\bar{c}}\big(d_{is_1} + d_{s_1,s_2} - d_{is_2}\big) - \delta^\omega_{s_1} - \phi^\omega_{o_\pi s_1} > 0$. Consider the path $\bar{\pi}$, which can be obtained from π by removing s_1 from the list ℓ_π. Note that $\bar{\pi}$ is a feasible path, since π is feasible. Also, we have $\hat{c}^\omega_\pi - \hat{c}^\omega_{\bar{\pi}} = A - \psi^\omega_{s_1 o_\pi} > 0$ (since $A > 0$, by assumption) and $\Psi^\omega_{s_1} \leq 0$ (by linear duality results). But $\hat{c}^\omega_\pi - \hat{c}^\omega_{\bar{\pi}} > 0$ contradicts π being a minimum reduced cost path in ω. Hence, the result follows. □

Proposition 4. *Let π be a feasible path with $\bar{s}$ as the last satellite in the sequence ℓ_π. The path π cannot be used in any optimal PCBP solution, if the per-kilometer BT vehicle cost $\bar{\bar{c}}$ is positive and $d_{\bar{s}o_\pi} < d_{\bar{s}d_\pi}$.*

Proof. Assume that the claim is false and there exists a π that is a part of the optimal PCBP solution and $d_{\bar{s}o_\pi} < d_{\bar{s}d_\pi}$. Consider the path $\bar{\pi}$, which we obtain from

π by replacing last PT connection with o_π. That is, we have $d_{\bar{\pi}} = o_{\bar{\pi}} = o_\pi$. Since $d_{\bar{s}o_\pi} < d_{\bar{s}d_\pi}$, we have $d_\pi > d_{\bar{\pi}}$. Observe that since π is feasible $\bar{\pi}$ is also feasible, since $\ell_{\bar{\pi}} = \ell_\pi$ and $\{d_{\bar{\pi}}\} \subset \{o_\pi, d_\pi\}$. This implies that $\bar{\pi}$ can replace π in the optimal solution and the objective will improve, which leads to a contradiction with π being used in the optimal solution. Hence the result follows. □

To solve the pricing problem at the k_{th} CG iteration, SA is initialized with an empty variable list V (list of routing variables to include in $\mathcal{P}_{LP}^{k+1}$) and set of partial paths Γ that contains an empty partial path γ_h for each $h \in H$. At each SA iteration, the partial path with the smallest reduced cost γ is removed from the candidate set Γ (ties are broken arbitrarily) and processed. To process a partial path γ (i.e., generate its augmentations) we consider the set of satellites S_γ^ω, which contains all the satellites $s \in S \setminus l_\gamma$ that satisfy the following conditions we impose to improve computational efficiency.

- After visiting s it is possible to reach a PT connection without violating the tour time and vehicle range limits, i.e., there exists some $h \in H$ such that $t_{\gamma_s} + t_{sh} \leq t_{\max}$ and $d_{\gamma_s} + d_{sh} \leq d_{max}$.

- In scenario ω, the satellite s is not covered by the PT connection γ_h, i.e., $s \notin S_h^\omega$ (imposed due to the Proposition-1).

- Starting from the PT connection h_γ, no reordering of ℓ_γ would provide an alternative path to reach s that is shorter both in distance and time (imposed due to the Proposition-2).

- The expression $\bar{\bar{c}} d_{s_\gamma s} - \delta_s^\omega - \phi_{h_\gamma s}$ is negative or there exits a satellite $\bar{s} \in S \setminus \ell_\gamma$ such that $d_{\gamma_s} + d_{s\bar{s}} + min_{h\in H}\{d_{\bar{s}h}\} < d_{max}$, $t_{\gamma_s} + t_{s\bar{s}} + min_{h\in H}\{t_{\bar{s}h}\} < t_{\max}$ and $\bar{\bar{c}}\big(d_{s_\gamma s} + d_{s,\bar{s}} - ds_\gamma\bar{s}\big) - \delta_{s_1}^\omega - \phi_{o_\pi s} < 0$ (imposed due to the Proposition-3).

Processing a partial path γ, we build the augmentations $\gamma_s, s \in S_\gamma^\omega$, and add them to Γ. At the same time, we also check whether any of the augmentations $\gamma_s, s \in S_\gamma^\omega$, can be "closed" as a feasible BT tour with a negative reduced cost. For that purpose, for each augmentation γ_s, we consider the set of possible *terminating PT connections* $H_s \subseteq H$, where H_s contains all the PT connections that satisfy the following feasibility and optimality conditions.

- $d_{\gamma_s} + d_{s_\gamma h} \leq d_{max}$ and $t_{\gamma_s} + t_{s_\gamma h} \leq t_{\max}$ (imposed to ensure the resulting path is feasible).
- $d_{\gamma_s h} < d_{\gamma_s h_{\gamma_s}}$ (imposed due to Proposition 4).

For each candidate termination $h \in H_s$ we check the reduced cost of the resulting BT tour π_h, and add the variable $v^{\omega}_{\pi_h}$ to V, if $\hat{c}^{\omega}_{\pi_h} < 0$. SA terminates when there is no partial path in Γ to process (i.e., $\Gamma = \emptyset$) or V is populated with at least m variables (i.e., $|V| = m$). With the parameter m, we aim to add multiple routing variables to the model in each CG iteration, which we observe to improve computational performance in our preliminary studies. A pseudocode of the algorithm SA is given in Algorithm 1.

Algorithm 1: SA

Input: δ, ϕ, ψ, H, Ω, S, m
Output: V

$V \leftarrow \{\}$
$\Gamma \leftarrow \{\gamma_h : h \in H\}$
do
  $\bar{\gamma} \leftarrow \arg\min_{\gamma \in \Gamma} \hat{c}_\gamma$
  Remove $\bar{\gamma}$ from Γ
  foreach $\omega \in \Omega$ **do**
    foreach $s \in S^{\omega}_{\bar{\gamma}}$ **do**
      Add $\bar{\gamma}_s$ to Γ
      foreach $h \in H_s$ **do**
        Let π be a route with $o_\pi = h_{\bar{\gamma}_s}$, $\ell_\pi = \ell_{\bar{\gamma}_s}$ and $d_\pi = h$.
        if $\hat{c}^{\omega}_{\pi} < 0$ **then**
          Add v^{ω}_{π} to V
while $\Gamma \neq \emptyset$ *and* $|V| < m$
return V

Initial Variable Pool

Let π_{hs} be the BT tour that starts at the PT connection $h \in H$, visits the satellite $s \in S$, and comes back to h. We initialize the CG procedure by including the routing variables $V^0 = \{v^{\omega}_{\pi_{hs}} : \omega \in \Omega, h \in H, s \in S \setminus S^{\omega}_h\}$, in the restricted formulation $\mathcal{P}^0_{LP}$. As the objective function coefficient of a variable $v^{\omega}_{\pi_{hs}} \in V^0$, we use c_π if the path π is feasible and use a sufficiently large value $M_{hs} = \bar{c}+\bar{\bar{c}}\big(max_{\bar{h}\in H}\{d_{\bar{h}s}\}+max_{\underline{h}\in H}\{d_{s\underline{h}}\}\big)$, which ensures that infeasible paths would not be utilized by an optimal solution, given that the PBCP instance is feasible. Including the routing variables V^0, we ensure that $\mathcal{P}^0_{LP}$ is feasible, as needed to initialize the CG procedure.

4.1.2 IP Solution

To build an efficient branch-and-price algorithm, it is critical to devise branching rules that eliminate any fractional solutions while keeping the pricing problem tractable (Barnhart et al., 1998). In the following, we define branching rules for both design variables and routing variables that achieve this goal. However, as we will show in the next chapter in more detail, our preliminary studies have shown that branching on the design variables is usually enough to obtain an optimal solution. Even when this is not the case, and branching solely on the design variables gives a solution $(\mathbf{x}^*, \mathbf{v}^*)$ with fractional values for some of the routing variables, one can obtain high-quality solutions with small optimality gaps (i.e., less than 0.5% in all our experiments) by simply fixing the design variable values to $\mathbf{x}^*$ and solving the resulting integer program by considering the routing variables generated during the BP iterations.

Branching rule for the design variables

We implement a rather straightforward branching rule for the design variables. For a BP node, after the PPs are solved and no additional BT route variable is found at the last iteration, we investigate x_h variables for $h \in H$ to see if any of them has a fractional value. Say $x_{\bar{h}} \notin \{0, 1\}$ for $\bar{h} \in H$. Then with the following branching cuts, we branch on the design variable $x_{\bar{h}}$.

$$x_{\bar{h}} \leq 0 \tag{4.2}$$

$$x_{\bar{h}} \geq 1 \tag{4.3}$$

If there is more than one design variable with a fractional value, for branching purposes, we prioritize the PT connections that belong to more than one PT line. Among the fractional design variables that belong to the highest priority level, we choose the one with a value closest to 0.5. Clearly, branching on the design variables does not disrupt the structure of the pricing problem and we can simply use the SA algorithm to solve it, without any modification.

Branching rule for the routing variables

Suppose, we solved $\mathcal{P}_{LP}$ in a Branch-and-Price node and obtained $v_\pi^\omega \notin \{0, 1\}$ for some $\omega \in \Omega$ and $\pi \in P$. This implies that we need to branch on v_π^ω to abide by the binary domain of the variable. Although setting it to 0 and 1 in two different branches provides a valid dichotomy, such a rule is hard to implement as it is computationally difficult to impose that a specific BT tour should not be used in the branch that we set the respective routing variable to zero. Therefore we suggest using a branching rule that is first proposed by Barnhart et al. (2000) and successfully used in various branch-and-price algorithms to solve network design problems in telecommunication and transportation networks (Yildiz and Karacsan, 2017, Yildiz et al., 2018). The fundamental idea is as follows. If there is at least one $v_\pi^\omega \notin \{0, 1\}$ for some $\omega \in \Omega$ and $\pi \in P$, we know that at least one satellite node $s \in S$ is visited by more than one fractional BTs. For demonstration purposes let us say there are only two fractional routing variables as $v_{\pi_1}^{\omega_1} = 0.5$, $v_{\pi_2}^{\omega_1} = 0.5$. Furthermore, let $s_1 \in S$ be a satellite that is included in both routes π_1 and π_2. Two observations can be made in this situation.

- π_1 and π_2 are different from each other by at least one node in the sequence. That is, if $o_{\pi_1} = o_{\pi_2}$, and $d_{\pi_1} = d_{\pi_2}$ it must be the case that $\ell_{\pi_1} \neq \ell_{\pi_2}$. In other words, the two tours must differ by at least one arc in the route, and,

- π_1 and π_2 have at least one node in common, one of which being s_1 since $s_1 \in \ell_{\pi_1}$ and $s_1 \in \ell_{\pi_2}$.

Based on these observations, we define the concept of *divergence node* in Definition 3.

Definition 3. *A divergence node of two or more BTs is the first node that the routes split after the last common node.*

For our given BTs π_1 and π_2 and generally, for any given two routes, we find the divergence node by tracing the paths arc by arc starting from the last common node until we find two different arcs called a_1 and a_2 for each route. Then we call the origin node of these two arcs divergence node, or d.

Let $A(d)$ denote the arcs originating from d. Furthermore, let $A(d, a_1) \subset A(d)$ and $A(d, a_2) \subset A(d)$ be any two subsets that include the arcs a_1 and a_2 respectively and satisfying $A(d, a_1) \cup A(d, a_2) = A(d)$. To obtain the subsets $A(d, a_1)$ and $A(d, a_2)$, one can partition $A(d)$ into two –arbitrarily– and check if a_1 and a_2 are in different subsets. If this is not the case, either a_1 or a_2 can be moved to other subset to obtain $A(d, a_1)$ and $A(d, a_2)$.

We define a new notation $P(s, E)$ that denotes a BT set including all the routes that stop by satellite $s \in S$ and include one arc from arc set E. Then, we go to the following branching for our given example:

$$\sum_{\pi \in P(s_1, A(d, a_1))} v_\pi^{\omega_1} = 0 \tag{4.4}$$

$$\sum_{\pi \in P(s_1, A(d, a_2))} v_\pi^{\omega_1} = 0 \tag{4.5}$$

Observe that in the first branch we do not allow the use of arcs in $A(d, a_1)$ and similarly, in the second branch we do not allow arcs in $A(d, a_2)$. This idea can be incorporated easily into our PP solving scheme, i.e., SA as follows. For example, for the paths including s_1, when we reach node d we simply remove the arcs of set $A(d, a_1)$ or $A(d, a_2)$ based on whether we are in the first or the second branch.

The given branching rule is valid because if we come across a fractional v_π^ω, we are guaranteed to find at least one common node $s \in S$, a divergence node d, and a set $A(d)$ that will eliminate the fractional solution. Furthermore, the number of branches is finite since there is a finite number of arcs (Barnhart et al., 2000).

4.2 Solving PCBP with Decomposition Branching

In this section, we explore the use of decomposition branching in the BP algorithm to take advantage of a specific nearly decomposable structure one may have in real-world PCBP instances. As a recent innovation in mixed-integer programming, decomposition branching combines two essential techniques for solving mixed-integer programs (MIPs): branch-and-bound and decomposition. It exploits the decomposable or nearly decomposable structures present in many MIPs to enhance computational efficiency (Yildiz et al., 2018). The main idea behind the suggested approach is to consider the linking constraints as allocation of a "common resource" among the sub-problems and build a branch-and-bound approach to find the optimal allocation. In the process, the aim is to solve a large MIP by solving smaller MIPs (sub-problems) several times for different resource allocations until the optimal allocation is found. In the sequence, we discuss the emergence of decomposable structures in PCBP and explain how we integrate decomposition branching into the BP framework.

Large cities consist of several divisions, disconnected from each other by either topographical features (rivers, lakes, forests, etc.) or administrative boundaries. It would be too costly or too time-consuming to have a BT vehicle to visit several points in different divisions of a city. We investigate one such example in our numerical experiments with problem instances from Istanbul, where the Bosphorus divides the city into two sides: European and Asian. Regarding this, we describe *service region*s in Definition 4.

Definition 4. *A service region (SR) is an area designated within a city to implement* *delivery operations independently from other parts of the city. All SRs must be interconnected with public transport lines.*

Let us assume that the area of focus is composed of $n > 1$ SRs. Note that, in PCBP we have a predetermined k value, i.e., we know beforehand the number of PT connections to be opened in the city. On the other hand, we do not know k_i, which is the number of PT connections that should be opened for a given SR $i = 1, 2, \ldots, n$ to minimize z. If we knew the optimal k_i such that $\sum_{i=1}^{n} k_i = k$, we would have a completely decomposable problem with n independent sub-problems. Therefore, we can think of the number of connections as a resource to be shared between n SRs.

Let I be the set of service regions in the city of focus, i.e., $I = \{1, 2, \ldots, n\}$. Using the described decomposable structure, we can write a new formulation $\bar{\mathcal{P}}$ that is equivalent to $\mathcal{P}$ given in Chapter 3. Most of the notation for $\mathcal{P}$ still holds for $\bar{\mathcal{P}}$ as well, but some of the set notations require some modifications which are given as follows. We define $S(i)$, $H(i)$, $P(i)$ as the set of satellites, PT connections, and BT tours of SR i, respectively. Similarly, we define $P(i, s)$ as the set of BT tours of SR i that visit satellite s. Furthermore, $\hat{P}(i, s, h)$ is defined as the set of BT tours of SR i that visit satellite s and start in h, and $\hat{\hat{P}}(i, s, h)$ is the set of BT tours of SR i that visit satellite s and end in h.

$$(\bar{\mathcal{P}}) \ \min \ z = \sum_{i=1}^{n} z_i \tag{4.6}$$

$$\text{s.t. } z_i \geq \sum_{\omega \in \Omega} \sum_{\pi \in P(i)} p^{\omega} c_{\pi} v_{\pi}^{\omega} \qquad \forall i \in I \tag{4.7}$$

$$\sum_{\pi \in P(i,s)} v_{\pi}^{\omega} + \sum_{h \in H(i,\omega,s)} x_h \geq 1 \qquad \forall \omega \in \Omega,\ \forall s \in S(i),\ \forall i \in I \tag{4.8}$$

$$\sum_{\pi \in \hat{P}(i,s,h)} v_{\pi}^{\omega} - x_h \leq 0 \qquad \forall \omega \in \Omega,\ \forall h \in H(i),\ \forall s \in S(i),\ \forall i \in I \tag{4.9}$$

$$\sum_{\pi \in \hat{\hat{P}}(i,s,h)} v_{\pi}^{\omega} - x_h \leq 0 \qquad \forall \omega \in \Omega,\ \forall h \in H(i),\ \forall s \in S(i),\ \forall i \in I \tag{4.10}$$

$$\sum_{i \in I} \sum_{h \in H(i)} x_h = k \tag{4.11}$$

$$x_h + x_m - x_n \leq 1 \qquad \forall h, m \in H(i), h \neq m, \forall n \in O(h, m),\ \forall i \in I \tag{4.12}$$

$$x_h \in \{0,1\} \qquad \forall h \in H(i),\ \forall i \in I \tag{4.13}$$

$$v^{\omega}_{\pi} \in \{0,1\} \qquad \forall \omega \in \Omega,\ \forall \pi \in P(i),\ \forall i \in I \tag{4.14}$$

Observe that in the formulation above, each set of constraints is decomposed except for the Constraint (4.11). In other words, except for the constraint on the number of PT connections to be opened, each set of constraints can be divided into smaller subsets, each defining limitations specific to different SRs.

Note that a BT tour π belongs to one and only one service region. Therefore, to write the reduced cost of a path variable v_π in scenario ω, we need to introduce only one new dual variable μ associated with the Constraints (4.7). Again, for the same reason, the previously defined dual variables associated with the Constraints (3.2)-(3.4) of $\mathcal{P}$ will be still valid for Constraints (4.8)-(4.10) to calculate the reduced cost of a path variable in formulation $\bar{\mathcal{P}}$. In particular, the reduced cost of a path variable v_π in scenario ω is given as follows.

$$\hat{\hat{c}}^{\omega}_{\pi} = \mu\ p^{\omega} c_{\pi} - \left(\sum_{s \in \ell_{\pi}} [\delta^{\omega}_{s} + \phi^{\omega}_{o_{\pi} s} + \psi^{\omega}_{d_{\pi} s}] \right) \tag{4.15}$$

The pricing problem defined under Section 4.1.1 continues to hold here with respect to its structure and solution methodology (i.e., extensive search algorithm). The only slight modification is on how we compute the reduced costs, which can be reflected into Algorithm 1 in a quite straightforward way by replacing $\hat{c}^{\omega}_{\pi}$ by $\hat{\hat{c}}^{\omega}_{\pi}$ as given in Equation (4.15) and $\hat{c}_{\gamma}$ by $\hat{\hat{c}}_{\gamma} = \mu\ p^{\omega} c_{\gamma} - \left(\sum_{s \in \ell_{\gamma}} \delta^{\omega}_{s} + \phi^{\omega}_{h_{\gamma} s} \right)$.

Noting that the Constraint (4.11) can also be written as $\sum_{i \in I} k_i = k$ and assuming a given k_i for a specific service region $i \in I$, we call the sub-problem associated with SR i as $DP_i(k_i)$. We formally define $DP_i(k_i)$ as follows.

$$(DP_i(k_i))\ \min\ z_i \tag{4.16}$$

$$\text{s.t.}\ z_i \geq \sum_{\omega \in \Omega} \sum_{\pi \in P(i)} p^{\omega} c_{\pi} v^{\omega}_{\pi} \tag{4.17}$$

$$\sum_{\pi \in P(i,s)} v^{\omega}_{\pi} + \sum_{h \in H(i,\omega,s)} x_h \geq 1 \quad \forall \omega \in \Omega,\ \forall s \in S(i) \tag{4.18}$$

$$\sum_{\pi \in \hat{P}(i,s,h)} v_\pi^\omega - x_h \le 0 \qquad \forall \omega \in \Omega,\ \forall h \in H(i),\ \forall s \in S(i) \tag{4.19}$$

$$\sum_{\pi \in \hat{\hat{P}}(i,s,h)} v_\pi^\omega - x_h \le 0 \qquad \forall \omega \in \Omega,\ \forall h \in H(i),\ \forall s \in S(i) \tag{4.20}$$

$$\sum_{h \in H(i)} x_h = k_i \tag{4.21}$$

$$x_h + x_m - x_n \le 1 \qquad \forall h, m \in H(i), h \neq m, \forall n \in O(h,m) \tag{4.22}$$

$$x_h \in \{0,1\} \qquad \forall h \in H \tag{4.23}$$

$$v_\pi^\omega \in \{0,1\} \qquad \forall \omega \in \Omega,\ \forall \pi \in P \tag{4.24}$$

If a sub-problem $DP_i(k_i)$ is feasible we denote its optimal solution value as $z^*_{ik_i}$. For computational convenience, we define $z^*_{ik_i} = \infty$ if the sub-problem is infeasible. We next show how to implement decomposition branching in the BP framework we develop to solve PCBP, by solving $DP_i(k_i)$ repeatedly for different k_i values, as needed.

Decomposition branching

At each branch-and-bound node with a solution $(\bar{\mathbf{x}}, \bar{\mathbf{v}})$ we first check if the quantities $\hat{k}_i = \sum_{h \in H(i)} \bar{x}_h,\ \forall i \in I$ are integers.

- If $\hat{k}_i$ is not integer for some $i \in I$, we generate two branches with the following cuts.

$$\sum_{h \in H(i)} x_h \le \lfloor \hat{k}_i \rfloor \ \wedge \ z_i \ge z^*_{i\lfloor \hat{k}_i \rfloor} \tag{4.25}$$

$$\sum_{h \in H(i)} x_h \ge \lceil \hat{k}_i \rceil \tag{4.26}$$

  The validity of the suggested dichotomy is obvious, and it cuts off the infeasible solution that implies opening a fractional number of PT connections in SR i. The branching cut (4.25) is further strengthened with an additional inequality, which imposes that there cannot exist any feasible solution in this part of the solution space that is better than the optimal solution of the sub-problem

$DP_i(k_i)$. Clearly, this is a valid cut for the respective branch-and-bound node, by the definition of $DP_i(k_i)$. Also observe that, the added cuts do not change the initial structure of the pricing problem as they do not include routing variables.

- If $\hat{k}_i$ are all integers for all $i \in I$, we start from the first SR (i.e., $i = 1$), and for each $i \in I$ we check whether the optimal solution of the sub-problem $DP_i(\hat{k}_i)$ is larger than $\bar{z}_{i\hat{k}_i} = \sum_{\omega \in \Omega} \sum_{\pi \in P(i)} p^\omega c_\pi \bar{v}^\omega_\pi$. If we detect this is the case for some sub-problem i (i.e., if there exists an i such that $z^*_{i\hat{k}_i} > \bar{z}_{i\hat{k}_i}$), we generate two branches with the following cuts.

$$\sum_{h \in H(i)} x_h \leq \hat{k}_i \;\wedge\; z_i \geq z^*_{i\hat{k}_i} \tag{4.27}$$

$$\sum_{h \in H(i)} x_h \geq \hat{k}_i + 1 \tag{4.28}$$

Clearly, if we have $z^*_{i\hat{k}_i} = \bar{z}_{i\hat{k}_i}$ for all $i \in I$, there is no need to explore the current branch-and-bound node further, and we can prune it after forming an integer solution by combining the sub-problem solutions and updating the upper-bound, if necessary. On the other hand, if there exists an SR i, for which we have $z^*_{i\hat{k}_i} > \bar{z}_{i\hat{k}_i}$, then it must be the case that some of the design variables related to the PT connections in i have fractional values in the given solution. In that case, the suggested branching rule cuts off this infeasible solution by imposing that z_i should be either as large as the optimal solution of the sub-problem $DP_i(k_i)$, i.e., $z_i \geq z^*_{i\hat{k}_i}$, or the number of PT connections opened in the SR i is at least $k_i + 1$.

The pseudocode for BP with DB (or shortly DB) is given in Algorithm 3 in Appendix C. Note that, we solve the sub-problems using another branch-and-price tree created according to the rules given in Section 4.1. Also, as we solve the sub-problems, we store their optimal solutions, in case they are needed later on in our main branch-and-price tree. We conclude this chapter by underlining that, as a significant advantage, the branching cuts in DB contains only the design variables.

Hence, they do not distort or alter the structure of the pricing problem, which we keep solving with the SA algorithm as described in Section 4.1.1 only with a small modification on reduced cost calculations.

Chapter 5

COMPUTATIONAL RESULTS

In this chapter we describe our experimental design, present the implementation details and discuss the computational and managerial insights we derive from our numerical experiments.

5.1 Problem instances

To obtain realistic spatial data to use in our problem instances, we consider the city of Istanbul and its public transportation system. Being the most populated city in Europe with more than 15 million inhabitants (Turkish Statistical Institute, 2019), Istanbul has an extensive and widely used urban public transport system spread over a large geographical area. PT network in Istanbul consists of different modes, such as metros (operated by Metro Istanbul), bus and bus-rapid-transit (BRT) services (operated by IETT), commuter trains (e.g., Marmaray, operated by Turkish State Railways), and ferryboats. In our experiments, we narrow down our focus to railway-based public transport services (RBPTS, namely, metros and Marmaray) for the following reasons.

- RBPTS consists of several interconnected lines that can serve all regions of the city seamlessly. This implies that, no matter how far the origin and destination satellites are from each other, every transfer request can be fulfilled thanks to this extensive coverage.

- Rapid transit and commuter trains offer a fast, reliable service that is not affected by traffic jams and road closures. Moreover, trains used for RBPTS usually operate with an underutilized capacity outside rush hour. This excess capacity can be easily dedicated to freight transport, and additional capacity can be employed by attaching new freight-only cars if needed. It is also worth noting that metro stations are convenient places to locate the equipment and staff required in a PT connection center. RBPTS lines are ideal candidates with these unique properties to serve as a backbone network to support express shipments within the city, as envisioned by our suggested system.

- Besides their popularity, metro lines are also well-integrated with other public transport modes and highly utilized as interchange points. This especially holds for our case of Istanbul, where passengers usually use RBPTS as the backbone of their journey combined with other modes such as buses, ferryboats, personnel vehicles, and private vehicles. Naturally, this integration offers ample opportunities to insert package transfers to crowd's planned trips, from satellites to PT connections, or vice versa. The recent studies also verify the strong potential of employing RBPTS riders as crowd-shippers (Dietmann, 2020).

The spatial data remain the same in all our experiments and is given as follows. We include seven RBPTS lines in our problem instances, namely, M1A, M2, M4, M5, M6, F1, and Marmaray (MA) (Figure 5.1a). Considered lines consist of 111 stations: 49 in the European and 62 in the Asian side. We select 40 of those stations as candidate locations for PT connections: 23 in the European and 17 in the Asian side (Figure 5.1b).

Figure 5.1: RBPTS lines and candidate locations for PT connections.

As for the satellites, we choose 110 points, which can ideally serve as ASP locations (e.g. popular points of interest such as big supermarkets, shopping malls, universities, dormitories, schools, or some other central locations in neighborhoods). As depicted in Figure 5.2, 65 of the 110 satellites reside in the European side, while the remaining 45 are in the Asian side. Ultimately, our instances consist of 150 vertices, 40 being candidate locations for PT connections and 110 being satellites. The distances between the vertices are obtained via *OpenRouteService API* and incorporated in our graph. We are interested in seeing the effects of number of open PT connections, therefore we consider the cases with $k \in \{6, 8, 10, 12, 14\}$.

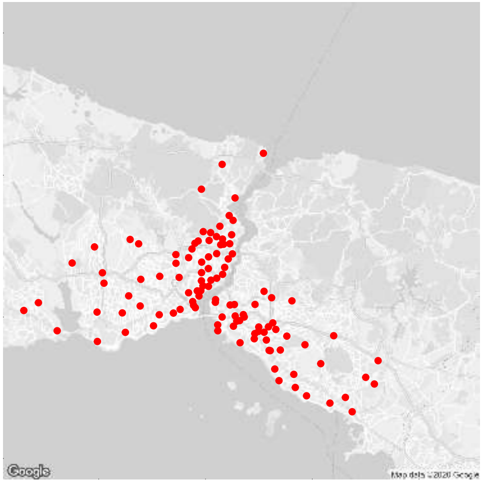


Figure 5.2: Satellite locations.

Investigating the costs of different on-demand transportation services offered in Istanbul, we consider a fixed cost of 10 Turkish Liras (TRY) and a per-kilometer cost of 4 TRY for BT. We note that the per-kilometer cost also incorporates the hourly compensation paid to the drivers. The maximum BT range is designated as 100 kilometers, in line with the approximate ranges given in the related literature for the e-motorcycles (Matsuda, 2014). Considering the nature of in-city use, the speed of BT vehicles is determined as 45 kilometers per hour. For the cycle durations (BT touring times) we consider the following five levels in our experiments $t_{\max} \in \{60, 65, 80, 90\}$ minutes.

The set of scenarios we use in our instances are built as follows. To build a specific scenario ω, we iterate over all pairs of candidate locations for PT connections and satellites, namely $\langle h, s \rangle$, $\forall h \in H, \forall s \in S$. For each pair $\langle h, s \rangle$, we draw a random number from a Bernoulli distribution with parameter q. If the drawn number is 1, we say opening a PT connection in h enables CS for satellite s in scenario ω. The parameter q is determined by the distance between s and h, denoted by $d(s, h)$, as well as a CS parameter $\zeta > 0$ as shown in (5.1). To see the effects of the different CS availability levels within the system, we consider different CS multiplier values as $\zeta \in \{0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$. Note that we use $p^{\omega_1} = p^{\omega_2}$ for any $\omega_1, \omega_2 \in \Omega$.

$$q = \begin{cases} \min\{0.9\zeta, 1.0\}, \ d(s,h) \leq 5 \\ \min\{0.8\zeta, 1.0\}, \ 5 < d(s,h) \leq 10 \\ \min\{0.7\zeta, 1.0\}, \ 10 < d(s,h) \leq 15 \\ \min\{0.6\zeta, 1.0\},, \ 15 < d(s,h) \leq 20 \\ 0.0, \ \text{otherwise} \end{cases} \tag{5.1}$$

5.2 Experiment Design

To investigate the main system dynamics and set a baseline to derive different experiment groups, we first define a base case setting (BC). Consequently, we create four different experiment groups, each focusing on different aspects of the system by perturbing the values of one parameter from the BC value. The BC is characterized by the parameter values $|\Omega| = 30, t_{\max} = 60$ minutes, $k = 8$ and $\zeta = 1$. Our experiment design consists of the following four groups of experiments.

- In the first group of experiments (E1), we vary the number of scenarios in the model between 10 and 80. By perturbing this parameter, we would especially like to gain insights into how our proposed solution algorithm handles cases with large numbers of scenarios. Similarly, we wish to find out how many scenarios one should consider in $\mathcal{P}$ to ensure that the resulting solution value closely approximates the expected BT costs, hence the design decisions are properly taken considering uncertainties in the CS delivery capacity. In particular, E1 consists of the problem instances with $|\Omega| \in \{10, 20, 30, 40, 50, 60, 70, 80\}$.

- Experiments of the second group (E2) focus on the changes in the distribution cycle duration $t_{\max}$. In particular, we aim to see the impact of the BT tour times on the system performance to assess the viability of the suggested delivery model to support express shipment demands. Experiments in this group consider the following four tour times: $t_{\max} \in \{60, 65, 80, 90\}$ minutes.

- The third group of experiments (E3) is built to measure the effects of the number of PT connections on the system performance. The sensitivity of the system to the number of PT connections is important as construction costs associated with the connections can be significant. Furthermore, we would like to see if a widespread CS coverage can be ensured by a limited number of PT connections, which imply less interference with the regular PT operations and better user experience for the PT riders. In this regard, the experiments within the group are comprised of the problem instances with as $k \in \{6, 8, 10, 12, 14\}$.

- Experiments in the fourth group (E4) consider different CS multipliers to assess the system's sensitivity against the level of CS participation/coverage. By increasing ζ, we simulate a more readily-available CS capacity for our system. Similarly, by decreasing ζ, we try to see if such a system will be cost-efficient or even feasible with some scarce supply on the CS side. To answer these questions in detail, the six experiments of this group is comprised of the instances with $\zeta \in \{0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3\}$.

Experiments in all groups are solved with plain BP as well as DB to investigate whether DB can provide any significant computational advantages. Both of the algorithms were implemented in JAVA programming language. All experiments were run on a 64-bit Windows 7 workstation with an i7-7600U processor at 2.80 GHz and 16 GB of RAM. We used CPLEX 12.10 for solving LPs.

5.3 Experiment Results

Detailed solutions statistics for all our experiments are reported in Table D.1, in Appendix D. We analyze our experiment results in two parts. In the first part, we focus on the computational issues, and in the second part, our focus is on the managerial and operational insights.

5.3.1 Computational performance analysis

The first question we want to answer in this section is how strong the lower-bound one can obtain by disregarding the routing variables' integrality requirements in $\mathcal{P}$. As discussed earlier in Section 4.1.2, to simplify the branch-and-price implementation, we solve $\mathcal{P}$ disregarding the integrality requirements for the routing variables. If the relaxed problem's solution has a fractional value for any of the routing variables, we then apply a simple heuristic approach (as explained in Section 4.1.2) to obtain a feasible solution. Figure 5.3 shows the resulting optimality gaps, calculated as the percentage difference between the optimal solutions of the relaxation (lower-bound) and the solutions obtained after the heuristic approach (upper-bound) if needed. As we see from these results, for 10 instances out of the total 21, the optimality gap is zero (i.e., the optimal solution of the relaxed problem already has binary solutions for the routing variables), and the largest optimality gap is less than 0.5% (average disregarding the zero values is 0.13%). As we can infer from these results, for the problem instances we consider in this study, the lower-bound provided by the linear relaxation of $\mathcal{P}$ without the integrality requirements is quite tight.

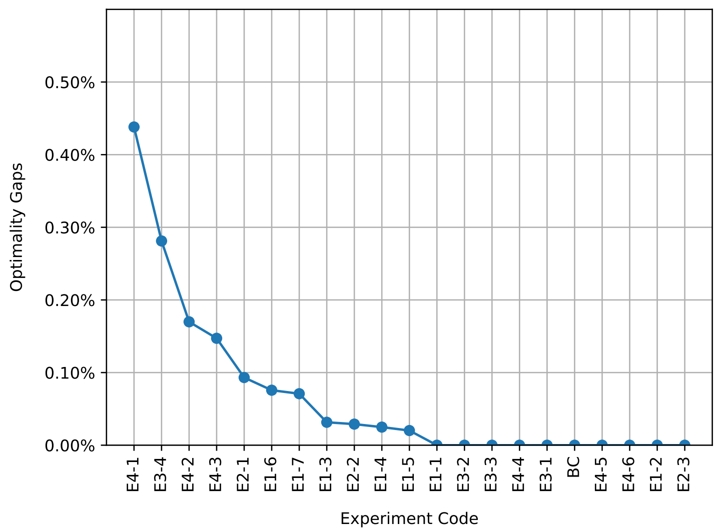


Figure 5.3: Optimality gaps for the problem instances.

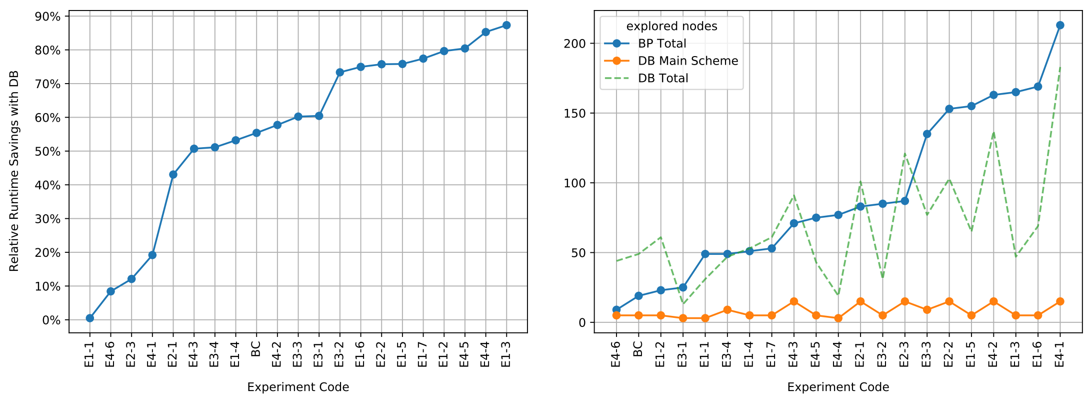
Secondly, we analyze the balance between the computational tractability and capability of the stochastic program $\mathcal{P}$ to adequately capture the uncertainties in the demand and crowd provided delivery capacity. In that regard, critical solution statistics for the E1 experiments are presented in Table 5.1. Columns *BP.time* and *DB.time* indicate the run time of the BP and DB algorithms to find the optimal solutions. The column *sol.val* indicates the optimal solution value, and the optimal PT connections to open are shown in the column *PT.con*. We test the performance of the optimal design decisions for each problem instance against a *test set (TE)* of 100 new scenarios (different than those contained in the scenario sets of the problem instances) and report the average BT cost in the column *exp.cost*. E1 results clearly show the critical importance of considering an "adequate" number of scenarios in the model to find a network design that minimizes the expected operational cost of the system as desired. As we see in Table 5.1, the resulting delivery network topology strongly depends on the scenarios considered in the model. Failing to capture the impact of demand and capacity uncertainties by considering a small number of scenarios can result in a sub-optimal network design with much higher operational costs. For example, we see that the resulting expected operational cost for the network design with $|\Omega| = 10$ is more than 12% higher than the expected cost of the network design with $|\Omega| = 80$. We also see that that the optimal solution values for the problems with 10 and 20 scenarios significantly underestimates the expected operational costs (i.e., *sol.val* is substantially less than *exp.cost*), indicating that the uncertainties in demand and capacity are not adequately captured in the resulting models. We also observe that beyond a threshold (i.e., $|\Omega| = 30$), increasing the number of scenarios in the model does not have a significant impact on the network design and resulting expected cost while the computational effort required to solve the problem increases significantly. Therefore, it is important to observe the balance between the solution quality and the computational tractability while deciding the scenario set to consider in the model. In that regard, these results also imply that using 30 scenarios as a base case parameter in our experiments is indeed a reasonable approach.

Finally, we analyze the computational advantages of DB over BP. We present our findings in two graphs. Figure 5.4a shows the relative percent runtime savings

Table 5.1: Solution statistics for E1 experiments.

ID	$\|\Omega\|$	*BP.time*	*DB.time*	*sol.val*	*exp.cost*	*PT.con*
E1-1	10	654.70	177.48	330.36	426.44	0, 5, 14, 28, 38, 50, 85, 94
E1-2	20	871.86	651.53	338.18	378.12	0, 5, 14, 28, 38, 50, 85, 94
BC	30	2858.39	1275.00	378.36	378.12	0, 3, 14, 28, 38, 50, 86, 94
E1-3	40	10427.67	1692.38	376.36	378.28	0, 3, 14, 28, 38, 50, 84, 94
E1-4	50	13348.43	2605.49	381.62	378.28	0, 3, 14, 28, 38, 50, 84, 94
E1-5	60	16359.74	3955.77	392.45	378.28	0, 3, 14, 28, 38, 50, 84, 94
E1-6	70	21427.03	5365.68	401.95	378.12	0, 3, 14, 28, 38, 50, 86, 94
E1-7	80	32556.96	7356.17	397.04	378.12	0, 3, 14, 28, 38, 50, 86, 94

obtained by DB when compared to BP. We see that, for the problem instances we consider in this study, DB can offer up to 87% reduction in the solution time, and on the average DB finds the optimal solution 56.3% faster than BP. Figure 5.4b, which displays the number of explored nodes both for BP and DB algorithms, reveals important insights into the computational advantages of the DB approach.



(a) (b)

Figure 5.4: Comparing DB and BP algorithms' performances.

The nodes stored in the *stack* structure of Algorithm 3 are given in the graph as nodes explored by the *DB Main Scheme*. This quantity does not include the nodes explored in the solving processes of the decomposed smaller problems (i.e., black boxes). On the other hand, we report another metric on the same graph with dashed lines, namely *DB total* which counts every explored node whether it is encountered in the main scheme or in either one of the smaller problems. We see that the main branch-and-bound tree of DB is much smaller than that of BP. However, the same conclusion does not always hold when we compare the metrics of *BP Total* and *DB Total*. Despite that, even for the experiments in which DB explores more nodes than BP, the runtime of DB still proves advantageous. Therefore, we infer that the improvement in the runtime does not directly depend on the number of nodes explored but on the node processing speed. DB algorithm speeds up the processing of a tree node by decomposing a bigger problem into two more tractable smaller problems that also allows concurrent solving.

It is also worthwhile to note here that DB makes the solution process of PPs much faster. That is because decomposed PPs are easier to solve, and they allow parallel processing. In addition to that, the algorithm can actually decrease the number of times PPs need to be solved in the majority of the instances as depicted in Figure 5.5. Even if the number of PP solvings is more in DB than BP in some instances, this is compensated by the easier and decomposed nature of the PP to give DB a clear advantage over BP in total solving time of PPs.

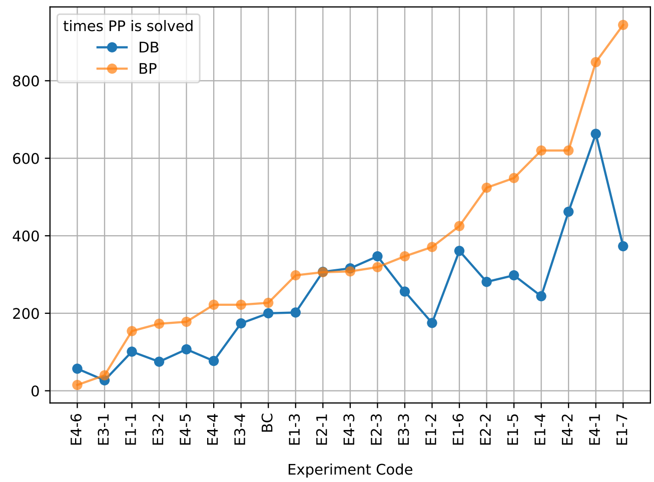


Figure 5.5: Number of PPs solved in DB and BP algorithms.

5.3.2 Managerial insights

We first focus on the BC solution to gain insights into the structures of the optimal solutions to PCBP and their managerial implications. PT opening decisions associated with the BC solution are presented in Figure 5.6. In the figure, the sizes of the PT connections are drawn proportional to the number of BT tours that originate or terminate at that PT connection. Similarly, the sizes of the satellites are drawn proportional to the number of times they have been visited by BT vehicles. This optimal solution consists of 112 BT tours in total (used in 30 scenarios).

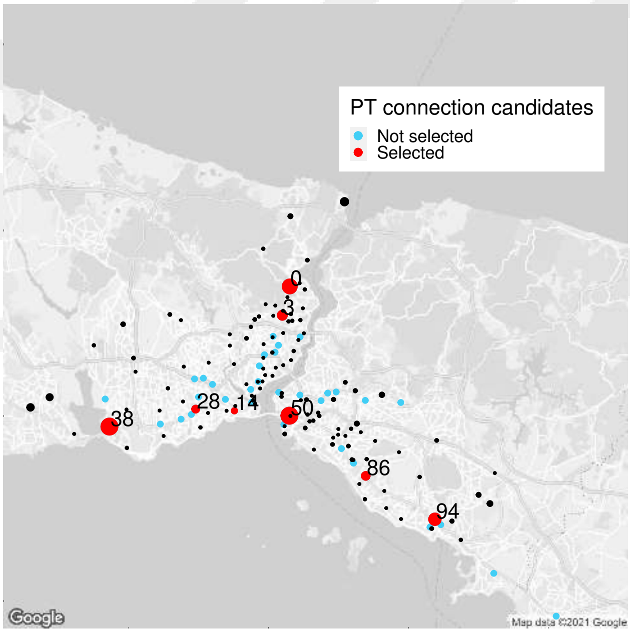
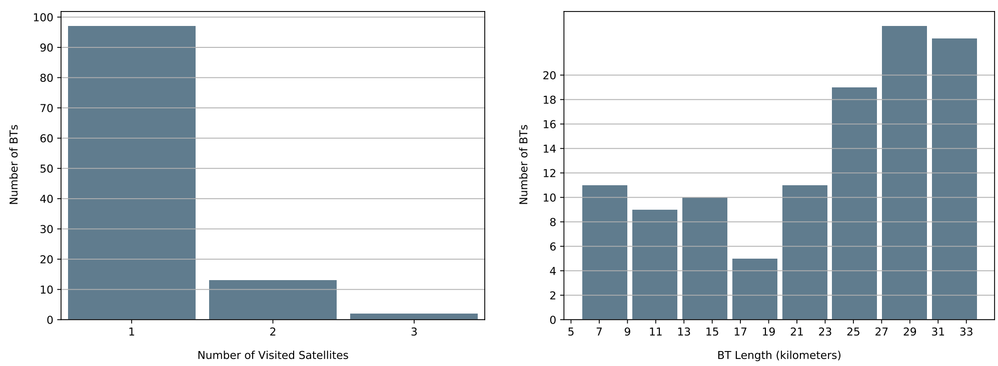


Figure 5.6: Opened hubs as per the solution of BC. Satellites are shown with smaller black dots. Numbers next to PT connections indicate the ID numbers.

As we see in Figure 5.7a, many of these tours visit only one satellite (97 tours out of 112), while the maximum number of visited satellites is three (there are two such BT tours). However, as we see in Figure 5.7b the tour lengths (hence durations) are not quite small. To the contrary, almost two thirds of the BT tours that are used in the optimal solution are longer than 20 km. As we discuss in more detail when we analyze the results of the E3 experiments, these results imply that, by strategically locating eight PT connections in the service area, it is possible to reduce the distances between satellites and the PT connections enough to cover most of the satellites by CS, most of the time. Hence, BT tours would be necessary (occasionally) to serve those satellites at the far ends of the service area. As a result, in most cases, the satellites to visit by BT are not only far from PT connections, but also located in places with less satellite densities, which we can see in Figure 5.6. Therefore, most of the BT tours visit only one or two satellites, not mainly because of the tour time/length limits. As an important managerial insight, according to these results, we can infer that the longer tour time limits may not help much with reducing the cost of BT usage, but will certainly increasing the delivery lead times. It is also worthwhile to note that, although different origin and destination PT connections are allowed for the tours, only five BTs in this solution end the tour in a different PT connection than the starting one. We also observe that the number of the BT tours that start or end at a PT connection varies much (size of the PT connections in Figure 5.6 varies a lot) showing that some of the PT connections are chosen to extend CS coverage, some are chosen to reduce the BT distances or establish interconnectivity of PT lines. For example, the Yenikapi PT connection (ID number 14) enables only one BT tour yet is selected by the model because this PT connection is required for the connectivity of several PT lines, as discussed with Figure 1.1.

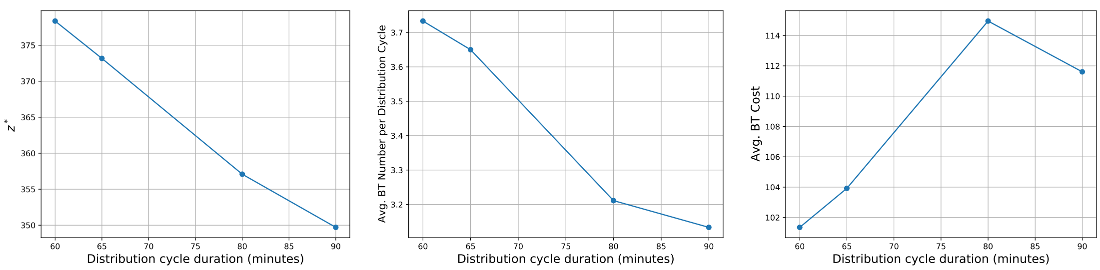


(a) (b)

Figure 5.7: Details on the BC solution.

The critical information about the E2 results is presented in Figure 5.8. The figure depicts the relations between the allowed BT touring time and statistics such as the solution value (z^*), average BT number in a distribution cycle, and the average cost of a BT tour. We observe in Figure 5.8a that when the system allows longer touring times for BT vehicles, overall costs associated with the routes in a distribution cycle decrease. Such a result is directly related to the number of BTs required to fulfill demands. When the touring time is relaxed, fewer BTs can meet the same demand as can be inferred from Figure 5.8b. Figure 5.8c suggests a rather interesting pattern in average BT costs that can be explained as follows. As $t_{\max}$ increases, the system tries to avoid routes with fewer stops to decrease overall costs. Consequently, with the increasing touring times, the average cost of a utilized BT tour rises, while the total number of BT tours decreases. Until a threshold (i.e., $t_{max} = 80$), a large number of relatively short BT tours with mostly single satellite visits are eliminated from the optimal solution (to save fixed costs); hence the average cost of a utilized BT tour rises although overall BT cost keeps decreasing.

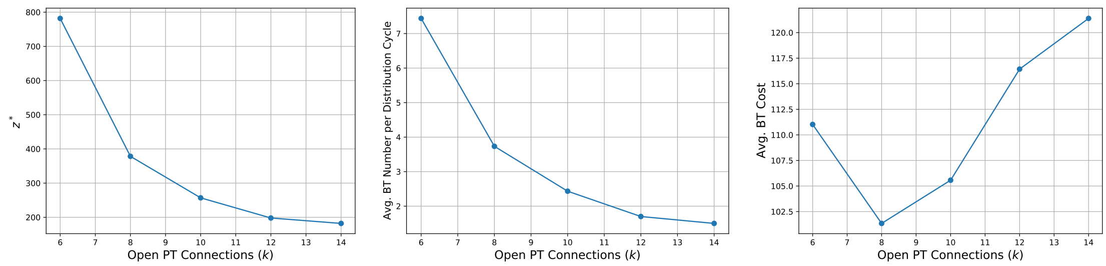
Another interesting result we observe from E2 results is that increasing allowed touring time from $t_{\max} = 60$ to $t_{\max} = 90$ reduces the cost by only around 6%. This implies that, as long as PT network transfer costs are negligible, compared to the BT costs, sparing some of the cycle time for the repositing of the packages in the PT network and thus giving less time for BT tours can be a good strategy, as giving less time to the tours of BT vehicles without any significant increase in BT costs.



(a) (b) (c)

Figure 5.8: Solution statistics for E2 experiments.

The critical information about the E3 results is given in Figure 5.9, which presents the same solution statistics as in Figure 5.8. As expected, Figure 5.9a shows that increasing the number of PT connections decreases the expected costs of BT tours in a distribution cycle, as opening more PT connections increase the CS coverage (reduces the need for BT). Figure 5.9b also confirms this reasoning by depicting a monotonously decreasing average number of BTs per distribution cycle which in turn implies that most of the demand can be met by CS supply. However, Figure 5.9c presents an intriguing trend on average BT costs. One might expect that as the number of PT connections increases the average distances that the BT tours cover would decrease; hence, the average BT tour costs would be less. However, a deeper thinking about the system dynamics reveals that the increasing k means that satellites previously traversed by low cost BTs now can be served by the extended CS coverage. Therefore the need for lower cost BTs became obsolete whereas the satellites that are far away from the PT lines are still visited by BTs with higher variable costs. This, in turn, leads to higher average BT costs.

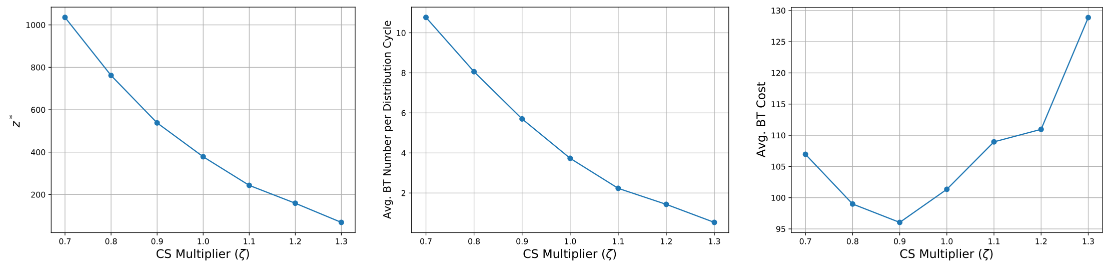


(a) (b) (c)

Figure 5.9: Solution statistics for E3 experiments.

The rate of decline in the expected BT costs by the increasing number of PT connections we see in Figure 5.9a provides an important managerial insight. It shows that the platform should carefully choose the number of PT connections to operate considering the trade-off between the associated set-up and operational costs. As we see in Figure 5.9a, up to a certain threshold (i.e., up to 8-10 PT connections) increasing the number of PT connections provides significant savings in the operational costs. Beyond this threshold, the contributions of additional PT connections may not worth the required set-up costs. As an interesting insight, we see that it is not required to have a large number of PT connections to obtain the benefits of an extensive CS coverage. With a relatively small number of PT connections (i.e., $k = 8$ or $k = 10$) it is possible to obtain a robust CS coverage that significantly reduces the need for costly BT operations. Here we want to emphasize that having a network design with fewer PT connections is not only desirable due to less set-up costs. It is also important to impose a minimal burden on the regular PT operations. In that regard, improving CS coverage by means other than opening more PT connections should be considered carefully, as we analyze next.

Solutions of E4 experiments are depicted in Figure 5.10 which shows unequivocally clear trends about the advantages of increasing the CS capacity in our system. Specifically, Figure 5.10a indicates that even a 10% increase in CS multiplier ζ results in a reduction of 35% in overall costs. The number of required BTs per distribution cycle also decreases exponentially as CS becomes more prevalent, as shown in Figure 5.10b. The varying average BT cost in a distribution cycle given in Figure 5.10c can be interpreted as similar to the cases in E2 and E3. That is, an increasing CS multiplier only leaves the satellites that are far from the PT lines as inaccessible by CS. On the other hand, an increasing ζ means more of the near satellites will be reachable by CS. Since those far-away satellites have more costly corresponding BTs because of the per-kilometer cost, the average BT cost increases as ζ increases from 0.9. On the other hand, we also observe an increase in average BT cost as ζ decreases from 0.9. That is because a low CS availability adds a significant number of new members to the set of satellites that need to be served by CS. As BT tours are limited in duration, the rented BT vehicles will have to visit the satellites in multiple BT tours in a given DC, which increases average BT costs by increasing total fixed costs.



(a) (b) (c)

Figure 5.10: Solution statistics for E4 experiments.

Results of this experiment group provide especially significant insights for the network management highlighting the sensitivity the system has to CS supply. We see that increasing ζ from 0.7 to 1.3, the expected cost of a distribution cycle (z^*) drops by 94%, implying that establishing and retaining a solid CS supply is of at most importance for the viability of the suggested system. As such, network management should take any measures necessary to increase CS availability as much as possible. In that regard, choosing the system's PT base as metro networks is a good measure in this regard as they are suitable for enabling CS operations (Dietmann, 2020).

Chapter 6

CONCLUSION

In this study, we present and investigate a novel public transport-based crowd-shipping system in which intracity package transfer demands are met with collaborative efforts of public transport vehicles, occasional carriers (i.e., the supply side of CS), and –where needed– backup transfers. Focusing on the strategic level design problem for such a delivery system, we introduce PCBP to the literature, which aims to find optimal locations for PT connections to minimize the expected costs associated with using BTs, as the most expensive mode of transfer. We consider a scenario-based approach and develop a two-stage stochastic programming formulation to deal with the uncertainty in delivery capacity on top of the uncertainty associated with the demand. To solve this challenging problem for realistic size problem instances, we develop a branch-and-price algorithm. Taking advantage of a specific structure one can observe in many real-world applications, we explore the idea of boosting the computational performance of the basic BP algorithm by using a novel branching framework: decomposition branching. Presenting the first example of employing this powerful idea in a branch-and-price algorithm, we show that incorporating DB can indeed provide substantial computational benefits.

Besides attesting to the efficacy of our solution approaches, our extensive numerical study reveals significant managerial insights that shed light on how to design such a network to maximize its efficiency. One such critical insight we derive from those experiments is the crucial importance of a strong CS delivery capacity for the viability of the last-mile delivery system we propose. Regarding this inference, we suggest that the decision-makers should consider strategies to increase CS delivery capacity. These can be in the form of tangible incentives for crowd-shippers (i.e., financial compensation, reward programs) and advertisement and highlighting the system's environmental and social goods such as reduced carbon emissions.

In this initial study, we focus on the network design problem assuming that the network management has already evaluated the customer expectations, the performance of the PT network (including but not limited to the metrics of time between consecutive PT vehicle arrivals and the longest trip time in the network) and the potential of CS delivery capacity together to decide the frequencies of distribution cycles and the length of the BT tours. Due to their critical impact on the operational costs and the service quality of the suggested delivery system, further research effort is required to help decision-makers determine distribution cycle frequencies and the BT tour times. Moreover, problems related to the real-time management of the suggested delivery network present interesting new research directions we plan to follow in future studies.

Evident from our results, it is also crucial for the network management to correctly model the behavior of the potential crowd-shippers based on the incentives to determine the best policies to ensure a solid CS capacity with the minimum cost. Investigating this critical problem for effective real-world implementation of the suggested delivery system opens up important new directions for future studies.

We want to conclude by underlining that our envisioned system uses the PT network as a backbone of a CS-based system. This approach has a conceptual significance as it provides an example for developing innovative solutions for transportation problems that are inspired by successful applications in communication networks, similar to the current developments in the literature on the physical internet. We hope our study's promising results would motivate new research in this direction to develop innovative solutions to address the chronic (and growing) transportation-related problems such as pollution/carbon emissions, congestion, noise, and accidents.

BIBLIOGRAPHY

Alumur, S. and Kara, B. Y. (2008). Network hub location problems the state of the art. *European journal of operational research*, 190(1):1–21.

Alumur, S. A., Campbell, J. F., Contreras, I., Kara, B. Y., Marianov, V., and O'Kelly, M. E. (2020). Perspectives on modeling hub location problems. *European Journal of Operational Research*.

Alumur, S. A., Kara, B. Y., and Karasan, O. E. (2009). The design of single allocation incomplete hub networks. *Transportation Research Part B Methodological*, 43(10):936–951.

Archetti, C., Savelsbergh, M., and Speranza, M. G. (2016). The vehicle routing problem with occasional drivers. *European Journal of Operational Research*, 254(2):472–480.

Arslan, A. M., Agatz, N., Kroon, L., and Zuidwijk, R. (2019). Crowdsourced deliverytextemdasha dynamic pickup and delivery problem with ad hoc drivers. *Transportation Science*, 53(1):222–235.

Arvidsson, N., Givoni, M., and Woxenius, J. (2016). Exploring Last Mile Synergies in Passenger and Freight Transport. *Built Environment*, 42(4):523–538.

Barnhart, C., Hane, C. A., and Vance, P. H. (2000). Using branch-and-price-and-cut to solve origin-destination integer multicommodity flow problems. *Operations Research*, 48(2):318–326.

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W., and Vance, P. H. (1998). Branch-and-price column generation for solving huge integer programs. *Operations research*, 46(3):316–329.

Bi, K., Yang, M., Zahid, L., and Zhou, X. (2020). A new solution for city distribution to achieve environmental benefits within the trend of green logistics: A case study in china. *Sustainability*, 12(20):8312.

Briffaz, M. and Darvey, C. (2016). Crowd-shipping in geneva exploratory and descriptive study of crowd-shipping. mathesis, University of Gothenburg.

Bruzzone, F., Cavallaro, F., and Nocera, S. (2021). The integration of passenger and freight transport for first-last mile operations. *Transport Policy*, 100:31–48.

Campbell, J. F. and O'Kelly, M. E. (2012). Twenty-five years of hub location research. *Transportation Science*, 46(2):153–169.

Cardenas, I. D., Dewulf, W., Beckers, J., Smet, C., and Vanelslander, T. (2017). The e-commerce parcel delivery market and the implications of home b2c deliveries vs pick-up points. *International journal of transport economics*, (XLIV).

Caspersen, E. and Navrud, S. (2021). The sharing economy and consumer preferences for environmentally sustainable last mile deliveries. *Transportation Research Part D: Transport and Environment*, 95:102863.

Cochrane, K., Saxe, S., Roorda, M. J., and Shalaby, A. (2016). Moving freight on public transit best practices, challenges, and opportunities. *International Journal of Sustainable Transportation*, 11(2):120–132.

Contreras, I., Cordeau, J.-F., and Laporte, G. (2011). Stochastic uncapacitated hub location. *European Journal of Operational Research*, 212(3):518–528.

Dahle, L., Andersson, H., Christiansen, M., and Speranza, M. G. (2019). The pickup and delivery problem with time windows and occasional drivers. *Computers and Operations Research*, 109:122–133.

De Langhe, K., Meersman, H., Sys, C., de Voorde, E. V., and Vanelslander, T. (2019). How to make urban freight transport by tram successful? *Journal of Shipping and Trade*, 4(1).

Dietmann, K. (2020). Crowdshipping the last-mile delivery - an empirical investigation into the crowd's willingness to participate as crowdshipping drivers. mathesis, HEC-Ecole de gestion de l'Université de Liège.

Ermagun, A., Shamshiripour, A., and Stathopoulos, A. (2019). Performance analysis of crowd-shipping in urban and suburban areas. *Transportation*, 47(4):1955–1985.

Farahani, R. Z., Hekmatfar, M., Arabani, A. B., and Nikbakhsh, E. (2013). Hub location problems a review of models, classification, solution techniques, and applications. *Computers and Industrial Engineering*, 64(4):1096–1109.

Feng, B., Li, Y., and Shen, H. (2015). Tying mechanism for airlines' air cargo capacity allocation. *European Journal of Operational Research*, 244(1):322–330.

Gatta, V., Marcucci, E., Nigro, M., and Serafini, S. (2019). Sustainable urban freight transport adopting public transport-based crowdshipping for b2c deliveries. *European Transport Research Review*, 11(1).

Ghilas, V., Demir, E., and Van Woensel, T. (2013). Integrating passenger and freight transportation model formulation and insights. *Proceedings of the 2013 Beta Working Papers; Technische Universiteit Eindhoven Eindhoven, The Netherlands*, 441:1–23.

Ghilas, V., Demir, E., and Woensel, T. V. (2016). The pickup and delivery problem with time windows and scheduled lines. *INFOR Information Systems and Operational Research*, 54(2):147–167.

Hasiak, S. (2019). Access mobility to local railway stations current travel practices and forecast. *Cybergeo*.

Iwan, S., Kijewska, K., and Lemke, J. (2016). Analysis of Parcel Lockers' Efficiency as the Last Mile Delivery Solution – The Results of the Research in Poland. *Transportation Research Procedia*, 12:644–655.

Jaller, M., Pineda, L., and Ambrose, H. (2018). Evaluating the Use of Zero-Emission Vehicles in Last Mile Deliveries. *ITS Reports*, 2017(33).

Kikuta, J., Ito, T., Tomiyama, I., Yamamoto, S., and Yamada, T. (2012). New subway-integrated city logistics szystem. *Procedia - Social and Behavioral Sciences*, 39:476–489.

Li, Z., Shalaby, A., Roorda, M. J., and Mao, B. (2021). Urban rail service design for collaborative passenger and freight transport. *Transportation Research Part E: Logistics and Transportation Review*, 147:102205.

Macrina, G., Pugliese, L. D. P., Guerriero, F., and Lagan'a, D. (2017). The vehicle routing problem with occasional drivers and time windows. In *Springer Proceedings in Mathematics and Statistics*, pages 577–587. Springer International Publishing.

Macrina, G., Pugliese, L. D. P., Guerriero, F., and Laporte, G. (2020). Crowd-shipping with time windows and transshipment nodes. *Computers & Operations Research*, 113:104806.

Masson, R., Trentini, A., Lehu'ed'e, F., Malh'en'e, N., P'eton, O., and Tlahig, H. (2017). Optimization of a city logistics transportation system with mixed passengers and goods. *EURO Journal on Transportation and Logistics*, 6(1):81–109.

Matsuda, Y. (2014). A study of electric motorcycle. In *SAE Technical Paper Series*. SAE International.

Millward, H., Spinney, J., and Scott, D. (2013). Active-transport walking behavior destinations, durations, distances. *Journal of Transport Geography*, 28:101–110.

Mladenow, A., Bauer, C., and Strauss, C. (2016). "Crowd logistics" the contribution of social crowds in logistics activities. *International Journal of Web Information Systems*, 12(3):379–396.

Mourad, A., Puchinger, J., and Van Woensel, T. (2019). Combining people and freight flows using a scheduled transportation line with stochastic passenger demands.

Munoz, P. and Cohen, B. (2015). Sharing Cities and Sustainable Consumption and Production Towards an Integrated Framework. *Journal of Cleaner Production*, 134:87–97.

Murray, C. C. and Chu, A. G. (2015). The flying sidekick traveling salesman problem Optimization of drone-assisted parcel delivery. *Transportation Research Part C Emerging Technologies*, 54:86–109.

Nocera, S., Pungillo, G., and Bruzzone, F. (2020). How to evaluate and plan the freight-passengers first-last mile. *Transport Policy*.

O'Kelly, M. E. and Miller, H. J. (1994). The hub network design problem. *Journal of Transport Geography*, 2(1):31–40.

Paloheimo, H., Lettenmeier, M., and Waris, H. (2016). Transport reduction by crowdsourced deliveries – a library case in Finland. *Journal of Cleaner Production*, 132:240–251.

Punel, A. and Stathopoulos, A. (2017). Modeling the acceptability of crowdsourced goods deliveries Role of context and experience effects. *Transportation Research Part E Logistics and Transportation Review*, 105:18–38.

Quak, H., Nesterova, N., van Rooijen, T., and Dong, Y. (2016). Zero Emission City Logistics Current Practices in Freight Electromobility and Feasibility in the Near Future. *Transportation Research Procedia*, 14:1506–1515.

Rougè, J.-F. and Montreuil, B. (2014). Crowdsourcing delivery new interconnected business models to reinvent delivery. In *1st international physical internet conference*, pages 1–19.

Savelsbergh, M. and Van Woensel, T. (2016). 50th anniversary invited article—city logistics challenges and opportunities. *Transportation Science*, 50(2):579–590.

Serafini, S., Nigro, M., Gatta, V., and Marcucci, E. (2018). Sustainable crowdshipping using public transport a case study evaluation in rome. *Transportation Research Procedia*, 30:101–110.

Stiglic, M., Agatz, N., Savelsbergh, M., and Gradisar, M. (2018). Enhancing urban mobility integrating ride-sharing and public transit. *Computers and Operations Research*, 90:12–21.

Turkish Statistical Institute (2019).

van Duin, J. H. R., Wiegmans, B. W., van Arem, B., and van Amstel, Y. (2020). From home delivery to parcel lockers a case study in Amsterdam. *Transportation Research Procedia*, 46:37–44.

Van Duin, R., Wiegmans, B., Tavasszy, L., Hendriks, B., and He, Y. (2019). Evaluating new participative city logistics concepts the case of cargo hitching. *Transportation Research Procedia*, 39:565–575.

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

Wong, W. H., Zhang, A., Hui, Y. V., and Leung, L. C. (2009). Optimal baggage-limit policy: Airline passenger and cargo allocation. *Transportation Science*, 43(3):355–369.

Yaman, H., Kara, B. Y., and Tansel, B. c. (2007). The latest arrival hub location problem for cargo delivery systems with stopovers. *Transportation Research Part B Methodological*, 41(8):906–919.

Yildiz, B. (2021). Express package routing problem with occasional couriers. *Transportation Research Part C Emerging Technologies*, 123:102994.

Yildiz, B., Boland, N., and Savelsbergh, M. (2018). Decomposition branching for mixed integer programming. Technical report.

Yildiz, B. and Karacsan, O. E. (2015). Regenerator location problem and survivable extensions a hub covering location perspective. *Transportation Research Part B Methodological*, 71:32–55.

Yildiz, B. and Karacsan, O. E. (2017). Regenerator location problem in flexible optical networks. *Operations Research*, 65(3):595–620.

Yildiz, B., Karacsan, O. E., and Yaman, H. (2018). Branch-and-price approaches for the network design problem with relays. *Computers and Operations Research*, 92:155–169.

Yildiz, B. and Savelsbergh, M. (2019). Service and capacity planning in crowd-sourced delivery. *Transportation Research Part C Emerging Technologies*, 100:177–199.

Yıldız, B. (2021). Package routing problem with registered couriers and stochastic demand. *Transportation Research Part E Logistics and Transportation Review*, 147.

Zhou, F. and Zhang, J. (2020). Freight transport mode based on public transport taking parcel delivery by subway as an example. In *ICTE 2019*, pages 745–754. American Society of Civil Engineers Reston, VA.

Appendix A

BUILDING THE CONNECTIVITY SETS

Connectivity sets $O(h, m), h, m \in H$ can be obtained easily with a polynomial-time algorithm whose pseudo-code is given as Algorithm 2.

Algorithm 2: Simple Algorithm to Build the Sets $O(h, m)$

Input: H, PT lines

Output: $O(h, m),\ \forall h \in H, \forall m \in H,\ h \neq m$

A // Creating an empty set to store $O(h,m)$ sets for different h and m's
$\hat{G} \leftarrow$ empty undirected graph
$\hat{G}$.addNodes(H)
foreach $h \in H$ **do**
    foreach $m \in H$ **do**
        if $h \neq m$ *AND h and m are in the same PT line* **then**
            $\hat{G}.addEdge(\{h, m\})$
foreach $h \in H$ **do**
    foreach $m \in H$ **do**
        if $h \neq m$ **then**
            $N \leftarrow$ depthFirst(h, m) // Suppose this calls depth first search algorithm that returns the path including origin and destination
            $O(h, m) \leftarrow N \setminus \{h, m\}$
            Add $O(h, m)$ to A
return A

Recall that H denotes the candidate locations for PT connections, which can also be interpreted as PT stations. We create a new network, namely $\hat{G} = (H, E)$, whose node set consists of the candidate locations for connections. Initially, we set $E = \emptyset$. We iterate through every node pair (h_1, h_2) in $\hat{G}$ and create a direct edge $\{h_1, h_2\}$ with weight 1 between them if and only if h_1 and h_2 are in the same PT line. Then, to compute $O(h, m)$ for any given $h, m \in H, h \neq m$ we run a basic graph search algorithm in graph $\hat{G}$ (e.g., depth-first search). Suppose the path between h and m goes through the nodes N. Then we set $O(h, m) = N \setminus \{h, m\}$.

Appendix B

SOLVING PCBP IN A PT NETWORK WITH CYCLES

Here we present a modification of $\mathcal{P}$ to relax our assumption that the PT network does not have any cycles. For a PT connection candidate $h \in H$ we define $C(h)$ as the set of PT connections $\bar{h} \in H \setminus \{h\}$ that are directly connected to h in the PT network (i.e., there is a PT line that visits both points). The main idea is to consider a connectivity network (known as communication graph in the literature (Yildiz and Karacsan, 2015)) which is denoted by $\bar{G} = (H, E)$ with the node set H and edge set $E = \{(h, \bar{h}) : h \in H, \bar{h} \in C(h)\}$. Considering the binary variables $w_e, e \in E$ we define the cut formulation $\mathcal{P}_{cut}$ as follows.

$$(\mathcal{P}_{cut}) \quad \min \; z = \sum_{\omega \in \Omega} \sum_{\pi \in P} p^{\omega} c_{\pi} v_{\pi}^{\omega} \tag{B.1}$$

$$\text{s.t.} \quad \sum_{\pi \in P(s)} v_{\pi}^{\omega} + \sum_{h \in H(\omega, s)} x_h \geq 1 \qquad \forall \omega \in \Omega, \; \forall s \in S \tag{B.2}$$

$$\sum_{\pi \in \hat{P}(s,h)} v_{\pi}^{\omega} - x_h \leq 0 \qquad \forall \omega \in \Omega, \; \forall h \in H, \; \forall s \in S \tag{B.3}$$

$$\sum_{\pi \in \hat{\hat{P}}(s,h)} v_{\pi}^{\omega} - x_h \leq 0 \qquad \forall \omega \in \Omega, \; \forall h \in H, \; \forall s \in S \tag{B.4}$$

$$\sum_{h \in H} x_h = k \tag{B.5}$$

$$w_{hm} \leq x_h \qquad \forall \{h, m\} \in E \tag{B.6}$$

$$w_{hm} \leq x_m \qquad \forall \{h, m\} \in E \tag{B.7}$$

$$w_{hm} \geq x_h + x_m - 1 \qquad \forall \{h, m\} \in E \tag{B.8}$$

$$\sum_{e \in \delta(U)} w_e \geq x_h + x_m - 1 \qquad \forall U \subset V, h \in U, m \notin U \tag{B.9}$$

$$x_h \in \{0, 1\} \qquad \forall h \in H \tag{B.10}$$

$$v_{\pi}^{\omega} \in \{0, 1\} \qquad \forall \omega \in \Omega, \; \forall \pi \in P \tag{B.11}$$

Constraints (B.6)-(B.8) are simply the linear representation of the relation $w_hm = x_hx_m$, meaning that for e_{hm} to be used for connectivity connections must be opened at locations h and m. The main connectivity constraint is given in (B.9), enforcing that the PT connection network (a sub-network in $\bar{G}$) is connected. Since the number of constraints in (B.9) grows exponentially with the size of the number of candidate PT connections, we suggest solving it with a branch-and-price-and-cut (BPC) procedure by iteratively adding the inequalities (B.9), as needed. In that regard, we want to note that the separation problem can be efficiently solved with a basic graph search as successfully implemented in the literature (Yildiz and Karacsan, 2015). Since these connectivity-cuts only contains design variables, the BP algorithm we develop to solve $\mathcal{P}$ can be extended in a straightforward way to solve $\mathcal{P}_{cut}$ with a BPC algorithm.

Appendix C

PSEUDOCODE FOR DB ALGORITHM

The pseudocode for the DB algorithm is outlined in the next page.

Algorithm 3: DB Algorithm

Input: $\mathcal{P}$
Output: z^*
1 $z^* \leftarrow \infty$
2 *stack* $\leftarrow \mathcal{P}^0_{LP}$
3 **while** *stack is not empty* **do**
4   *current_node* ← the top node of *stack*
5   Remove *current_node* from *stack*
6   **do**
7   Solve *current_node* to obtain its optimal solution $(\bar{x}, \bar{v})$, optimal value $\bar{z}$ and the dual variable values δ, ϕ and ψ
8         $V \leftarrow SA(\delta, \phi, \psi, H, S, \omega, m)$
9   Add routes of V to the model
10   **while** V *is not empty*
11   **if** *current_node is infeasible* **then**
12   Prune by infeasibility and go to line 4
13   **if** $\bar{z} \geq z^*$ **then**
14   Prune by bound and go to line 4
15   **if** $\hat{k}_i = \sum_{h \in H(i)} \bar{x}_h$ *is integer for* $i = 1, 2, \ldots n$ **then**
16   *branching* ← false
17   **foreach** $i = 1, 2, \ldots, n$ **do**
18   **if** $z^*_{i\hat{k}_i} > \bar{z}_{i\hat{k}_i}$ **then**
19   *first_child* ← BP node with the cut $\sum_{h \in H(i)} x_h \leq \hat{k}_i \;\wedge\; z_i \geq z^*_{i\hat{k}_i}$
20   *second_child* ← BP node with the cut $\sum_{h \in H(i)} x_h \geq \hat{k}_i + 1$
21   Add *first_child* and *second_child* to *stack*
22   *branching* ← true
23   Break
24   **if** *not branching* **then**
25   Prune *current_node* after forming an integer solution by combining sub-problem solutions
26             $\hat{z}$ ← the solution obtained by combining subproblem solutions
27   **if** $\hat{z} < z^*$ **then**
28                 $z^* \leftarrow \hat{z}$
29   **else**
30   **foreach** $i = 1, 2, \ldots, n$ **do**
31   **if** $\hat{k}_i$ *is not integer* **then**
32   *first_child* ← BP node with the cut $\sum_{h \in H(i)} x_h \leq \lfloor \hat{k}_i \rfloor \;\wedge\; \bar{z}_i \geq z^*_{i\lfloor \hat{k}_i \rfloor}$
33   *second_child* ← BP node with the cut $\sum_{h \in H(i)} x_h \geq \lceil \hat{k}_i \rceil$
34   Add *first_child* and *second_child* to *stack*
35   Break
36 return z^*

Appendix D

DETAILED SOLUTIONS

In Table D.1, we report the algorithm used to solve the problem instances (alg), the experiment ID (ID), the parameter and its value that differentiates the instance from BC (*param*), the run time in seconds (*time*), the number of *total* branch-and-bound nodes explored (#*node*), the number of BT tours utilized in the best integer solution (#*endbt*), the solution value of $\mathcal{P}^0_{LP}$ (LP), the lower bound found by the algorithm (*lb*), the upper bound found as described in Section 4.1.2 (*ub*), the optimality gap (column headed *gap*), the average cost of a BT tour (*costavg*) and the average number of BT tours per distribution cycle (#*dcbt*).

alg	*ID*	*param*	*time*	*#node*	*#endbt*	*LP*	*lb*	*ub*	*gap*	*costavg*	*#dcbt*
DB	BC	-	1275.00	49	45471	292.62	378.36	378.36	0.00%	101.35	3.73
	E1-1	$\Omega = 10$	177.48	31	10985	248.86	330.36	330.36	0.00%	103.24	3.2
	E1-2	$\Omega = 20$	651.53	61	34736	248.25	338.18	338.18	0.00%	105.68	3.2
	E1-3	$\Omega = 40$	1692.38	47	66359	291.52	376.36	376.48	0.03%	101.86	3.78
	E1-4	$\Omega = 50$	2605.49	53	81953	292.93	381.62	381.72	0.02%	101.08	3.84
	E1-5	$\Omega = 60$	3955.77	65	116983	298.67	392.45	392.53	0.02%	101.16	3.93
	E1-6	$\Omega = 70$	5365.68	69	111455	297.29	401.95	402.26	0.08%	105.06	3.94
	E1-7	$\Omega = 80$	7356.17	61	133580	297.55	397.04	397.32	0.07%	104.59	3.91
	E2-1	$t = 65$	3321.35	101	105779	296.2	373.17	373.52	0.09%	103.92	3.65
	E2-2	$t = 80$	18755.81	103	366918	283.48	357.09	357.19	0.03%	114.96	3.21
	E2-3	$t = 90$	235906.99	121	7847912	272.52	349.7	349.7	0.00%	111.61	3.13
	E3-1	$k = 6$	206.74	13	1851	179.36	182.08	182.08	0.00%	121.39	1.5
	E3-2	$k = 10$	576.67	31	3116	190.39	197.94	197.94	0.00%	116.43	1.7
	E3-3	$k = 12$	1745.09	77	60880	218.91	256.89	256.89	0.00%	105.57	2.43
	E3-4	$k = 14$	1415.99	47	120737	556.41	781.77	783.98	0.28%	111.02	7.44
	E4-1	$\zeta = 0.7$	29956.46	183	658505	792.89	1035.36	1039.92	0.44%	106.95	10.77
	E4-2	$\zeta = 0.8$	7329.58	137	363482	596.67	761.64	762.94	0.17%	98.99	8.06
	E4-3	$\zeta = 0.9$	2739.66	91	199818	432.31	537.87	538.66	0.15%	96.04	5.7
	E4-4	$\zeta = 1.1$	339.59	19	9781	188.88	243.32	243.32	0.00%	108.95	2.23
	E4-5	$\zeta = 1.2$	338.48	43	13809	130.59	159.04	159.04	0.00%	110.95	1.43
	E4-6	$\zeta = 1.3$	158.45	44	2101	64.72	68.73	68.73	0.00%	128.87	0.53
BP	BC	-	2858.39	19	22695	292.62	378.36	378.36	0.00%	101.35	3.73
	E1-1	$\Omega = 10$	654.70	49	8917	248.86	330.36	330.36	0.00%	103.24	3.2
	E1-2	$\Omega = 20$	871.86	23	22695	248.25	338.18	338.18	0.00%	105.68	3.2
	E1-3	$\Omega = 40$	10427.67	165	46848	291.52	376.36	376.48	0.03%	101.86	3.78
	E1-4	$\Omega = 50$	13348.43	51	100847	292.93	381.62	381.72	0.02%	101.08	3.84
	E1-5	$\Omega = 60$	16359.74	155	122894	298.67	392.45	392.53	0.02%	101.16	3.93
	E1-6	$\Omega = 70$	21427.03	169	62489	297.29	401.95	402.26	0.08%	105.06	3.94
	E1-7	$\Omega = 80$	32556.96	53	187009	297.55	397.04	397.32	0.07%	104.59	3.91
	E2-1	$t = 65$	5832.28	83	36979	296.2	373.17	373.52	0.09%	103.92	3.65
	E2-2	$t = 80$	77316.69	153	197149	283.48	357.09	357.19	0.03%	114.96	3.21
	E2-3	$t = 90$	268393.95	87	7343328	272.52	349.7	349.7	0.00%	111.61	3.13
	E3-1	$k = 6$	522.50	25	1766	179.36	182.08	182.08	0.00%	121.39	1.5
	E3-2	$k = 10$	2163.69	85	3767	190.39	197.94	197.94	0.00%	116.43	1.7
	E3-3	$k = 12$	4386.55	135	20006	218.91	256.89	256.89	0.00%	105.57	2.43
	E3-4	$k = 14$	2896.67	49	82306	556.41	781.77	783.98	0.28%	111.02	7.44
	E4-1	$\zeta = 0.7$	37070.24	213	332044	792.89	1035.36	1039.92	0.44%	106.95	10.77
	E4-2	$\zeta = 0.8$	17343.93	163	155420	596.67	761.64	762.94	0.17%	98.99	8.06
	E4-3	$\zeta = 0.9$	5557.48	71	68533	432.31	537.87	538.66	0.15%	96.04	5.7
	E4-4	$\zeta = 1.1$	2307.97	77	25781	188.88	243.32	243.32	0.00%	108.95	2.23
	E4-5	$\zeta = 1.2$	1728.94	75	7447	130.59	159.04	159.04	0.00%	110.95	1.43
	E4-6	$\zeta = 1.3$	173.03	9	1451	64.72	68.73	68.73	0.00%	128.87	0.53

Table D.1: Details on solutions obtained by DB and BP algorithms.