УНИВЕРЗИТЕТ У БАЊОЈ ЛУЦИ
UNIVERSITY OF BANJA LUKA
ПРИРОДНО-МАТЕМАТИЧКИ ФАКУЛТЕТ
FACULTY OF NATURAL SCIENCES AND MATHEMATICS

Dimitrije D. Čvokić

EXISTENCE AND SOLUTION METHODS FOR STACKELBERG EQUILIBRIA IN THE $(r|p)$ HUB–CENTROID PROBLEM UNDER THE PRICE WAR

A thesis submitted in partial fulfillment of the requirements for the degree of *Doctor of Philosophy*
in
Mathematics

Banja Luka, 2021.

УНИВЕРЗИТЕТ У БАЊОЈ ЛУЦИ
UNIVERSITY OF BANJA LUKA
ПРИРОДНО-МАТЕМАТИЧКИ ФАКУЛТЕТ
FACULTY OF NATURAL SCIENCES AND MATHEMATICS

Димитрије Д. Чвокић

ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ $(r|p)$ ХАБ–ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ

ДОКТОРСКА ДИСЕРТАЦИЈА

Бања Лука, 2021.

Acknowledgements

I can freely say that my Ph.D. studies were indeed a challenging period for me. An old saying says: ”No pain, no gain!” However, in addition to my efforts and struggles, this thesis could not have been performed and finished if not for the assistance, patience, and support of many individuals, to whom I am sincerely grateful.

Firstly, I would like to acknowledge my Ph.D. thesis supervisors, Professor Yuri A. Kochetov and Professor Aleksandar Savić, for their encouragement and guidance over the last few years. Their knowledge of contemporary trends in operations research and refined ability to distinguish important from unimportant has lead to the original research proposal – examining the Stackelberg competition of hub location problems that include pricing. They have helped and supported me through tough times over the course of this research, for which I owe them a special note of appreciation.

I am also grateful to Professor Aleksandr V. Plyasunov for the tremendous support he has–and continues–to provide, especially when it comes to the revision process that has lead to this document. His knowledge and understanding of the written word have allowed me to fully express the concepts behind this thesis.

My appreciation is also extended to my supervisors on undergraduate and master studies, Professor dr Milan Jovanović and Professor dr Đorđe Dugošija, who served as a voice of quiet wisdom in matters ranging from the most fundamental aspects of mathematics to the paths that my career has eventually taken.

When it comes to the financial support, I gratefully acknowledge the funding received towards my Ph.D. from the Government of Russia and the Ministry of Scientific and Technological Development, Higher Education and Information Society of the Government of Srpska Republic. I also appreciate that the University of Banja Luka and Faculty of Natural Sciences and Mathematics have exempted me from tuition fees.

Finally, I would like to express my deepest gratefulness to my parents Dušan and Stoja Čvokić, my sisters Milica and Jelena, as I firmly believe that without their love, support and understanding I could never succeed in ending this academic journey.

Захвалница

Слободно могу рећи да је докторски студиј био један поприлично изазован период за мене, у правом смислу те ријечи. Давно се у народном уму родио усклик: „Без муке нема науке!“. Ипак, без обзира на све уложене напоре и пратећа унутрашња трвења, предочени истраживачки подухват не би могао бити спроведен и завршен без помоћи, стрпљења, и подршке многих мени драгих особа, те им се овим путем свима захваљујем.

На првом мјесту, желио бих да истакнем улогу својих ментора, професора др Јурија А. Кочетова и професора др Александра Савића, као научно–истраживачких руководилаца, али и због њиховог постојаног бодрења током протеклих неколико година. Њихово познавање савремених токова у развоју операционих истраживања и профињено разликовање битног од небитног, је у мом случају резултовало проучавањем Штаклберговог надметања размијештања разводних тачака које у обзир узима и утицај цијена. Професори су ми били незамјењива подршка у неким поприлично незгодним периодима током самог студија, на чему им се од свег срца захваљујем.

Посебно бих се захвалио и професору др Александру В. Пљасунову на његовој несебичној помоћи, како током самог истраживања, тако и током писања и израде ове дисертације. Његово разумијевање писане ријечи ми је омогућило да у потпуности изразим идеје и концепте у позадини свог досадашњег рада.

Захвалност дугујем и својим менторима на дипломском и мастер студију, професору др Милану Јовановићу и професору др Ђорђу Дугошији, који су служили као тихи глас искуства и разборитости по многим питањима, почевши од формирања личног погледа на математику, до путева којим је моја каријера пошла.

Када је ријеч о материјалној подршци, срдачно се захваљујем на финансијској помоћи коју сам током мојих докторских студија и за израду дисертације добио од Владе Русије и Министарства за научнотехнолошки развој, високо образовање и информационо друштво Владе Републике Српске. Такође, посебно цијеним излазак у сусрет руководстава Универзитета у Бањој Луци и Природно–математичког факултета да се ослободим трошкова студирања.

На крају, желим да изразим посебну захвалност мојим родитељима Душану и Стоји Чвокић, мојим сестрама Милици и Јелени, без чије љубави, подршке, стрпљења, и разумијевања не бих привео крају ово своје академско путовање.

Supervisor: Dr Yuri A. Kochetov, Full Professor at Novosibirsk State University (Novosibirsk, Russia)
Co–supervisor: Dr Aleksandar Savić, Associate Professor at University of Belgrade (Belgrade, Serbia)

Title: Existence and solution methods for Stackelberg equilibria in the $(r|p)$ hub–centroid problem under the price war

Abstract

A new logistic problem is introduced in which two competing transportation companies want to enter the market, while they are aware of each other. The objective for both is to maximize their profits by finding the best hub and spoke networks and price structures. The first one that enters the market, a leader, wants to establish p hubs. The second one, a follower, is planning to locate r of them. After setting their networks, it is expected that the competing companies will engage in the price war, assuming the customers patronize routes by their corresponding prices. The logistic regression based model is used to estimate how the market share will be split among the competitors. This problem is called the $(r|p)$ hub–centroid problem under the price war. A mathematical model for finding its Stackelberg equilibrium is provided as a bi–level non–linear mixed–integer program. It is shown that in this setting for each origin–destination pair a unique finite Bertrand–Nash price equilibrium exists. Based on this result, the existence of Stackelberg equilibria is shown, new equations for the best response pricing are proposed, and the characterization of optimal routes under the price equilibrium is given. The problem itself is also addressed from the computational complexity aspect. It is shown that this bi–level optimization problem is NP–hard. Therefore, the use of metaheuristics is a natural choice for solving real–size instances. An alternating heuristic and a variable neighborhood search algorithms were designed as a solution approaches for the leader. When it comes to the follower, we have provided a way to reformulate the corresponding problem linearly. The computational experiments were conducted using CAB instances. The results are thoroughly discussed, highlighting the effects of different parameters and providing some interesting managerial insights. Finally, some possible future research directions are outlined.

Keywords: hub location, pricing, logit model, Stackelberg equilibria, Bertrand–Nash price equilibrium, reformulation, complexity, alternating heuristic, variable neighborhood search

Scientific Field: Natural Sciences
Scientific Area: Mathematics
CERIF Classification Code: P 160

Ментор: проф. др Јуриј А. Кочетов, редовни професор на Националном истраживачком државном универзитету у Новосибирску (Новосибирск, Русија)
Коментор: проф. др Александар Савић, ванредни професор на Универзитету у Београду (Београд, Србија)

Наслов: Постојање Штаклбергових еквилибријума у проблему $(r|p)$ хаб–центроида са цјеновним надметањем и алгоритми за њихово проналажење

Резиме

Уводи се нови логистички проблем у којем двије конкурентне транспортне компаније–такмаци једна за другом улазе на тржиште, при чему се сматра да су обје у потпуности и савршено обавијештене. Циљ и једне и друге је максимизација профита образовањем транспортне мреже с разводним тачакама и пратећом цјеновном структуром. Прва која улази на тржиште има намјеру да размјести p разводних тачака, док друга планира да постави њих r. Подразумијева се да корисници услуга бирају руте превасходно према њиховим цијенама, тј. према одговарајућем моделу подјеле потражње заснованом на логистичкој регресији. Очекује се да ће компаније, по формирању транспортних мрежа, кренути са цјеновним надметањем. Предочени сукоб интереса се назива проблемом $(r|p)$ хаб–центроида са цјеновним надметањем. Представљен је математички модел за проналажење одговарајућег Штаклберговог еквилибријума као нелинеарни мјешовито–цјелобројни математички програм у два нивоа. Показано је да у овој поставци постоји јединствен коначни Бертранд–Нешов цјеновни еквилибријум. На основу тог резултата доказано је постојање Штаклберговог еквилибријума, представљене су нове једначине за најбољу цјеновну реакцију, и дата је карактеризација оптималних рута. Такође, проблем је разматран из угла рачунске сложености. Пошто је показано да је NP–тежак, коришћење метахеуристика за његово рјешавање се наметнуло као природан избор. Конструисани су алгоритми на основу алтернацијске хеуристике и метода промјенљивих околина. Што се тиче друге компаније–такмаца, показано је како се њен модел може линеаризовати. Рачунски огледи су извршени над стандардним у литератури CAB-инстанцама. Резултати емпиријског исљеђивања су детаљно дискутовани, истакнути су утицаји различитих параметара, и издвојени неки интересантни увиди. На крају, размотрени су и неки од могућих будућих праваца за наставак истраживања.

Кључни појмови: размијештање разводних тачака, формирање цијена, логит-модел, Штаклбергов еквилибријум, Бертранд–Нешов цјеновни еквилибријум, рачунска сложеност, реформулација, алтернацијска хеуристика, метод промјенљивих околина

Научно поље: Природне науке
Научна област: Математика
Класификациони код према CERIF-у: P 160

Contents

List of Abbreviations

AH	alternating heuristic
AM	auxiliary model
API	application programming interface
BC	before Christ
BVNS	basic variable neighborhood search
BPNE	Bertrand–Nash price equilibrium
C	capacitated
CAB	Civil Aeronautics Board
CDF	cumulative distribution function
CPU	central processing unit
CIC	cyclic inter–hub connection
DCM	discrete choice model
DDR	double data rate
DM	decision maker
ECDF	empirical CDF
FOC	First Order Condition
GHz	gigahertz
GiB	gibibyte
GRASP	Greedy randomized search procedure
HCLP	hub–center location problem
HCovLP	hub covering location problem
HCP	hub–centroid problem
HCPuPW	hub–centroid problem under the price war
HLP	hub location problem
HMLP	hub–median location problem
HMPuPW	hub–medianoid problem under the price war
iid	independently and identically distributed
IM	inter–modal logistics

LS	local search
LPR	linear programming relaxation
MA	multiple allocation
MAD	mean absolute deviation
MC	maximum cover
Mod	modular arc costs
NP	non–deterministic polynomial–time
O–D	origin–destination
P	deterministic polynomial–time
PDF	probability density function
P	pricing
QM	queue estimation model
RAM	random–access memory
R	robustness
SA	single allocation
SC	Stackelberg competition
U	uncapacitated
VNS	variable neighborhood search
w.l.o.g.	with out loss of generality

List of Figures

List of Tables

List of Codes

List of Algorithms

The intentionally blank page paradox.

Chapter 1

Introduction

> *The chief forms of beauty are order and symmetry and definiteness, which the mathematical sciences demonstrate in a special degree.*
>
> Aristotle, *Metaphysics*

Optimality can be considered as a fundamental principle that defines the laws of physics, dominates in the etiology, or which orchestrates our social lives. Therefore, we can freely say that the optimization, among many other things, is a stepping stone of our civilization. The first known attempt to investigate optimality in mathematics is contributed to a Greek mathematician Euclid of Alexandria (325-265 BC). He was interested in geometry optimization, like the problems of minimal distance or the greatest area [47]. As might be expected, for a long period since then, mathematicians have not been aware of any general methods for finding optimal solutions. Only some special techniques were at their disposal. According to many historians of mathematics, a new period for optimization started in 1636, when Pierre de Fermat proposed in his paper [36] a general approach to compute local extreme points of a differentiable function by setting its derivative to zero. This idea lead to the now well–known first–order condition. The third period in optimization started not long ago with the discovery of linear programming (LP) by Leonid Kantorovich in 1939 [70, 41].

Today, the interest in mathematical optimization is constantly rising among the researches. All corners of our daily lives, economy, and science are intertwined with our intentions to make the best decisions, i.e., to find the optimal solutions. For example, cutting unnecessary costs or careful resource usage is directly affecting the company's profit. Formulating the appropriate mathematical model to represent the problem and designing an adequate solution approach is of enormous significance in the real–world, not just in academia.

When it comes to logistics, all its activities and fields require optimization, taking into account their specific features determined by the goals, resources, infrastructure, available strategies, market situation analysis, etc. This is the reason why even in the logistics each optimization problem is addressed and approached separately.

However, after obtaining the precise and well–designed mathematical model, the next step is to choose the appropriate solution approach. Contemporary optimization literature is very rich when it comes to different algorithms for optimization problems. In other words, the ground is very solid for designing new algorithms or adjusting the existing ones to find a high–quality solution in a reasonable amount of time.

The primary goal of this work is to provide more realistic model for a specific network design. The understanding of competition and pricing effects allow us to obtain a deeper managerial insights, provides basis for further refinements, and finally, enables more efficient and useful transportation services. In order to make all this, we need to provide adequate mathematical models, describe its properties, and to come up with a good solution approach.

Here, we will investigate mutual effect of hub location and pricing decisions in a competitive environment. The focus of research is on the existence of Stackelberg equilibria and design of solution approach for a newly introduced model.

The remainder of this chapter is structured as follows. Firstly, a small introduction to hub location problems is given, including the overview of current research trends in this field. After that, the general story of Stackelberg competition, $(r|p)$ hub–centroid problem, Bertrand competition, and logit model are presented.

1.1 Hub Location Problem

Following the global growth in trade volumes, a spoke–hub distribution paradigm has emerged, as a sort of antipode to the point-to-point networks. From a historic perspective, it seems that the paradigm was firstly used in the telecommunication industry. However, in logistical systems, the airline industry and postal companies are one of the main users of this concept. Delta Air Lines, an airline transportation company from the USA, pioneered the use of spoke–hub network model in 1955 [37]. Nevertheless, this system has become the norm in the airline industry after the Airline Deregulation Act introduced by the U.S. federal government in 1978 [8]. Simultaneously, the information technology sector adopted this paradigm under a dub name “star network” [10].

Under the spoke–hub paradigm traffic planners organize routes as series of “spokes” which connect outlying points, called origin–destination pairs (O–D pairs), through series of “hubs”. The set of such routes, which forms a connected graph, is denoted as a hub and spoke network in the operations research literature.

An immediate consequence of using this concept is a flow concentration on hub-to-hub

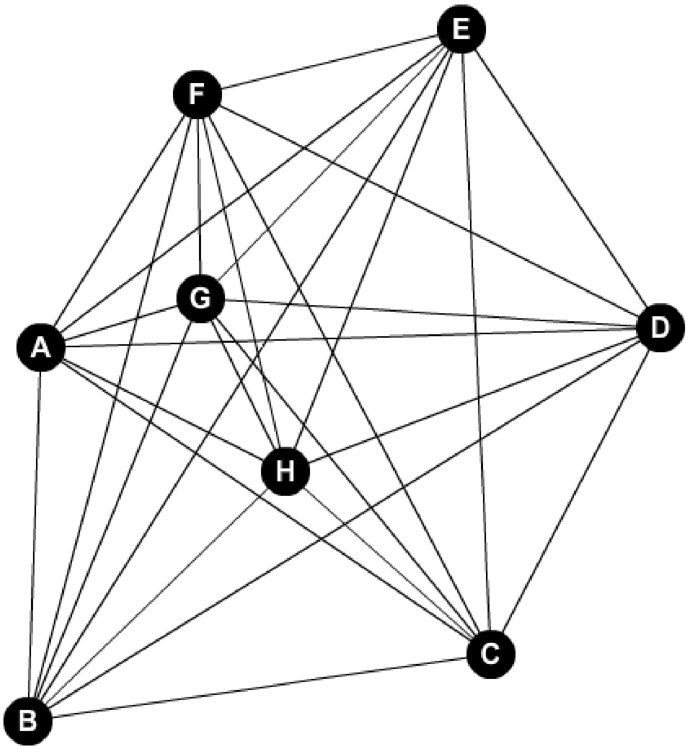


Figure 1.1: Point-to-point network of eight nodes. Each edge represents two connections of opposite directions.

connections, which has several advantages [110]:

- lower network installation costs;
- economies of scale on connections;
- economies of scale at hubs;
- economies of scope in the use of shared transshipment facilities.

For example, an interconnected network with k nodes and no hubs has $k(k-1)$ spokes. However, selecting one node to serve as a hub, i.e., to connect all other non–hub nodes with each other, will result in a need for only $2(k-1)$ spokes.

In Fig. 1.1, a point-to-point network is presented. It involves 56 direct connections, depicted with 28 edges in the interest of clarity, where an edge represents two connections of opposite directions. Each connection requires an infrastructure and vehicle service. On the other hand, a hub and spoke structure presented in Fig. 1.2, requires only 20 connections, represented as 10 edges. Two hubs are distinguished with red color, alongside with the inter–hub connection. The cost reduction is obvious. Furthermore, using larger and more efficient vehicles on inter–hub connections does not only reduce the transportation costs per unit of flow but is possibly more friendly for the environment.

However, it is worth noting that some potential disadvantages may also occur such as additional transshipment because fewer point-to-point services are offered. For some

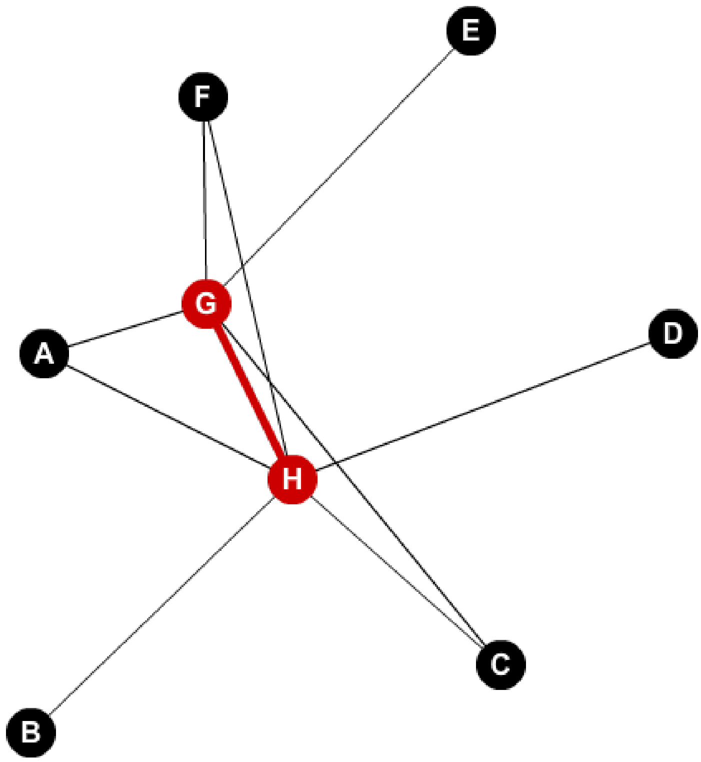


Figure 1.2: A hub and spoke network of eight nodes. Each edge represents two connections of opposite directions. Hubs and inter–hub connections are distinguished with red color.

connections, this may involve delays and potential congestions at hubs. As the demand and network load grow, more point-to-point services become feasible. Thus, hub and spoke networks are an intermediate stage in network development because the service preference remains direct connection.

In the location theory, problems concerned with these kinds of networks are called hub location problems (HLPs). Due to their high importance in academic and engineering fields, HLPs are widely studied by researchers [22, 48, 20].

The main goal is to seek out for the optimal locations of hubs and allocations of non–hub nodes to hubs concerning a given objective. Hubs serve as collection, consolidation, and dissemination points when routing the flow between O–D pairs. The location part is concerned with finding the most suitable position to establish a hub in the network architecture. On the other hand, the allocation subproblem seeks to explore the connection relationship in the network architecture.

The traditional classification of hub location problems is usually done according to the following criteria:

- solution domain: discrete or continuous;
- objective: maximization or minimization;
- a decision about the number of hubs: exogenous or endogenous;

- hub capacity: uncapacitated (unlimited) or capacitated (limited);
- hub location cost: fixed, variable, or no cost at all;
- inter–hub connectivity: total or partial.
- the allocation of a non–hub node to hub nodes: multiple or single;
- the cost of direct connection: fixed, variable, or no cost at all;
- problem type: median, center, covering, etc.;
- other restrictions.

Usually, it is assumed that the hubs are interconnected. Various mathematical models for hub location problems can be formulated by considering this classification.

In the research literature, the concept of concentrator in a network was firstly addressed by Goldman in his paper [54], published in 1969. The first HLP was introduced by O'Kelly in 1987 [103]. It was a so–called single allocation HLP. In 1994, Campbell has formulated a multiple allocation variant [19]. Soon, many researchers become interested in HLP. For example, Skorin–Kapov et al. have investigated, in [118], tight linear programming relaxations of uncapacitated p-hub median problems. Ebery et al., in 2000, considered and presented formulations and solution approaches for the capacitated multiple allocation hub location problem [44]. Mayer and Wagner have provided, in their paper [92] from 2002, a new and improved branch-and-bound procedure for the uncapacitated multiple allocation HLP. The interest in HLP is growing every year. A more detailed review of HLPs, their classification, and solution approaches can be found in [22, 48, 5, 82].

1.1.1 Two Classic Hub Location Problems

In the following lines, as an illustration, two classic hub location problems are going to be presented: p–hub median problem (p–HMLP) and p-hub center problem (p–HCLP).

p–Hub Median Location Problem

The formulation of this problem is similar to the p–median one, hence the name origin. Because every non–hub node could be allocated to one or more hubs this model is also named *the multiple allocation p–HLP*, to build a connection with the (single allocation) p–HLP of O'Kelly.

Formally, the problem is defined over a complete digraph $G = (N, A)$, where N is the non–empty set of nodes and A is the set of arcs. A hub can be established only at some node $k \in N$. The number of hubs p is determined exogenously. There are no capacity constraints imposed on hubs and there are no costs for locating hubs. The multiple allocation scheme involved in p–HMLP means that each non–hub node $i \in N$ can be allocated to several existing hubs. There are no costs for establishing connections on arcs. There exists a non–negative transportation cost per unit of flow along the arc $(i, k) \in A$, denoted as $c_{ik} \geq 0$. Usually, it is assumed that c_{ik} is obtained by a non–decreasing function of distance. The so–called fundamental approach in modeling the economies of scale is used by involving several standard assumptions from the hub location theory:

- the resulting hub backbone must be totally interconnected;
- transportation cost factors χ, α and δ for the flow consolidation in the collection (origin to hub), transfer between hubs, and distribution (hub to destination), are respectively appertain to;
- concatenation of arcs composes a route, in which hubs are located at the joints;
- at most two hubs are allowed to be on a single route, meaning that at most two stops are permitted.

The ordered quadruple (i, j, k, l) represents the route $i \to k \to l \to j$. The transportation cost $c_{ij,kl}$ for a route $i \to k \to l \to j$ is computed as $c_{ij,kl} = \chi c_{ik} + \alpha c_{kl} + \delta c_{lj}$. A non–negative demand $w_{ij} \geq 0$ for every O–D pair $(i, j) \in A$ is assumed to be perfectly inelastic (basically, it is an arc weight in G). The resulting optimal hub and spoke network must cover all nodes in G.

The non–negative real variable $x_{ij,kl}$ represents a fraction of flow going from i to j through hubs k and l. The binary decision variable y_k is taking value 1 if and only if k is chosen to be a hub, otherwise $y_k = 0$.

The p–HMLP can be represented as the following mixed–integer linear program [19]:

$$\min \sum_{i,j,k,l \in N} w_{ij} x_{ij,kl} c_{ij,kl} \tag{1.1}$$

$$\text{s.t.} \sum_{k,l \in N} x_{ij,kl} = 1, \quad \forall i, j \in N, \tag{1.2}$$

$$\sum_{k \in N} y_k = p, \tag{1.3}$$

$$x_{ij,kl} \leq y_k, \quad \forall i, j, k, l \in N, \tag{1.4}$$

$$x_{ij,kl} \leq y_l, \quad \forall i, j, k, l \in N, \tag{1.5}$$

$$x_{ij,kl} \geq 0, \quad \forall i, j, k, l \in N, \tag{1.6}$$

$$y_k \in \{0, 1\}, \quad \forall k \in N. \tag{1.7}$$

The objective function (1.1) sums the transportation costs for all established routes. Constraints (1.2) assure that for each O–D pair $(i, j) \in N$ every customer is going to be serviced. The number of hubs p is stipulated by constraint (1.3). Constraints (1.4) and (1.5) assure us that flows are routed via hubs. Variables' domains are given by (1.6) and (1.7).

Remark 1.1. In a way, an O–D pair in the p–HMLP is analogous to a demand point (node) in a p–median location problem.

p–Hub Center Location Problem

The p–HCLP was introduced by Campbell in 1994 [19]. The goal is to establish p hubs in a network, in order to minimize the maximum cost (or distance) between O–D pairs. We can find this problem in the transportation of decomposable or sensitive goods through a network, or when the delivery time frame is very tight [64].

In 2000, Kara and Tansel provided a linear formulation of Campbell's quadratic model and showed that its corresponding decision problem is NP–complete [74]. One of the latest and most promising linear formulations of the p–HCLP was provided in 2009 by Ernst et al. [46].

Formally, the setting for the p–HCLP is the same as for the p–HMLP. Here, we will provide a classic four–index non–linear mixed–integer formulation:

$$\min \ \max \sum_{k,l \in N} w_{ij} x_{ij,kl} c_{ij,kl} \tag{1.8}$$

$$\text{s.t.} \sum_{k,l \in N} x_{ij,kl} = 1, \quad \forall i, j \in N, \tag{1.9}$$

$$y_k = p, \tag{1.10}$$

$$x_{ij,kl} \leq y_k, \quad \forall i, j \in N, \tag{1.11}$$

$$x_{ij,kl} \leq y_l, \quad \forall i, j \in N, \tag{1.12}$$

$$x_{ij,kl} > 0, \quad \forall i, j, k, l \in N, \tag{1.13}$$

$$y_k \in \{0, 1\}, \quad \forall k \in N, \tag{1.14}$$

The objective function (1.8) is to minimize the maximal variable cost among established transportation routes. The constraint (1.9) assure that for each O–D pair $(i, j) \in N$ every customer is going to be serviced. The number of hubs p is indicated with (1.10). Constraints (1.11) and (1.12) assure us that flows are routed via hubs. Finally, the constraints (1.13) and (1.14) define the variables' domains.

1.1.2 Current Trends in HLP Research

Roughly, current trends in HLP research can be classified into model extensions and solution approach developments. For the first class, the classic HLP models are extended in order to grasp more reality mainly by:

- accounting capacities [45, 25];
- considering different flow cost models [76, 96];
- examining partial interconnection between hubs [18, 111, 109];
- considering location with reliability [75, 2];
- taking into the account pricing [86, 28, 29];
- examining competition and collaboration [112, 86, 28];
- considering stochastic and robust variants [9, 117, 29].

Table 1.1: Prominent studies with exact solution approaches for different HLPs.

Year	Problem code	Solution approach	# of nodes	Reference
1994	U-MA-p–HCLP/HCovLP	Integer linear programming	25	[19]
1996	U-MA/SA-p–HMLP	Linear programming relaxation	25	[118]
2000	U-MA-p–HCLP	Mixed–integer linear programming	25	[74]
2001	U-SA-p–HMLP	Mixed–integer linear programing	200	[43]
2005	C-SA-p–HLP	Integer linear programming	40	[91]
2008	U-MA-HMLP	Benders decomposition	200	[34]
2009	U-SA-p–HMLP	Integer linear programming	81	[6]
2009	U-SA-p–HMLP	Integer linear programming	550	[104]
2011	U-MA-HMLP	Enhanced Benders decomposition	500	[24]
2011	C-SA-p-HLP	Mixed–integer linear programming	140	[83]
2012	C-SA-p–HMLP	Generalized Benders decomposition	100	[35]
2014	U-MA/SA-HLP+R	Second–order cone programming	25	[117]
2016	U-MA-HLP+SC	Enumeration–based algorithm	25	[117]
2017	U-MA-HLP+CIC	Branch and cut with mixed–dicut inequalities	100	[26]
2017	U-SA-HLP+MAC	Branch and bound with Lagrangian relaxation	75	[121]
2020	U-SA-HLP+R+P	Second-order cone programming	25	[29]

When it comes to the second class of HLP research trends, a large variety of algorithms have been proposed to solve different types of HLPs during the past three decades. Table 1.1 presents some of the prominent studies (and definitely not all) in which the exact optimization methods have been used. The first column refers to the year of publishing, in the second one a code name for the addressed problem is presented, the third columns gives the name of solution approach (it could be many, but some of the most interesting ones are chosen), the fourth tells us the maximum instance size that was empirically investigated, and the last column provides a literature reference. The code names are composed using the following abbreviations in a given order:

- capacity: uncapacitated (U) or capacitated (C);
- non–hub node allocation: multiple allocation (MA) or single allocation (SA);

- type: (p-)HLP—as a general abbreviation, hub median location problem (HMLP), hub center location problem (HCLP), hub cover location problem (HCovLP), and HLP+ for the extended models of HLP.

Among many extensions that were addressed in the scientific literature, Table 1.1 refers only to a few of them as its purpose is to, basically, illustrate the shift in research trends. Particularly, extensions concerning robustness (R), Stackelberg competition (SC), cyclic inter–hub connection (CIC), modular arc costs (MACs), and pricing (P) are referred.

Table 1.2: Prominent studies with heuristic solution approaches for different HLPs.

Year	Problem code	Solution approach	# of nodes	Reference
1986	U-SA-p–HMLP	Nearest neighborhood	∞ (plane)	[102]
1992	U-MA-p–HMLP	Tabu search and GRASP	25	[77]
1996	U-SA-p-HLP	Simulated annealing	20	[119]
1999	U-MA-p–HMLP	Greedy algorithm	50	[115]
2000	C-MA-p-HLP	Shortest paths based heuristic	200	[44]
2002	U-MA-p–HMLP	Tabu search and GRASP	50	[78]
2003	U-MA-p–HMLP	Tabu search	50	[89]
2005	U-SA-p–HLP	Genetic algorithm	200	[123]
2006	C-MA-p–HCovLP	Greedy algorithm	400	[15]
2007	U-SA-p–HMLP	Genetic algorithm	200	[81]
2008	U-SA-p–HMLP	Lagrangian relaxation and local search	81	[126]
2009	U-SA-p-HLP	Genetic algorithm	200	[49]
2009	U-MA-p-HCovLP	Evolutionary approach	82	[107]
2009	U-SA-p–HCLP	Heuristic based on aggregation	1000	[53]
2009	U-SA-p–HCLP	Ant colony optimization	1000	[94]
2010	U-SA-p–HMLP	General variable neighborhood search	1000	[67]
2012	U-MA-p–HMLP+IM	Tabu search	25	[68]
2013	U-SA-HLP	Memetic algorithm	400	[90]
2017	U-MA-p–HCLP+R	Hybrid metaheuristic algorithm	900	[97]
2017	U-SA-p–HCLP	General variable neighborhood search	423	[16]
2018	U-SA-HCovLP+QM	Particle swarm optimization	30	[63]
2020	U-SA-HMLP+R+P	2–phase matheuristic	25	[29]

This table is not exhaustive at all. It is dominantly based on [48], where an interested reader can find references to many important and interesting studies, regarding HLP.

Although we can apply exact algorithms, they are usually used to solve problems with a smaller number of nodes. Empirical studies show that lager instances need to be solved by heuristic procedures. While large–size instances can be somewhat dealt with specialized exact methods (e.g., Benders decomposition and branch-and-price methods), the development of heuristic approaches (specialized ones, metaheuristics, or matheuristics) has helped many real–world applications, in which optimal or near–optimal solutions can be found in a very short execution time. Moreover, in the most of HLP studies usually the (meta)heuristic algorithms were applied for the empirical investigation. Some prominent research papers concerning heuristic solution approaches applied to different HLPs are presented in Table 1.2. The organization of Table 1.2 is the same as in the previous one. Also, extensions like inter–modal logistics (IM), queue estimation model (QM), are pointed out.

It is worth to note that sometimes even heuristics might not be able to properly address the large–size cases. For example, the very large–scale systems which demand a large number of investments. In some cases, even 1% gaps are not tolerable. In other words, there is a lack of appropriate solution approaches for solving situations of this kind [48].

1.2 Stackelberg Competition

Heinrich von Stackelberg (1905–1946) proposed in his book "Marktform und Gleichgwich" [120], published in 1934, a new model of market economy based on the following assumptions:

- it is more reasonable to expect that companies will enter the market in a sequential fashion, rather than simultaneously;
- strategic decisions that company makes are usually hard to change and easily observable;
- it is natural to investigate the possibility of first move advantage in cases of excess capacities.

Using the Game Theory terms, the players of this 2–stage game are called leader and follower. The leader moves first by choosing her strategy, knowing *ex ante* that the follower observes her decision. After making a move, it can not be revoked, i.e., the leader is under the power of commitment. The follower is not committed to a future non–Stackelberg leader's decision and the leader is aware of this. Each player intends to maximize their respective utility function payoffs. All strategy profiles are the common knowledge. In other words, the complete and perfect information is assumed.

Remark 1.2. The assumption of perfect information is important, as otherwise the game could be reduced to Cournot duopoly.

Remark 1.3. In the literature, for the leader and follower, usually pronouns "she" and "he" are used, respectively.

Let us consider the following strategic game

$$\Gamma = \langle P, S, u \rangle, \tag{1.15}$$

where:

- $P = \{L, F\}$ represents the set of players in which L stands for the leader and F stands for the follower;

- S_L and S_F are the corresponding non–empty (usually finite) strategy sets of leader and follower, respectively;
- $S = S_L \times S_F$ represents the set of strategy profiles;
- $u_L : S \to \mathbb{R}$ and $u_F : S \to \mathbb{R}$ are finite utility functions for the leader and follower, respectively;
- $u = (u_L, u_F)$ represents the game utility function.

Remark 1.4. For two functions $f : X \to Y$ and $g : X \to Z$ the notation (f, g) stands for the function $h : X \to Y \times Z$, characterized by $\pi_1(f, g) = f$ and $\pi_2(f, g) = g$, where $\pi_1 : Y \times Z \to Y$ and $\pi_2 : Y \times Z \to Z$ denote the projections.

In other words, for every leader's strategy s_L we have the set of follower's best responses

$$BR_F(s_L) = \arg\max_{s_F \in S_F} u_F(s_L, s_F). \tag{1.16}$$

On the basis of follower's best response, the solution of Stackelberg competition may be conceived in the set

$$\tilde{S} = \arg\max_{\substack{(s_L, s_F):\\ s_L \in S_L,\ s_F \in BR_F(s_L)}} u(s_L, s_F). \tag{1.17}$$

Definition 1.1. Any pure strategic profile in $\tilde{S}$ of game Γ is called *the unsafe Stackelberg equilibrium*.

Remark 1.5. The Stackelberg duopoly solutions can be perceived as subgame perfect Nash equilibria of a 2–stage game.

Unfortunately, the practical achievement of unsafe equilibria is not ensured if the follower's utility functions are surjective, i.e., when they can attain the same value for different particular strategies. For example we can have that $BR_F(s_L) = \{s_F^1, s_F^2\}$ and $u_L(s_L, s_F^1) < u_L(s_L, s_F^2)$. Both strategy profiles (s_L, s_F^1) and (s_L, s_F^2) belong to $\tilde{S}$, but obviously that the first one is not suitable as a solution for the Stackelberg game. Namely, in these situations we can say that the game is ill–posed. Therefore, in order to exceed such confusing and unwanted situation, the notion of a safe Stackelberg equilibrium is introduced, as in [125].

Firstly, the follower's behavior needs to be properly defined. For that, we consider two extreme cases:

(1) the leader has optimistic expectations about follower's behavior: among all strategies in the best response set, the follower chooses one which is the best for leader (altruistic/benevolent follower);

(2) the leader has pessimistic expectations about the follower's behavior: among all strategies in the best response set, the follower chooses one which is the worst for leader (selfish/malicious follower).

In the first case, for every leader's strategy s_L we consider the following subset of follower's best responses

$$\overline{BR}_F(s_L) = \arg\max_{s_F \in BR_F(s_L)} u_L(s_L, s_F). \tag{1.18}$$

For the second case, instead of $\arg\max$ function we need to use $\arg\min$.

Remark 1.6. The follower's behavior can be defined for the intermediate cases, too.

Now we can introduce a new solution set

$$\overline{S} = \arg\max_{\substack{(s_L, s_F): \\ s_L \in S_L,\ s_F \in \overline{BR}_F(s_L)}} u_L(s_L, s_F). \tag{1.19}$$

Definition 1.2. Any pure strategy profile in $\overline{S}$ of game Γ is called *the (safe) Stackelberg equilibrium* for a benevolent follower.

The safe Stackelberg equilibrium for the malicious follower is defined in a similar way.

Definition 1.3. For each Stackelberg equilibrium the corresponding leader's strategy is called the Stackelberg strategy.

Remark 1.7. In a simultaneous game, given in a matrix form, there may not exists a pure Nash equilibrium, but a Stackelberg strategy always exists.

Remark 1.8. When the sum of utility functions for the leader and follower is a constant, i.e., $u_L(s_L, s_F) + u_F(s_L, s_F) = const.$, for $s_F \in BR_F(s_L)$, then the follower's utility function is surjective.

Remark 1.9. For the set of best responses BR, we will denote the particular best response with lowercase letters br. For example, for a set $BR_F(s_L)$, the particular best response can be denoted as $br_F(s_L)$.

Example 1.1. The market price p depends on total quantity Q produced and it is estimated by function $P(Q)$, i.e., as

$$p = P(Q) = a - bQ, \tag{1.20}$$

where a and b are known coefficients. Denote the leader's and follower's production quantities as q_L and q_F, respectively. We have that $Q = q_L + q_F$.

The production cost function is given as $C(q) = cq$, and it is the same for both competitors.

The follower's profit, once the leader's quantity q_L is known can be computed as

$$u_F(q_L, q_F) = (a - b(q_L + q_F) - c)q_F. \tag{1.21}$$

The first–order condition $\frac{\partial u_F(q_L,q_F)}{\partial q_F} = 0$ yields $a - bq_L - 2bq_F - c = 0$, i.e., the best response of follower is given as

$$BR_F(q_L) = \frac{1}{2}\left(\frac{a-c}{b} - q_L\right). \tag{1.22}$$

We can see that his best response function is surjective. In other words, the behavior of follower does not need additional specification, which means that the set $BR_F(q_L)$ is singleton. Because of this, we can write $br_F(q_L) = \frac{1}{2}\left(\frac{a-c}{b} - q_L\right)$.

The leader's profit is given by

$$u_L(q_L, q_F) = (a - b(q_L + q_F) - c)q_L. \tag{1.23}$$

Substituting the follower's best response expression in the right hand side of last equation we obtain a new one–parameter function

$$\bar{u}_L(q_L) = \left(a - b\left(q_L + \frac{1}{2}\left(\frac{a-c}{b} - q_L\right)\right)\right) q_L - cq_L, \tag{1.24}$$

i.e.,

$$\bar{u}_L(q_L) = \frac{q_L}{2}\left(a - c - bq_L\right). \tag{1.25}$$

The first–order condition yields $a - bq_L - \frac{a-c}{2} - c = 0$, from which we get the Stackelberg strategy for the leader

$$q_L^* = \frac{a-c}{2b}. \tag{1.26}$$

Finally, substituting the right hand side of last equation with q_L in Eq. (1.21) we obtain

$$\left(\frac{a-c}{2b}, \frac{a-c}{4b}\right). \tag{1.27}$$

as the strategy profile which is a safe Stackelberg equilibrium with payoffs $\left(\frac{(a-c)^2}{8b}, \frac{a-c}{4b}\right)$. A graphical explanation of this example is given at Fig. 1.3. ▲

Remark 1.10. The follower could announce a deviation from the equilibrium by choosing a non–optimal strategy in order to reduce the game to Cournot. Under the assumption of rationality this is *a non–credible threat*, because when the leader has chosen her equilibrium strategy, any follower deviation would hurt him, too.

Remark 1.11. In our illustrative example, the leader has the first move advantage, but in general that does not need to be the case.

Remark 1.12. The Stackelberg competition can be viewed as a dynamic variant of Cournot duopoly. The best response in Cournot duopoly is the same as the follower's best response in Stackelberg model.

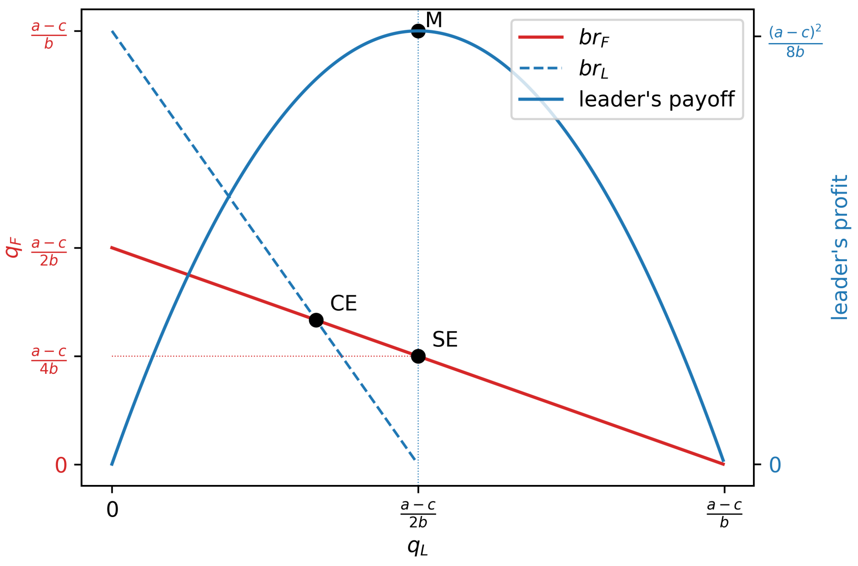


Figure 1.3: The depiction of Stackelberg competition from Example 1.1. M denotes the Stackelberg strategy payoff. The Stackelberg equilibrium SE is given as the intersection of follower's best response function and the Stackelberg strategy (dotted blue line). The Nash equilibrium CE of Cournot duopoly, given as the intersection of best response lines, is presented for comparison.

Usually, the Stackelberg competition analysis is accompanied with the investigation of *entry deterrence* scenario. Particularly, we are interested in determining weather the leader can deter the follower from entering the market and what is the optimal deterrence decision. In our illustration above, this means to set $br_F(q_L) = 0$, which implies $q_L = \frac{a-c}{b}$. In other words, when leader sets it production equal to $\frac{a-c}{b}$, it would not be profitable for the follower to enter the market.

The concept of Stackelberg competition is applied in a number of real–world problems in the domain of decision science, transportation, engineering, military, etc. Some prominent areas are the toll setting [72], natural gas imbalance cash–out problem [71], traffic light optimization [95], synthesis of reactor networks [55, 12], metabolic engineering [55, 21], attacker–defender Stackelberg games [17], central economic planning [1], and facility location [79]. In mathematical optimization research literature the Stackelberg game is usually represented as a bi–level mathematical program.

1.3 The $(r|p)$ Hub–Centroid Problem

Competition in the location theory has been studied for almost several decades. Nevertheless, in the case of hub location problems, the literature is rather scarce. Recently, Mahmutogullari and Kara presented in [88] a new competitive hub location problem, called the $(r|p)$ hub–centroid problem ($(r|p)$HCP), for which they even provided an exact solution approach. Particularly, they investigated a Stackelberg competition in which two transportation companies sequentially enter the market by deploying their hub and spoke networks in order to maximize their respective market shares.

The basic mathematical setting for the problem is a complete digraph $G = (N, A)$, where N is the non–empty node–set and $A \subseteq N^2$ is the set of arcs. A hub can only be established at a node $k \in H$, where $H \subseteq N$ is the subset of nodes that are available to locate hubs. A hub can be shared, and there are no capacity constraints. All hubs should be mutually interconnected. The leader and follower intend to open p and r hubs, respectively. It is assumed that for both competitors the number of hubs to locate is greater than two, as otherwise the economies of scale are not generated.

Multiple allocations of non–hub nodes to hubs are allowed. For each O–D pair $(i, j) \in N^2$, only one route can be established. The concatenation of arcs composes a route, where hubs are located at the joints. At most two hubs are allowed to be on a single route, i.e., at most two stops are permitted.

For every arc $(i, j) \in A$ there is a transportation cost per unit of flow $c_{ij} \geq 0$. The transportation factor α is already known for the market and it corresponds to the transfer between hubs. Transportation cost on a route $i \to k \to l \to j$ is given as $c_{ij,kl} = c_{ik} + \alpha c_{kl} + c_{lj}$, for all $i, j, k, l \in N$.

Customers prefer the leader or follower concerning provided service levels $c_{ij,kl}$. A customer prefers the follower's route if his service level is strictly lower than the leader's. Otherwise, it is assumed that the demand is captured by the leader. In the case of equal service levels, ties are broken in the leader's advantage, since the customers were already served by her.

The demand $w_{ij} \geq 0$ for every O–D pair is taken to be perfectly inelastic. Every customer must be served by one of the competitors. There are no budget constraints. Both competitors have a sufficiently large amount of resources to cover all network installation costs.

The following variables are used to describe the decisions made by the leader:

- $U_{ijk} = \begin{cases} 1, & \text{if a flow from } i \in N \text{ to } j \in N \text{ is routed through } k \in H \text{ as} \\ & \quad \text{the first hub} \\ 0, & \text{otherwise} \end{cases}$

- $V_{ijm} = \begin{cases} 1, & \text{if a flow from } i \in N \text{ to } j \in N \text{ is routed through } m \in H \text{ as} \\ & \quad \text{the second hub} \\ 0, & \text{otherwise} \end{cases}$

- $H_k = \begin{cases} 1, & \text{if the leader locates a hub at node } k \in H \\ 0, & \text{otherwise} \end{cases}$

- $a_{ij} = \begin{cases} 1, & \text{if a flow from node } i \in N \text{ to node } j \in N \text{ is captured by the} \\ & \quad \text{follower} \\ 0, & \text{otherwise} \end{cases}$
- $\beta_{ij} \geq 0$ is a service level for an O–D pair $(i, j) \in N^2$ provided by the leader.

Below are the variables used to represent the follower's network choice:

- $u_{ijk} = \begin{cases} 1, & \text{if a flow from } i \in N \text{ to } j \in N \text{ is routed through } k \in H \text{ as} \\ & \quad \text{the first hub} \\ 0, & \text{otherwise} \end{cases}$
- $v_{ijm} = \begin{cases} 1, & \text{if a flow from } i \in N \text{ to } j \in N \text{ is routed through } m \in H \text{ as} \\ & \quad \text{the second hub} \\ 0, & \text{otherwise} \end{cases}$
- $h_k = \begin{cases} 1, & \text{if the follower locates a hub at node } k \in H \\ 0, & \text{otherwise} \end{cases}$
- $\gamma_{ij} \geq 0$ is a service level for an O–D pair $(i, j) \in N^2$ provided by the follower.

If follower chooses $S \subseteq H$ as a hub backbone then the γ_{ij} is determined as

$$\gamma_{ij} = \min_{k,m \in S} \{c_{ik} + \alpha c_{km} + c_{mj}\} \tag{1.28}$$

In order to represent the sequence of variables, we will use a more compact notation: $U = \{U_{ijk}\}_{i,j,k \in N}$, $V = \{V_{ijm}\}_{i,j,m \in N}$, $H = \{H_k\}_{k \in N}$, $a = \{a_{ij}\}_{i,j \in N}$, $\beta = \{\beta_{ij}\}_{i,j \in N}$, $u = \{u_{ijk}\}_{i,j,k \in N}$, $v = \{v_{ijm}\}_{i,j,m \in N}$, $h = \{h_k\}_{k \in N}$, $\gamma = \{\gamma_{ij}\}_{i,j \in N}$. The set of follower's solutions for a given leader's solution, including the computation of service levels γ, is shortly denoted as $\mathcal{F}(H, U, V, a, \beta)$. The optimal solutions are denoted with asterisk.

The $(r|p)$HCP can be formulated as a bi–level mixed–integer mathematical program. For the leader, Mahmutogullari and Kara propose the following model in [88]:

$$\min \sum_{i,j \in N} w_{ij} a_{ij} \tag{1.29}$$

$$\text{s.t.} \sum_{k \in H} H_k = p, \tag{1.30}$$

$$\sum_{k \in H} U_{ijk} = 1, \quad \forall i, j \in N, \tag{1.31}$$

$$\sum_{m \in H} V_{ijm} = 1, \quad \forall i, j \in N, \tag{1.32}$$

$$U_{ijk} \leq H_k, \quad \forall i, j \in N \wedge \forall k \in H \tag{1.33}$$

$$V_{ijm} \leq H_m, \quad \forall i,j \in N \wedge \forall m \in H \tag{1.34}$$

$$\beta_{ij} \geq \sum_{k \in H} U_{ijk}(c_{ik} + \alpha c_{km}) + V_{ijm} c_{mj}, \quad \forall i,j \in N \wedge \forall m \in H \tag{1.35}$$

$$\beta_{ij} - \gamma^*_{ij} \leq a_{ij} M, \quad \forall i,j \in N \tag{1.36}$$

$$(h^*, u^*, v^*) \in \mathcal{F}(H, U, V, a, \beta) \tag{1.37}$$

$$H_k, U_{ijk}, V_{ijk}, a_{ij} \in \{0,1\}, \quad \forall i,j,k \in N \tag{1.38}$$

$$\beta_{ij} \geq 0, \quad \forall i,j \in N. \tag{1.39}$$

The leader's objective (1.29) is to maximize the amount of flow that she can capture, which is equivalent to the minimization of flow captured by the follower. Constraint (1.30) ensures that the leader will locate p hubs in H. Requirement that flow from $i \in N$ to $j \in N$ will go through at most two hubs k and m, in a given order, is stipulated with (1.31), (1.32), (1.33), and (1.34). Constraints (1.35) serve to calculate the leader's service levels:

- if $V_{ijm} = 0$, the constraint is redundant;
- otherwise, the right hand side becomes a service level provided by the leader for flow from i to j.

The flow takeover by follower is determined with (1.36):

- if the left hand side is positive, i.e., the follower provides a service level for the flow from i to j, better then the leader's one, the right hand side, for a sufficiently large M (BigM), must be positive and $a_{ij} = 1$;
- otherwise, the constraint is redundant.

Constraint (1.37) also states that γ^* is induced from the follower's model. The respective domains of variables are specified with (1.38) and (1.39).

Remark 1.13. The constraints (1.35) allow a situation in which multiple routes can be established for a given O–D pair (i, j), if the service levels are the same. In that case, if the leader captures the flow for (i, j), we can assume that the demand is being equally split among established routes.

For the follower's problem, the following mixed–integer linear program is proposed in the same paper [88].

$$\max \sum_{\substack{i,j \in N \\ m \in H}} w_{ij} v_{ijm} \tag{1.40}$$

$$\text{s.t.} \sum_{k \in H} h_k = r, \tag{1.41}$$

$$\sum_{k \in H} u_{ijk} = 1, \quad \forall i, j \in N \tag{1.42}$$

$$\sum_{m \in H} v_{ijm} \leq 1, \quad \forall i, j \in N \tag{1.43}$$

$$u_{ijk} \leq h_k, \quad \forall i, j \in N \wedge k \in H \tag{1.44}$$

$$v_{ijm} \leq h_m, \quad \forall i, j \in N \wedge m \in H \tag{1.45}$$

$$\sum_{k \in H} u_{ijk}(c_{ik} + \alpha c_{km}) + c_{mj} - \beta_{ij} + \epsilon$$

$$\leq (1 - v_{ijm})M, \quad \forall i, j \in N \wedge \forall m \in H \tag{1.46}$$

$$h_k, u_{ijk}, v_{ijk} \in \{0, 1\}, \quad \forall i, j \in N \wedge \forall k, m \in H. \tag{1.47}$$

The follower's objective (1.40) is to maximize the amount of flow that he can capture. Constraint (1.41) ensures that the follower will locate r hubs in H. Allocation of flow to the first hub is regulated with constraints (1.42). Constraints (1.43) state that flow from $i \in N$ to $j \in N$ can be captured by the follower using a hub located at node $m \in H$. Constraints (1.44) and (1.45) prevent flows going through non–hub nodes. Constraints (1.46) determine captured flows in the following manner.

The left hand side represents the difference in service level provided by the follower and service level provided by the leader. To break ties, a sufficiently small positive value ϵ is added. If the left hand side is non–negative, the corresponding variable v_{ijm} is forced to take 0 value, because the follower cannot provide a better service level than the leader for a given O–D pair (i, j). Otherwise, there are no restrictions on v_{ijm} and M is a large positive value (Big M). When there are no restriction on v_{ijm}, and taking into account the follower's objective function, v_{ijm} is assigned 1. The variables' domains are defined by (1.47).

Remark 1.14. $M = (2 + \alpha) \max_{i,j \in N} c_{ij}$ is large enough since the left hand side can not exceed that value.

Remark 1.15. In this setting, the follower is not obligated to cover all O–D pairs. For every non–hub node i, there must be an allocation to some hub $k \in H$ (because of (1.42)), which allows the transportation from i to k. However, (1.43) says that the transport in the opposite direction does not have to be allowed.

Remark 1.16. Although, it is not mentioned in the model, it is assumed that γ^* in (1.36) is computed from the optimal solution of follower's model using (1.28).

Besides the above model, the paper of Mahmutogullari and Kara [88] provides a few other interesting results:

- a one–level reformulation of $(r|p)$HCP based on min–max approach and total enumeration of possible follower's hub backbones;
- a proof that leader's and follower's problems are both NP–hard;
- the exact algorithm for solving the $(r|p)$HCP.

On Fig. 1.4 we can see how $(2|2)$ hub–centroid optimal solution differs from those of 2-hub median location and 2-hub center location problems, for a CAB instance from [103], when $\alpha = 0.6$. In the interest of clarity, only the (leader's) hub locations are labeled.

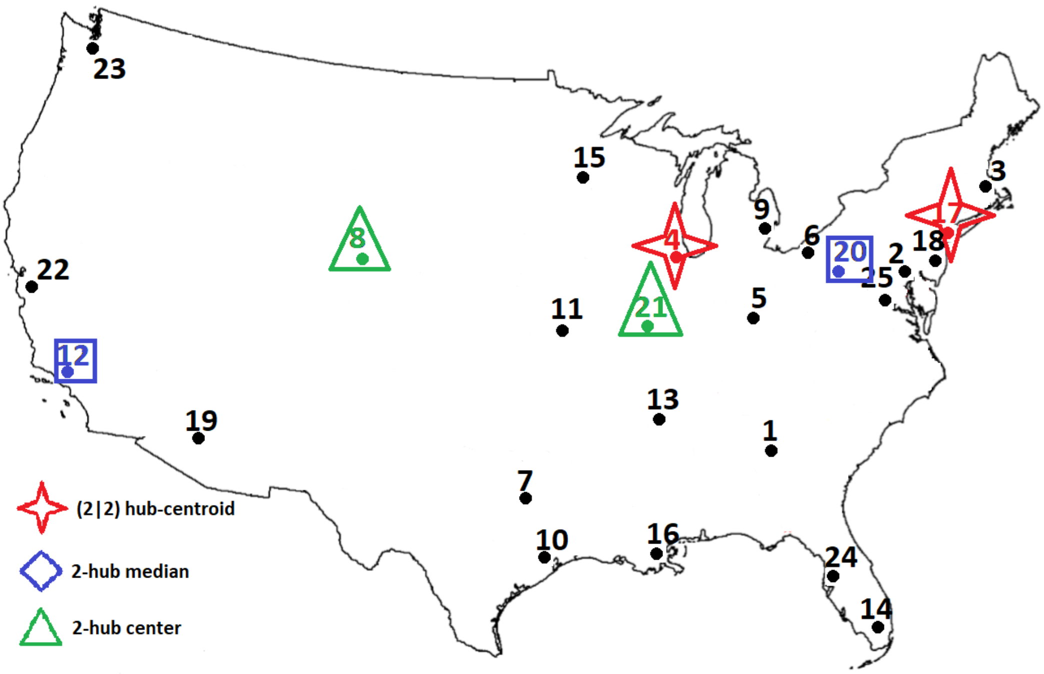


Figure 1.4: Optimal hub locations for $(2|2)$ hub–centroid, 2-hub median, and 2-hub center problems, concerning the CAB instance with $\alpha = 0.6$.

Besides the afore analysis, we can see that in $(r|p)$HCP model (1.29)–(1.47) $a_{ij} = \sum_{m \in H} v_{ijm}$, for $i, j \in N$. Therefore, we can write the follower's objective as

$$\max \sum_{i,j \in N} w_{ij} a_{ij}. \tag{1.48}$$

Furthermore, we note that the leader's objective can be equivalently transformed into the maximization one as

$$-\max - \sum_{i,j \in N} w_{ij} a_{ij}. \tag{1.49}$$

From here, we see that the sum of leader's and follower's function is a constant, i.e., this is a zero–sum game. Therefore, the follower's utility function is surjective and there is no need for defining the auxiliary problem.

Remark 1.17. Mahmutogullari and Kara have not considered in their paper [88] *the entry deterrence* scenario.

In 2020, de Araújo et al. published a paper [33] in which they further investigate the $(r|p)$HCP. Particularly, their results can be summarized as follows:

- a new one–level reformulation of $(r|p)$HCP;
- a proof that $(r|p)$HCP is Σ_2^P–hard;
- an improved exact algorithm for solving the $(r|p)$HCP.

1.4 Bertrand Competition

It is obvious that besides competing on quantities the companies can compete in pricing. In some cases, this seems more logical, especially in the short run. Taking this into account, a French mathematician Joseph Louis François Bertrand (1822–1900) investigated and challenged claims of the previously proposed Cournot model. The underlying assumptions of Bertrand competition are:

- a simultaneous game;
- consumers are indifferent between products/services (homogeneity);
- there are no capacity constraints;
- the strategic choice is on prices, rather than quantities;
- a price–dependent demand function $d(p)$ is known for the market;
- consumers will buy from the firm that offers the lowest price;
- consumers have perfect information;
- the same positive unit cost $c > 0$ for both players.

The Bertrand competition can be presented as the following *infinite game*

$$\Gamma = \langle P, S, u \rangle \tag{1.50}$$

where

- $P = \{A, B\}$ represents the set of players;
- $S = \mathbb{R}_0^{2+}$ represents the set of strategy profiles;

- $u : S \to \mathbb{R}^2$ represents the game utility function, which in this case corresponds to the profit pairs.

Strategy sets for both players are the same, i.e., $S_A = S_B = \mathbb{R}_0^+$. The same hold for the corresponding utility functions $u_A = u_B$, where $u_A, u_B : \mathbb{R}_0^{2+} \to \mathbb{R}_0$. In other words, this simultaneous game is symmetric.

Remark 1.18. This setting requires supply to be perfectly elastic and that the company can easily increase output in response to the sudden surge in demand.

Let the projected prices for players A and B are p_A and p_B, respectively. The demand d_A for competitor A is estimated as

$$d_A = D(p_A, p_B) = \begin{cases} \frac{1}{2}d(p_A), & p_A = p_B \\ 0, & p_A > p_B \\ d(p_A), & p_A < p_B \end{cases} \tag{1.51}$$

The demand d_B for competitor B is estimated in a similar way.

The profit is computed as a product of demand and price–cost difference, i.e., for player A we have that profit is equal to $d_A \cdot (p_A - c)$. Similarly, for player B the profit is equal to $d_B \cdot (p_B - c)$.

If player A decides to sell (provide) products (services) under the price p_A then the best response set of B is given as

$$BR_B(p_A) = \arg\max_{p_B \in \mathbb{R}_0^+} u_B(p_A, p_B) \tag{1.52}$$

$$= \arg\max_{p_B \in \mathbb{R}_0^+} D(p_B, p_A)(p_B - c) \tag{1.53}$$

Analyzing all three cases from (1.51) we easily conclude that the best response is

$$br_B(p_A) = \begin{cases} c, & p_A \leq c \\ p_A - \varepsilon, & p_A > c \end{cases} \tag{1.54}$$

where $\varepsilon > 0$ is an arbitrary small positive number, such that $p_A - \varepsilon > c$.

For this game, the Nash equilibrium, as a solution concept, is called the Bertrand–Nash price equilibrium. Recall that in a Nash equilibrium neither of players has incentive do deviate unilaterally. Therefore, if a strategy profile (p_A^*, p_B^*) is a Nash equilibrium than the following holds:

$$\begin{aligned} &br_A(p_B^*) = p_A^* \ \wedge \ br_B(p_A^*) = p_B^* \ \wedge \\ &br_A(br_B(p_A^*)) = p_A^* \ \wedge br_B(br_A(p_B^*)) = p_B^*. \end{aligned} \tag{1.55}$$

Consider the last equality in (1.55). We have that

$$p_B^* = \begin{cases} c, & br_A(p_B^*) \leq c \\ br_A(p_B^*) - \varepsilon, & br_A(p_B^*) > c. \end{cases} \tag{1.56}$$

which gives us two main cases:

(1) $p_B^* = c$;

(2) $p_B^* = br_A(p_B^*)$.

The first case requires that $br_A(p_B^*) \leq c$. From the first equality in Eq. (1.55) we obtain the condition $p_A^* \leq c$. Taking into account Eq. (1.54) and analyzing both options we realize that p_B^* must be equal to c. Symmetry gives us that p_A^* must equal to c, too. Therefore, a strategy profile (c, c) is one Nash equilibrium.

The second case requires that also $br_A(p_B^*) \geq c$. From the first equality in Eq. (1.55) we obtain the condition $p_A^* \geq c$ and $p_B^* = p_A^* - \varepsilon$. In other words, we have strict inequality $p_B^* < p_A^*$ (because $c > 0$). Symmetry gives us that $p_A^* < p_B^*$, which results in contradiction. Therefore, the only Nash equilibrium is the perfect competition (c, c). A graphical explanation is presented at Fig. 1.5.

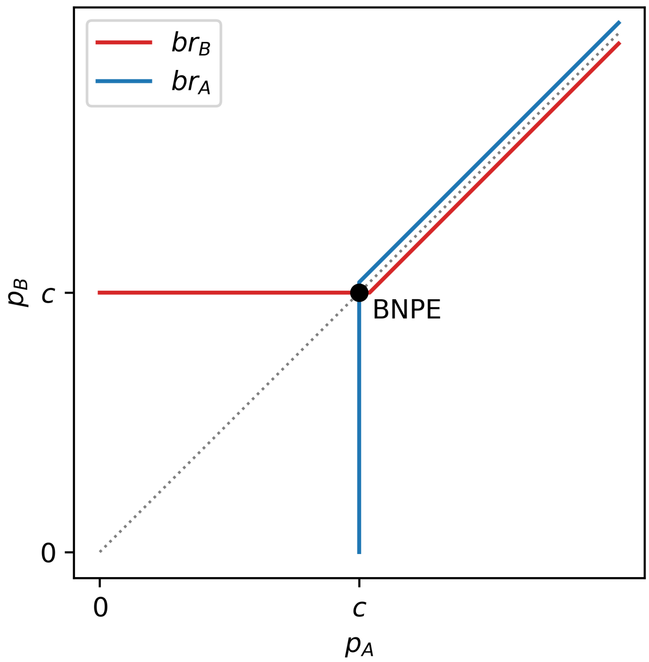


Figure 1.5: Graphical explanation of Bertrand–Nash price equilibrium (BNPE).

Remark 1.19. This equilibrium does not hold with asymmetric cost functions since the company with the lowest marginal costs would seize the entire market and become a monopolist.

Remark 1.20. In the literature, every simultaneous price game is usually called Bertrand's competition (in a broader sense).

The solution of this game, called the Bertrand–Nash price equilibrium, is a bit paradoxical (also known as the Bertrand's paradox). We have that in the case of imperfect competition, where there is a strong incentive to collude, the situation ends up with the same outcome as in perfect competition.

1.5 Logit Model

In econometrics, the discrete choice models (DCMs) are used to make predictions about the choice probabilities between two or more alternatives. Examples are entering or not entering the labor market, choosing between modes of transport, or choosing between transportation routes. The DCMs are in contrast with standard consumption models in which the optimal demand is determined. Loosely speaking, instead of asking "how much", DCM analysis asks "which one". The first applications of DCMs were in transportation planning.

Following the work of Domencich and McFadden [40], it is common to characterize the choice process by four elements: a decision–maker, the choice set, the attributes of alternatives, and a decision rule.

A decision–maker (DM) can be represented by any natural person (e.g., airline passenger) or a juridical one (e.g., a government agency).

All feasible alternatives define a finite set called *universal choice set* C. However, DM may select from a subset of C, defined as *choice set* C_{DM}. For example, the C can represent all routes that connect all pairs of cities in a market. On the other hand, the choice set of individual traveling from a city i to a city j, denoted as C_{ij} would contain only routes connecting i with j. The proper choice set must be collectively exhaustive, mutually exclusive, and finite. The set of available routes connecting cities i and j is finite and mutually exclusive. However, when it comes to collective exhaustiveness, this situation requires a perfectly inelastic demand, i.e., the passenger must take one of the proposed flight routes. Otherwise, the alternative "declining to travel" should be included to make the choice set exhaustive.

Attributes are characteristics of alternatives that DMs consider during the choice process. They can represent both deterministic and stochastic quantities. Examples in airline industry are average fare (price), schedule quality, connection time, aircraft type, etc. In reality, we cannot know all factors that affect the individual choice decisions as their determinants are partially observed or imperfectly measured. Therefore, DCMs rely on stochastic assumptions and specifications to account for unobserved factors.

When it comes to decision rules, it is assumed that they are based on rational behavior, which usually refers to a consistent and transitive preferences of DM. Consistent preferences mean that a DM will consistently choose the same alternative if it is presented with two identical choice situations. Transitive preferences refer to the fact that if alternative A is preferred to alternative B and alternative B is preferred to alternative C, then alternative A is preferred to alternative C. According to [52], the decision rules are grouped into categories: dominance, satisfaction, lexicographic, and utility. In practice, only the last category is considered, because the first three have some limitations: either the resulting choice is not unique or they do not capture how individuals make trade–offs among attributes. The motivation for using maximum utility theory as a decision rule is to represent how individuals make trade–offs among attributes

Utility U_{ni} represents a net benefit value of DM n when choosing an alternative i. It is assumed that DM aims to maximize the utility U_{ni}, i.e., i is chosen if and only if $U_{ni} \geq U_{nj}$, for all $j \neq i$. The internal structure of utility estimation is assumed to be decomposable in linear fashion as

$$U_{ni} = V_{ni} + \varepsilon_{ni}, \tag{1.57}$$

where V_{ni} and ε_{ni} represent observed and unobserved components, respectively. For the first component we have that $V_{ni} = \Theta x_{ni}$, where

- x_{ni} is a vector of observed variables relating to alternative i for DM n that depends on attributes of the alternative or DM (e.g., price, flight duration, gender);
- Θ is a corresponding coefficient vector for the observed attributes (i.e., variables).

The component ε_{ni} captures the impact of all unobserved factors that affect the DM's choice. It may be influenced by many factors, including measurement errors, omitting attributes that are important to choice process, or incorrectly specifying the functional form of attributes. DCM relies on statistical tests to identify violations in assumptions related to error distributions [124]. The probability that DM selects the alternative $i \in C_{ni}$ is given as:

$$\begin{aligned} P_{ni} &= P(U_{ni} \geq U_{nj}\ (\forall j \neq i)) && (1.58)\\ &= P(V_{ni} + \varepsilon_{ni} \geq V_{nj} + \varepsilon_{nj}\ (\forall j \neq i)) && (1.59)\\ &= P(\varepsilon_{nj} \leq V_{ni} - V_{nj} + \varepsilon_{ni}\ (\forall j \neq i)). && (1.60) \end{aligned}$$

The probability that ε_{nj} is less than $V_{ni} - V_{nj} + \varepsilon_{ni}$ is obtained from the cumulative distribution function (CDF), i.e., by integrating over the joint probability distribution function (PDF) of error terms $f(\varepsilon)$

$$P_{ni} = \int_{\varepsilon_i=-\infty}^{+\infty} \cdots \int_{\varepsilon_j=-\infty}^{V_i - V_j + \varepsilon_i} f(\varepsilon) d\varepsilon_{|C_n|} \ldots, d\varepsilon_{i+1} d\varepsilon_i. \tag{1.61}$$

Because CDF is continuous, the case when utility of two alternatives is identical, $U_{ni} = U_{nj}$, is irrelevant to the derivation of choice probabilities.

Specific choice probabilities for different DCMs are obtained by imposing various assumptions on the distribution of these error terms. The assumption that unobserved error components ε are *independently and identically distributed* (iid) and follow a Gumbel distribution leads to the binary logit in case of two alternatives, or the multinomial logit in case of more than two alternatives [93].

1.5.1 Some Properties of Gumbel Distribution

For many DCMs it is assumed that error terms follow the Gumbel distribution, including those which are the most often used for real–life scenarios. CDF and PDF of Gumbel distribution are given respectively as

$$F(\varepsilon; \mu, \Theta) = e^{-e^{-\Theta(\varepsilon-\mu)}}, \quad \Theta > 0, \tag{1.62}$$

$$f(\varepsilon; \mu, \Theta) = \Theta e^{-\Theta(\varepsilon-\mu)} e^{-e^{-\Theta(\varepsilon-\mu)}}, \tag{1.63}$$

where η and $\frac{1}{\Theta}$ are the mode and scale, respectively.

The Gumbel, albeit very similar to the Gaussian distribution, is not symmetric. It is skewed to the right, i.e., its mean is larger than its mode. The mean is equal to $\mu + \frac{\gamma}{\Theta}$, where γ is the Euler–Mascheroni constant. The median is $\mu - \frac{1}{\Theta}\ln(\ln 2)$ and variance is $\frac{\pi^2}{6\Theta^2}$. On Fig. 1.6 and 1.7 we can see how Gumbel and Gaussian distributions differ, having their mean and variance being the same.

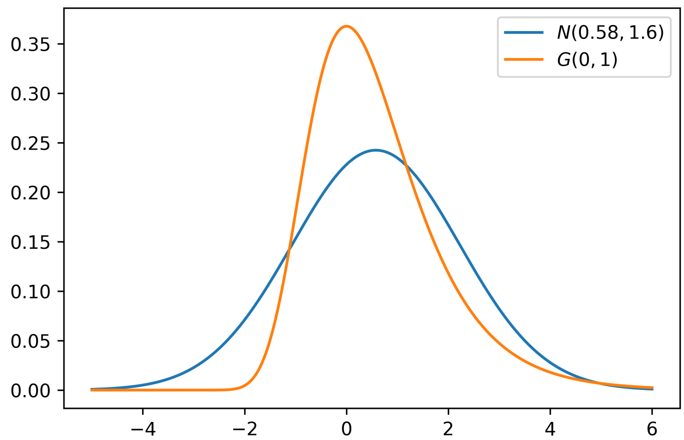


Figure 1.6: The plots of PDFs for normal and Gumbel distributions. The corresponding means and variances are the same.

Let us assume that $\varepsilon \sim G(\mu, \Theta)$ and x and Ω are positive constants. We have that $\varepsilon + x \sim G(\mu + x, \Theta)$ and $\Omega\varepsilon \sim G(\Omega\mu, \Theta/\Omega)$. In other words, every Gumbel distribution

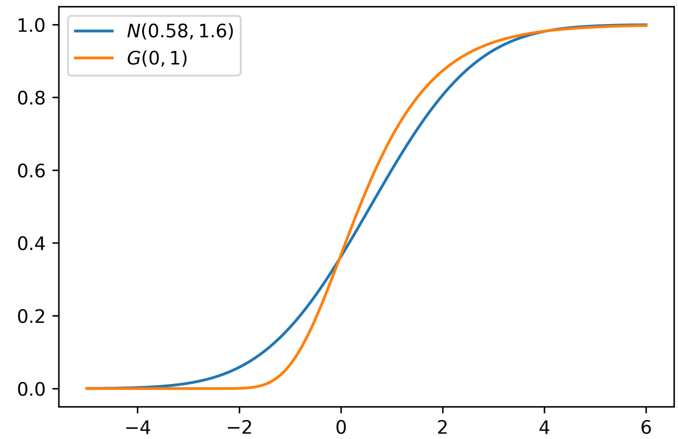


Figure 1.7: The plots of CDFs for Gaussian and Gumbel distributions. The corresponding means and variances are the same.

can be derived from the unit Gumbel distribution by applying the scale and translation. Fig. 1.8 depicts the application of these operations on unit Gumbel distribution.

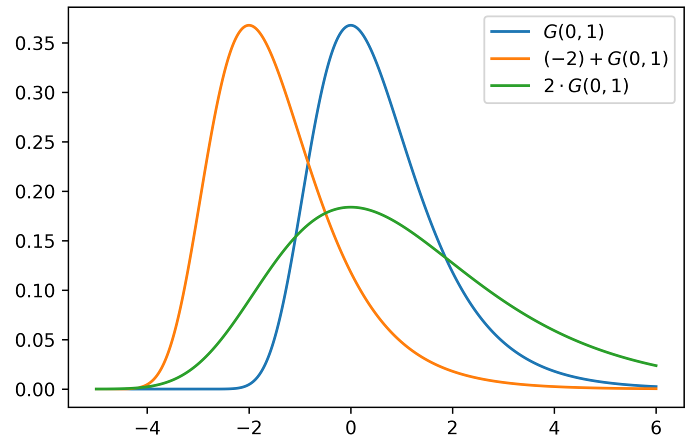


Figure 1.8: Translation (orange) and scale (green) of Gumbel $G(0, 1)$ (blue) distribution.

Under the assumption $\varepsilon_1 \sim G(\mu_1, \Theta)$ and $\varepsilon_2 \sim G(\mu_2, \Theta)$, $\bar{\varepsilon} = \varepsilon_2 - \varepsilon_1$ is logistically distributed with CDF and PDF

$$F(\bar{\varepsilon}) = \frac{1}{1 + e^{\Theta(\mu_2 - \mu_1 - \bar{\varepsilon})}}; \tag{1.64}$$

$$f(\bar{\varepsilon}) = \frac{\Theta e^{\Theta(\mu_2 - \mu_1 - \bar{\varepsilon})}}{1 + e^{\Theta(\mu_2 - \mu_1 - \bar{\varepsilon})}}, \quad \Theta > 0; \tag{1.65}$$

where μ is the mode and $\frac{1}{\Theta}$ is the scale. The logistic distribution is symmetric, i.e., the mean is equal to mode $\mu_2 - \mu_1$. The variance of logistic distribution is $\frac{\pi^2}{3\Theta^2}$.

1.5.2 Choice Probabilities of Logit Model

The multinomial logit model is used to describe how an individual chooses among two or more discrete alternatives. The probabilities are derived from the assumption that error terms are iid Gumbel $G(0,1)$. For that we use the following notation:

- n is DM index;
- C_n is the set of all alternatives for n;
- V_i is the deterministic utility for the i^{th} alternative;
- U_i is the total utility for the i^{th} alternative;
- ε_i is the error associated with the i^{th} alternative.

We have that

$$
\begin{aligned}
P_{ni} &= P(\varepsilon_j < V_i - V_j + \varepsilon_i \ (\forall j \neq i)) && (1.66)\\
&\text{(Treating } \varepsilon_i \text{ as a conditional variable.)}\\
&= \int_{-\infty}^{\infty} \Pi_{\substack{j\in C_n\\ j\neq i}} P(\varepsilon_j < V_i - V_j + \varepsilon_i | \varepsilon_i) f(\varepsilon_i) d\varepsilon_i && (1.67)\\
&= \int_{-\infty}^{\infty} f(\varepsilon_i) \Pi_{\substack{j\in C_n\\ j\neq i}} F(V_i - V_j + \varepsilon_i) d\varepsilon_i && (1.68)\\
&\text{(Assuming } \varepsilon_i \overset{iid}{\sim} G(0,1).)\\
&= \int_{-\infty}^{\infty} e^{-\varepsilon_i} e^{-e^{-\varepsilon_i}} \Pi_{j\in C_n, j\neq i} e^{-e^{-(V_i - V_j + \varepsilon_i)}} d\varepsilon_i && (1.69)\\
&\text{(Since, } e^{-e^{-\varepsilon_i}} = e^{-e^{-(V_i - V_i + \varepsilon_i)}}.)\\
&= \int_{-\infty}^{\infty} e^{-\varepsilon_i} \Pi_{j\in C_n} e^{-e^{-(V_i - V_j + \varepsilon_i)}} d\varepsilon_i && (1.70)\\
&= \int_{-\infty}^{\infty} e^{-\varepsilon_i} e^{-\left(\sum_{j\in C_n} e^{-(V_i - V_j + \varepsilon_i)}\right)} d\varepsilon_i && (1.71)\\
&= \int_{-\infty}^{\infty} e^{-\varepsilon_i} e^{-\left(e^{-\varepsilon_i} \sum_{j\in C_n} e^{-(V_i - V_j)}\right)} d\varepsilon_i && (1.72)
\end{aligned}
$$

In the interest of clarity we can introduce a new variable t, such that $t = -e^{-\varepsilon_i}$, $dt = e^{-\varepsilon_i} d\varepsilon_i$, and $t \in (-\infty, 0)$. Therefore, we have that

$$
\begin{aligned}
P_{ni} &= \int_{-\infty}^{0} e^{\left(t \sum_{j\in C_n} e^{-(V_i - V_j)}\right)} dt && (1.73)\\
&= \left. \frac{e^{\left(t \sum_{j\in C_n} e^{-(V_i - V_j)}\right)}}{\sum_{j\in C_n} e^{-(V_i - V_j)}} \right|_{-\infty}^{0} && (1.74)\\
&= \frac{1}{\sum_{j\in C_n} e^{-(V_i - V_j)}} && (1.75)
\end{aligned}
$$

$$= \frac{1}{e^{-V_i} \sum_{j \in C_n} e^{V_j}} \tag{1.76}$$

$$= \frac{e^{V_i}}{\sum_{j \in C_n} e^{V_j}}. \tag{1.77}$$

From here, taking that $|C_n|= 2$ we obtain the binary logit model probability choices

$$P_{ni} = \frac{1}{1 + e^{-(V_{ni} - V_{nj})}}. \tag{1.78}$$

There is an underlying sigmoid relationship between observed utility and choice probabilities. The relationship between service improvements and existing market position is a subtle point, yet one that is important to consider when making large infrastructure or service improvements [52]. Fundamental properties of the multinomial logit model are:

- only differences in utility are uniquely identified;
- adding a constant to utilities does not affect which alternative has the maximum utility and does not change choice probabilities;
- multiplying utilities by a constant does not affect which alternative has the maximum utility;
- multiplying utilities by a constant changes the relationship between observed utility and choice probabilities, as they are influenced by the variance;
- the independence of irrelevant alternatives, which states that the ratio of choice probabilities between any two alternatives is independent of the availability or attributes of other alternatives.

Remark 1.21. Because only differences in utility are uniquely identified, variables that do not vary over the choice set have to be included in the utility function by interacting with generic ones or by specifying them as alternative–specific. Moreover, this property requires that the location parameter associated with the Gumbel distribution should be normalized to a constant.

1.6 Outline of Dissertation

In Chapter 2, the newly introduced $(r|p)$ hub–centroid problem under the price war is presented. We propose a bi–level mathematical model and investigate its properties. The theoretical examination of proposed model leads to statements about solution existence, computational complexity, and optimal routes. At the end, linear reformulations of the follower's model are presented.

The theoretical investigation from Chapter 2 provides a foundation for a solution approach design, presented in Chapter 3. Because we proved that the $(r|p)$ hub–centroid problem under the price war is NP–hard, the heuristics seem to be natural choice to solve the corresponding bi–level model. Besides that, we show how commercial state-of-the-art solver, Gurobi Optimizer, can be used to solve the lower level model exactly. Having this tool at our disposal, we use it in our algorithms design, based on alternating heuristic and variable neighborhood search.

Chapter 4 presents computational experiments done on the CAB instances, using the algorithms designed in the previous chapter. Results of this empirical investigation help us to examine the stability of solution approach, difference in quality between proposed algorithms, the effect of different parameters, similarities of the resulting hub backbones, and the effect of auxiliary model. Managerial insights are based on the observed patterns in the aforementioned investigation.

Finally, Chapter 5 presents a concluding overview of this research, indicating some directions for the future work.

Chapter 2

The $(r|p)$ Hub–Centroid Problem under the Price War

The mathematician may be compared to a designer of garments, who is utterly oblivious of the creatures whom his garments may fit. To be sure, his art originated in the necessity for clothing such creatures, but this was long ago; to this day a shape will occasionally appear which will fit into the garment as if the garment had been made for it. Then there is no end of surprise and delight.

Tobias Dantzig

Lüer-Villagra and Marianov [86] have argued that the location or route opening decisions could be heavily dependent on the revenues that a company can obtain, while the revenues depend on the price structure. In their paper, they investigate the hub location and pricing from the follower's point of view. O'Kelly et al. [105] focused on a (non–competitive) hub location problem with the price–sensitive demands, deploying an improved Benders decomposition algorithm. Čvokić and Stanimirovic [29] introduced deterministic and robust variants of (non–competitive) single allocation hub location and pricing problem, in which the demand is the price–dependent. As a solution approach for the robust counterpart, besides conic–quadratic mixed–integer reformulation, a 2–phase matheuristic is proposed. However, in all these studies, the response of competition was not considered.

Sasaki and Fukushima have analyzed in [114] a *continuous* Stackelberg competition in which the incumbent competes with several entrants for profit maximization. For every route, only one hub was allowed. Afterward, Sasaki investigated the hub network design model in a competitive environment with a flow threshold [112]. Čvokić et al. introduced in their paper [27] a leader–follower hub location problem under fixed markups, deploying an alternating heuristic as a solution approach. As we already said, Mahmutogullari and Kara addressed an $(r|p)$ hub–centroid problem in which the goal is *a market share* maximization, designing an exact algorithm [88]. In their paper, the demand is divided among the competitors by the "the winner takes it all" rule: a competitor with lower route costs gets the whole demand. Recently, Čvokić et al. introduced the $(r|p)$ hub–centroid problem under the price war [28]. Here, we present this new problem.

Two competing transportation companies intend to enter the market. They are aware of each other. Both of them aspire to maximize *their profits* by finding the best hub and spoke networks and corresponding price structures. The management of one company wants to establish p hubs, and the other one has a plan to locate r hub–facilities. After setting their networks, it is expected that the competing companies will engage in the price war, which assumes responding to the current opponent's pricing with a more competitive one. The solution of price war, if it exists, is a Bertrand–Nash price equilibrium, in which none of the competitors has an incentive to change their price decisions unilaterally.

Usually, two scenarios are considered in the literature: simultaneous and sequential entrance to the market. In the first scenario, the price war is a natural assumption. The issue is that we could expect multiple Nash equilibria to exist when it comes to the competitors' hub and spoke topologies. Finding payoff–dominant equilibrium could be a daunting task. Moreover, the payoff–dominant equilibrium does not need to be composed of pure strategies, i.e., it can be characterized by the cycle of best responses. The standard interpretation of Nash equilibrium composed of mixed–strategies is not admissible. The company will not "flip a coin" to choose a network topology. In the second scenario, the price war is not assumed, i.e., the first competitor that enters the market is committed to its location and price decisions. Nevertheless, a finite Stackelberg price solution implies the existence of a feasible Bertrand–Nash price equilibrium, which opens the door to a cooperative price game with transferable utilities.

Taking all this into account, we may be interested in considering an intermediate variant, i.e., a Stackelberg competition under the price war. One company, the leader, enters the market as the first competitor, anticipating the entrance of the other company, the follower. The prices are set according to the solution of price war. This setting implies that the leader cannot impose prices on the follower, i.e., the pricing is mainly a result of the follower's entrance to the market. We could say that this scenario is equivalent to the search for the Stackelberg strategy if the game is simultaneous, and it may also be related to the search for a price *status quo* point when cooperative pricing is considered.

When it comes to our problem's name, the term "centroid" is taken loosely, i.e., in a more general sense, because this competition does not need to be a zero–sum game. Nevertheless, in this competition, the leader is under a direct attack from the follower. She must take the follower's moves into account and find the best position so that possible harm is minimized while maximizing the profit. The variant of our problem in which the leader is focused on minimizing the follower's profit does not make sense from the economic viewpoint. After all, we can look at using the term "centroid" similarly to how the semantics of term "center" was extended from geometry to the location theory. Of course, we should always leave some space for possibly better naming of this and similar problems.

2.1 Mathematical Formulation

The underlying mathematical setting for the problem is a complete digraph $G = (N, A)$, where N is the non–empty node–set and $A \subseteq N^2$ is the set of arcs. A hub can only be established at some node $k \in N$. Also, hubs can be shared, and there are no corresponding capacity constraints. All hubs should be mutually interconnected. Establishing a hub is considered as a strategic decision.

For every arc $(i, j) \in A$ there is a transportation cost per unit of flow $c_{ij} \geq 0$. For each O–D pair $(i, j) \in N^2$, only one route can be established. Multiple allocations of non–hub nodes to hubs are allowed. The transportation factors $\chi,\ \alpha$ and δ are already known for the market and they correspond to flow consolidation in collection (origin to hub), transfer between hubs, and distribution (hub to destination), respectively. Concatenation of arcs composes a route, where hubs are located at the joints. At most two hubs are allowed to be on a single route, i.e., at most two stops are permitted. Transportation cost on a route $i \to k \to l \to j$ is given as $c_{ij,kl} = \chi c_{ik} + \alpha c_{kl} + \delta c_{lj}$, for all $i, j, k, l \in N$.

It is assumed that the customers choose routes according to observed prices. Both competitors are using the mill pricing, i.e., the customers are paying their expenses. The logit model is used to resolve the issue with a discrete choice. As we know, it is essentially a rule that determines how much of the flow is going to be captured by a competitor. There is a sensitivity parameter $\Theta \geq 0$ with an already known non–negative value assigned. A higher Θ means that customers are susceptible to price differences, so they will mostly choose less expensive routes. On the other hand, a smaller Θ means that the customers are less sensitive to price differences.

The demand $w_{ij} \geq 0$ for every O–D pair is taken to be perfectly inelastic. Every customer must be served by one of the competitors. There are no budget constraints. Both competitors have a sufficiently large amount of resources to cover all network installation costs. Imposing such constraint when the demand is inelastic, and competitors are aiming to maximize their profits does not make much sense. An insufficient budget could lead to *a degenerate solution* in which the competitors are utilizing the nature of an inelastic demand by setting arbitrary high prices on some O–D pairs, which would make the problem trivial and not very interesting. One of the nice features of solution should be that we can compare the leader's and follower's profits. Moreover, the inelastic demand, combined with the competitive setting and high network installation costs, can produce similar problems, even without budget constraints.

Therefore, following the work presented in [113, 112], the study considers only settings where both players are forced to serve all nodes. This approach can also be justified by stating that lots of features, such as representations of economies of scale, are still crude approximations of actual dynamics in a hub and spoke topology, as recently observed by [20]. It should be noted that degenerate solutions are not possible if the demand function is downward–sloping.

The following variables are used to describe the choices made by the leader and follower:

- $x_k = 1$ if the leader has established a hub at node $k \in N$, and 0 otherwise;
- $\rho_{ij,kl} = 1$ if the leader has established a transportation route $i \to k \to l \to j$ from i to j, and 0 otherwise;
- $t_{ij,kl}$ is the price charged by the leader on a route $i \to k \to l \to j$;
- $y_k = 1$ if the follower has established a hub at node $k \in N$, and 0 otherwise;
- $\varsigma_{ij,kl} = 1$ if the follower has established a transportation route $i \to k \to l \to j$ from i to j, and 0 otherwise;
- $q_{ij,kl}$ is the price charged by the follower on a route $i \to k \to l \to j$.

In order to represent the sequence of variables, we will use a more compact notation: $c = (c_{ij,kl})_{i,j,k,l \in N}$, $x = (x_k)_{k \in N}$, $\rho = (\rho_{ij,kl})_{i,j,k,l \in N}$, $t = (t_{ij,kl})_{i,j,k,l \in N}$, $y = (y_k)_{k \in N}$, $\varsigma = (\varsigma_{ij,kl})_{i,j,kl \in N}$, and $q = (q_{ij,kl})_{i,j,k,l \in N}$. The follower's solutions for a given leader's solution are shortly denoted as $\mathcal{F}(x, \rho)$. The optimal solutions are denoted with an asterisk.

Bitran and Ferrer in [13] provided the closed form expression for the optimal response price q over a cost c, when the opponent's price t is known

$$q^* = c + \frac{1}{\Theta} \left(1 + W_0 \left(e^{-\Theta(c-t)-1}\right)\right). \tag{2.1}$$

W_0 is the principal branch of the Lambert W function. Lüer–Villagra and Marianov have generalized this expression in their study [86] by considering HLP with multiple routes of the same O–D pair. This motivates us to introduce the function to represent the optimal price response. Following the result in [86], this new function $\lambda_{ij,kl} : \mathbb{N} \times \mathbb{R}_+^{4|N|^4} \longrightarrow \mathbb{R}$ is defined as

$$\lambda_{ij,kl}(N, c, \rho, t, \varsigma) = c_{ij,kl} + \frac{1}{\Theta} \left(1 + W_0 \left(\frac{\sum_{u,v \in N} e^{-\Theta c_{ij,uv} - 1} \varsigma_{ij,uv}}{\sum_{u,v \in N} e^{-\Theta t_{ij,uv}} \rho_{ij,uv}}\right)\right). \tag{2.2}$$

The $(r|p)$HCPuPW can be represented as a bi–level mixed–integer non–linear mathematical program. For the leader, we propose the following model:

$$\max \sum_{i,j,k,l\in N} w_{ij}(t^*_{ij,kl} - c_{ij,kl}) \frac{\rho_{ij,kl} e^{-\Theta t^*_{ij,kl}}}{\sum_{u,v\in N} \rho_{ij,uv} e^{-\Theta t^*_{ij,uv}} + \sum_{u,v\in N} \varsigma^*_{ij,uv} e^{-\Theta q^*_{ij,uv}}} \tag{2.3}$$

$$\text{s.t.} \sum_{k\in N} \rho_{ij,kl} \leq x_l, \quad \forall i,j,l \in N, \tag{2.4}$$

$$\sum_{l\in N} \rho_{ij,kl} \leq x_k, \quad \forall i,j,k \in N, \tag{2.5}$$

$$\sum_{k,l\in N} \rho_{ij,kl} = 1, \quad \forall i,j \in N, \tag{2.6}$$

$$\sum_{k\in N} x_k = p, \tag{2.7}$$

$$(t^*, q^*, y^*, \varsigma^*) \in \mathcal{F}^*(x,\rho), \tag{2.8}$$

$$x_k, \rho_{ij,kl} \in \{0,1\}, \quad \forall i,j,k,l \in N. \tag{2.9}$$

The leader's profit (2.3) is calculated as a sum of all net incomes. Constraints (2.4) and (2.5) require that the nodes can be allocated solely to hubs. (2.6) stipulates that only one route can be established per O–D pair. The number of hubs to locate is exogenous and specified with Eq. (2.7). Constraint (2.8) denotes that for a given leader's solution only optimal follower's solutions are considered. The domain of decision variables is stated in (2.9).

Recalling the terminology for the bi–level problems, we present two definitions concerning solutions.

Definition 2.1. A solution $((x,\rho),(t,q,y,\varsigma))$ is *semi–feasible* if (x,ρ) satisfies (2.4)–(2.7) and (2.9) and $(t,q,y,\varsigma) \in \mathcal{F}(x,\rho)$.

In other words, in the semi–feasible solutions the optimality is not required for the follower's solution.

Definition 2.2. A solution $((x,\rho),(t,q,y,\varsigma))$ is *feasible*, if it semi–feasible and $(t,q,y,\varsigma) \in \mathcal{F}^*(x,\rho)$, i.e., the follower's solution is optimal.

For the deeper understanding of terminology the reader is referred to the book of Dempe [38].

When it comes to the follower's problem, we propose the following bi–objective mixed–integer non–linear program, for which the preferred solutions are obtained by an *a priori* lexicographic method. It is assumed that the follower's behavior is benevolent, i.e., the leader has optimistic expectations concerning the follower's attitude.

$$\max \sum_{i,j,k,l \in N} w_{ij}(q_{ij,kl} - c_{ij,kl}) \frac{\varsigma_{ij,kl} e^{-\Theta q_{ij,kl}}}{\sum_{u,v \in N} \rho_{ij,uv} e^{-\Theta t_{ij,uv}} + \sum_{u,v \in N} \varsigma_{ij,uv} e^{-\Theta q_{ij,uv}}} \tag{2.10}$$

$$\max \sum_{i,j,k,l \in N} w_{ij}(t_{ij,kl} - c_{ij,kl}) \frac{\rho_{ij,kl} e^{-\Theta t_{ij,kl}}}{\sum_{u,v \in N} \rho_{ij,uv} e^{-\Theta t_{ij,uv}} + \sum_{u,v \in N} \varsigma_{ij,uv} e^{-\Theta q_{ij,uv}}} \tag{2.11}$$

$$\text{s.t.} \sum_{k \in N} \varsigma_{ij,kl} \leq y_l, \quad \forall i,j,l \in N, \tag{2.12}$$

$$\sum_{l \in N} \varsigma_{ij,kl} \leq y_k, \quad \forall i,j,k \in N, \tag{2.13}$$

$$\sum_{k,l \in N} \varsigma_{ij,kl} = 1, \quad \forall i,j \in N, \tag{2.14}$$

$$\sum_{k \in N} y_k = r, \tag{2.15}$$

$$t_{ij,kl} = \lambda_{ij,kl}(N, c, \varsigma, q, \rho), \quad \forall i,j,k,l \in N, \tag{2.16}$$

$$q_{ij,kl} = \lambda_{ij,kl}(N, c, \rho, (\lambda_{ij,uv}(N, c, \varsigma, q, \rho))_{i,j,u,v \in N}, \varsigma), \quad \forall i,j,k,l \in N, \tag{2.17}$$

$$t_{ij,kl}, q_{ij,kl} \geq 0, \quad \forall i,j,k,l \in N, \tag{2.18}$$

$$y_k, \varsigma_{ij,kl} \in \{0,1\}, \quad \forall i,j,k,l \in N. \tag{2.19}$$

The follower's profit (2.10) is calculated as a sum of all net incomes. The behavior of follower as a benevolent and altruistic competitor is defined by (2.11). The constraints (2.12) and (2.13) require that the nodes can be allocated solely to hubs. (2.14) stipulate that only one route can be established per O–D pair. The number of hubs to locate is exogenous and specified with Eq. (2.15). The follower is setting the equilibrium prices (2.16)–(2.17), i.e., he does not have an incentive to change his own price decision. It should be noted that in the price equilibrium, the leader also does not have an incentive to change her price decision. The domains of price and network variables are stated in (2.18)–(2.19).

The main reason behind the bi–objective formulation of follower's model is that we can not claim that our problem is a zero–sum game. The sum of the leader's objective function with the first follower's one does not yield a constant. Ignoring this fact could lead to an ill–posed model.

The lower level model, regarding the first objective solely, is concerned with finding a medianoid affected by the price war, for which the leader's set of hubs H_p is fixed. It is called the $(r|H_p)$ *hub–medianoid problem under the price war* ($(r|H_p)$HMPuPW).

The follower's behavior says that he is impelled to increase the leader's wellbeing (the second objective of lower–level model). Simultaneously, he constrains himself from considering hub and spoke networks which will generate him a suboptimal profit. In the literature, this problem is called *the (follower's) auxiliary problem*, while the corresponding model is accordingly named as *the (follower's) auxiliary model* (AM) (please, see [4, 38]). We note that *a priori* lexicographic method in bi–objective formulation implicitly assumes that the following constraint must be satisfied when solving AM

$$\sum_{i,j,k,l\in N} \frac{w_{ij}(q_{ij,kl}-c_{ij,kl})\varsigma_{ij,kl}e^{-\Theta q_{ij,kl}}}{\sum_{u,v\in N}\rho_{ij,uv}e^{-\Theta t_{ij,uv}}+\sum_{u,v\in N}\varsigma_{ij,uv}e^{-\Theta q_{ij,uv}}} \geq F^*, \tag{2.20}$$

where F^* represents the optimal value of the prior $(r|H_p)$HMPuPW. In other words, $(r|H_p)$HMPuPW can have more than one optimizer and among all of them we are searching for the one which is the best for leader (benevolent behavior).

Remark 2.1. Changing the second objective from maximization to minimization we obtain a model which defines the malicious behavior of follower. In this case, the leader has pessimistic expectations concerning the follower's attitude.

Remark 2.2. If the Bertrand–Nash price equilibrium exists, then the equilibrium equation holds for the leader's prices, too. In other words,

$$t_{ij,kl} = \lambda_{ij,kl}(N, c, \varsigma^*, (\lambda_{ij,uv}(N, c, \rho, t, \varsigma^*))_{i,j,u,v\in N}, \rho). \tag{2.21}$$

In the following section we address the problem of Bertrand–Nash price equilibrium existence.

2.2 Existence of Bertrand–Nash Price Equilibrium

If a finite Bertrand–Nash price equilibrium does not exist, then obviously the set of feasible solutions is empty for both competitors. Then again, there may exist multiple price equilibria. The following theorem resolves this issue.

Theorem 2.3 (Čvokić, Kochetov, Plyasunov, Savić in [28]). *In $(r|p)$HCPuPW, for a given leader's and follower's networks, there exists a unique finite Bertrand–Nash price equilibrium.*

Proof. The objective functions (concerning profit) for both competitors are separable by O–D pairs. Taking into account that the networks are already given, we know which routes are established. Thus, we can focus on a particular O–D pair in our analysis and neglect the indexes entirely. Because each competitor can establish only one route per O–D pair, the best response price constraints are reduced to (2.1).

The derived closed form expression for the best response in terms of margins for the competitors are given as follows:

- the leader's best response margin $r_L(r_F) = \frac{1}{\Theta}\left(1 + W_0\left(Qe^{\Theta r_F - 1}\right)\right)$
- the follower's best response margin $r_F(r_L) = \frac{1}{\Theta}\left(1 + W_0\left(\frac{e^{\Theta r_L - 1}}{Q}\right)\right)$

where $Q = \frac{e^{-\Theta c_L}}{e^{-\Theta c_F}}$. The margins of best responses are bijective functions (continuous, monotone increasing) from a domain of non–negative real numbers, to corresponding co–domains, and vice versa for the inverses. We need to prove that the finite stable point, i.e., a Bertrand–Nash price equilibrium, always exists. In other words, we need to solve the following equation

$$r_L^* = r_L(r_F^*) = r_L(r_F(r_L^*)), \tag{2.22}$$

which is reduced to the system

$$\tau = W_0\left(Qe^{W_0\left(\frac{e^\tau}{Q}\right)}\right) \tag{2.23}$$

$$r_L^* = \frac{\tau + 1}{\Theta} \tag{2.24}$$

Algebra can also be done for the other player, in the same fashion. The principal branch of Lambert W function can be represented by an infinitely nested logarithm as $W_0(x) = \ln\left(\frac{x}{W_0(x)}\right)$. Using this, we can transform Eq. (2.23) into $W_0\left(Qe^{W_0\left(\frac{e^\tau}{Q}\right)}\right) e^\tau = Qe^{W_0\left(\frac{e^\tau}{Q}\right)}$. After multiplication of both sides by $W_0\left(\frac{e^\tau}{Q}\right)$ and simplifying the equation, we obtain the next system of equations with their corresponding constraints

$$W_0\left(Qe^\xi\right) = \frac{1}{\xi} \ \wedge\ \xi > 0 \tag{2.25}$$

$$\xi = W_0\left(\frac{e^\tau}{Q}\right) \tag{2.26}$$

$$r_L^* = \frac{\tau + 1}{\Theta} \ \wedge\ r_L^* \geq 0 \tag{2.27}$$

The first equation always has a solution on $(0, \infty)$. What remains is to check if the solution is feasible, i.e., if $r_L^* \geq 0$. The last two equations result in $e^\tau = Q\xi e^\xi \ \wedge\ \xi > 0 \ \wedge\ \tau \geq -1 \iff \xi \geq W_0\left(\frac{1}{Qe}\right)$. Therefore, we need to prove that $W_0\left(Qe^{W_0\left((Qe)^{-1}\right)}\right) \leq \frac{1}{W_0((Qe)^{-1})}$ for all $Q > 0$. To do that, we will analyze a function $f(Q) = W_0\left((Qe)^{-1}\right) W_0\left(Qe^{W_0\left((Qe)^{-1}\right)}\right)$.

We observe that $\lim_{Q\to\infty} f(Q) = 0$, which can be seen through the series expansion at $x = \infty$. Next, $W_0\left(\frac{1}{Qe}\right) W_0\left(Qe^{W_0\left(\frac{1}{Qe}\right)}\right)^2 = 0$ is representing the first order condition for $f(Q)$, which does not have a solution on $(0, \infty)$. At the end, $\lim_{Q\to 0+} f(Q) = \frac{1}{e}$, because $W_0((Qe)^{-1}) \to \infty$ when $Q \to 0+$, and $\lim_{x\to\infty} xW_0\left(\frac{a}{x}\right) = a$ for some real a, which can again be seen from the series expansion at $x = \infty$. In our case $a = \frac{1}{e}$. The plot of function $f(Q)$ is presented in Fig. 2.1. □

Remark 2.4. On a plot, the Bertrand–Nash price equilibrium can be represented as an intersection of the best response curves, as on Fig. 2.2.

The search for the optimal leader's solution can be based on finding the best feasible follower's hub and spoke topology for which the prices are computed by (2.28)–(2.29).

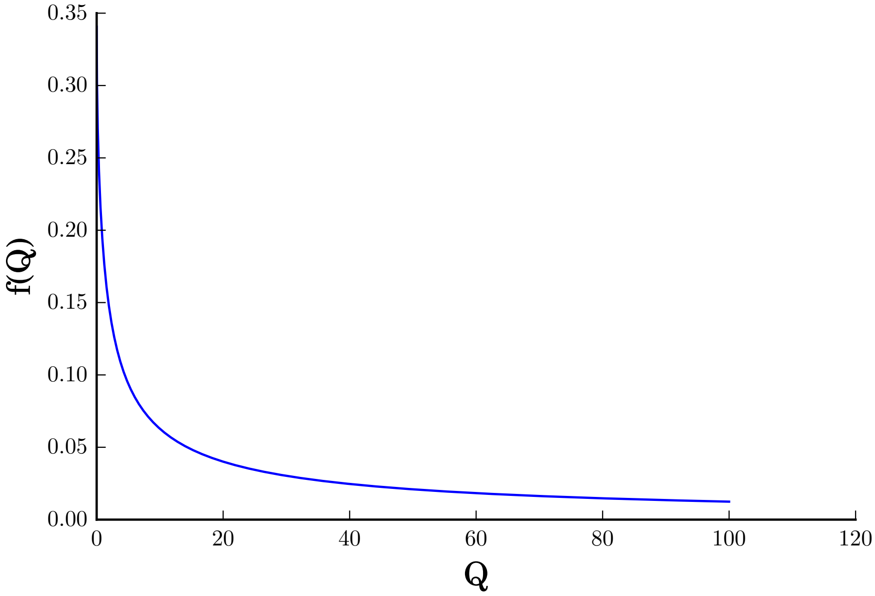


Figure 2.1: The graph of the function $f(Q)$ when $Q \in (0, \infty)$. The limit of $f(Q)$ when $Q \to 0+$ is $\frac{1}{e}$, and the limit when $Q \to \infty$ is 0.

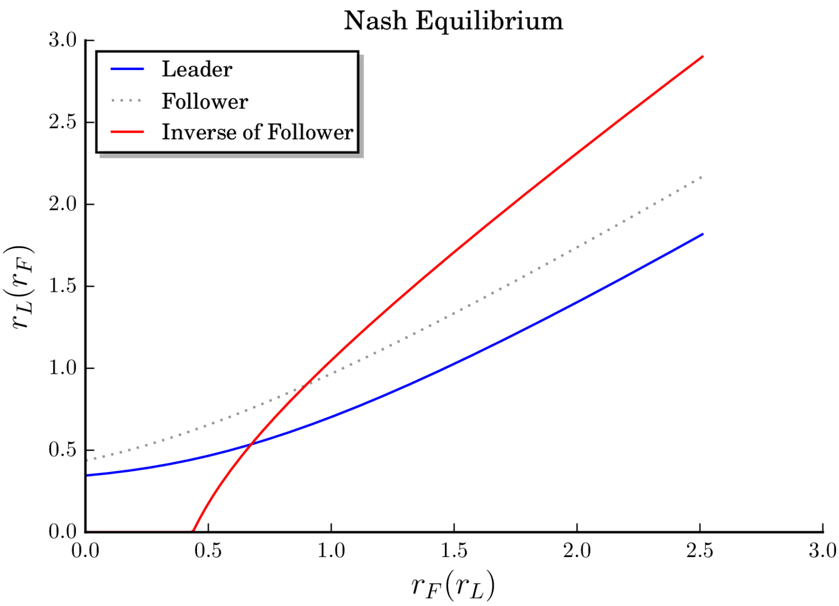


Figure 2.2: The Bertrand–Nash price equilibrium for $\Theta = 3.35$ and $Q = 0.5$, presented as the intersection of best response curves: the leader's (blue) and the follower's (red). Doted line represents the follower's best response in the same coordinate system as the leader's one (just for comparison).

Proposition 2.5 (Čvokić, Kochetov, Plyasunov, Savić in [28]). *For a given leader's network in ($r|p$)HCPuPW, the optimal follower's Bertrand–Nash equilibrium price $q^*_{ij,kl}$ on a route $i \to k \to l \to j$ can be computed by*

$$\tau_{ij,kl} = W_0\left(\frac{e^{\tau_{ij,kl}}}{W_0\left(e^{\tau_{ij,kl}+\Theta(c_{ij,kl}-c_{ij,uv})}\right)}\right) \tag{2.28}$$

$$q^*_{ij,kl} = c_{ij,kl} + \frac{\tau_{ij,kl}+1}{\Theta} \tag{2.29}$$

where (u, v) is the pair of hubs connecting the route established by the leader for the O–D pair (i, j).

Proof. The statements follows from the proof of Theorem 2.3, when the networks are fixed and equations are derived from the follower's point of view. To obtain Eq. (2.28) from Eq. (2.23) we exploit the identity $e^{W_0(x)} = \frac{x}{W_0(x)}$. □

Remark 2.6. Although presented from the follower's viewpoint, the *equilibrium* price equations (2.28)–(2.29) can be used to calculate the corresponding leader's price, too.

Remark 2.7. The new Bertrand-Nash equilibrium follower's pricing does not take into account the leader's price — only the route costs.

Remark 2.8. The logit model and possibly different route costs yield a Bertrand–Nash price equilibrium that is not *a perfect competition*.

Remark 2.9. The proof of Theorem 2.3 proposes the starting point for the numerical computation of the corresponding follower's price

$$\max\left\{W_0\left(e^{-\Theta(c_{ij,uv}-c_{ij,kl})-1}\right), \frac{1}{\Theta}\left(1+W_0\left(e^{-\Theta(c_{ij,uv}-c_{ij,kl})-1}\right)\right)\right\}. \tag{2.30}$$

Remark 2.10. The pair (∞, ∞) is also an equilibrium, but not a feasible one.

Definition 2.3. *The price war sequence* $\{(t_i, q_i)\}_{i=0}^{\infty}$ is composed from the ordered pairs in which t_0 is some arbitrary finite leader's price and

- $q_i = br_F(t_i)$ for $i \in \mathbb{N} \cup \{0\}$, i.e., q_i is the follower's best response to the leader's price t_i;
- $t_i = br_L(q_{i-1})$ for $i \in \mathbb{N}$, i.e., t_i is the leader's best response to the follower's price q_{i-1}.

The corresponding sequences $\{t_i\}_{i=0}^{\infty}$ and $\{q_i\}_{i=0}^{\infty}$ are called *the leader's and follower's induced price war sequences*, respectively. The starting price t_0 is called *the (leader's) war starting price*.

Corollary 2.11. *A price war sequence can be (re)constructed from either the leader's or the follower's induced price war sequence.*

Remark 2.12. We can easily take the follower's viewpoint, i.e., we can construct the price war sequence from his price, as the best response to the leader's one. Of course, in this case we should accordingly adjust the ordering and equalities in the previous definition.

Definition 2.4. The pair of prices (t, q) *is attainable from the leader's price* $t_0 \in [0, \infty)$ if there exists a price war sequence $\{(t_i, q_i)\}_{i=0}^{\infty}$ with t_0 as the leader's war starting price and $n \in \mathbb{N}$ such that $(t, q) = (t_n, q_n)$. Otherwise, we say that this pair of prices is *unattainable (or not attainable) from the leader's price* t_0.

Attainable means there is a *finite* sequence of price moves, starting from t_0 and ending with t, for the leader, and q, for the follower.

Remark 2.13. If $q \neq br_F(t)$ for a price pair (t, q), then (t, q) is not attainable from any war starting price.

Definition 2.5. If a finite pair of prices (t, q) satisfies that $q = br_F(t)$, than we say it is *incidentally attainable from the leader's price* t and t is *an incidental leader's war price for* (t, q).

Remark 2.14. For the sake of clarity, it is worth to mention that in the paper of Nunes and Boatwright [101], published in Journal of Marketing Research—an economics journal, the incidental prices are prices advertised, offered, or paid for unrelated products or goods that neither sellers nor buyers regard as relevant to the price of an item that they are engaged in selling or buying.

Remark 2.15. If (t^*, q^*) is a Bertrand–Nash price equilibrium, then the leader's equilibrium price t^* is incidental.

We can immediately conclude the following.

Corollary 2.16. *In* $(r|p)$*HCPuPW, the Bertrand–Nash price equilibrium* (∞, ∞) *is unattainable from any leader's war starting price.*

Similarly to Corollary 2.16, even the finite Bertrand–Nash price equilibrium is not attainable from any non–incidental leader's war starting price.

Theorem 2.17. *In* $(r|p)$*HCPuPW, for a particular O–D pair, the* finite *Bertrand–Nash price equilibrium* (t^*, q^*)*, is not attainable from any non–incidental leader's war starting price.*

Proof. If we assume the opposite, then we have a finite sequence of ordered pairs—a price war subsequence $(t_0, q_0), (t_1, q_1), \ldots, (t_n, q_n)$, for some $n \in \mathbb{N}$, where $(t_n, q_n) = (t^*, q^*)$, i.e., (t_n, q_n) is a finite Bertrand–Nash price equilibrium. W.l.o.g. we can focus our attention to the last two pairs (t_{n-1}, q_{n-1}) and (t_n, q_n), for which $t_{n-1} \neq t^*$. We know that $br_L(q_{n-1}) = t^*$. Two cases are possible:

(1) $q_{n-1} = q^*$;

(2) $q_{n-1} \neq q^*$.

Regarding the first case, we have that $q^* = br_F(t_{n-1})$ and $q^* = br_F(t^*)$, which implies that $br_F(t_{n-1}) = br_F(t^*)$. From Eq. (2.1), we know that the best response is a bijective function. Therefore, $t_{n-1} = t^*$, which is a contradiction.

The second case can be reduced to the first one. From the last two pairs, taking into account that the Bertrand–Nash price equilibrium is a fixed point and applying Remark (2.12), we can construct the price war sequence taking the follower's viewpoint: $(q_{n-1}, t^*), (q^*, t^*)$. Following the same reasoning as in the previous case, we obtain that $q_{n-1} = q^*$ which contradicts the second case premise. ☐

Not being attainable does not mean it can not be a limit of some price war sequence. However, the finite Bertrand–Nash price equilibrium is feasible and it can be computationally approximated quite well (e.g., by Halley's method). Moreover, we will soon see that it is a limit of every price war sequence. But before proving this, we will present one small suitable lemma.

Lemma 2.18. *For function* $f(r) = \frac{W_0\left(Qe^{W_0\left(\frac{e^{\Theta r-1}}{Q}\right)}\right)+1}{\Theta}$*, when* $Q > 0$ *and* $\Theta > 0$*, there exists* $\bar{r} \geq 0$ *such that* $f(r)$ *is a contraction on a closed bounded interval* $r \in [0, \bar{r}]$.

Proof. Function f is differentiable and monotone increasing on $[0, \infty)$. Its second derivative is

$$\Theta W_0\left(e^{W_0\left(e^{\Theta r-1}/Q\right)}Q\right)\left(\frac{e^{-2W_0\left(e^{\Theta r-1}/Q\right)+2\Theta r-2}}{Q^2\left(W_0\left(\frac{e^{\Theta r-1}}{Q}\right)+1\right)^2\left(W_0\left(e^{W_0(e^{\Theta r-1}/Q)}Q\right)+1\right)^3}\right. \tag{2.31}$$
$$\left.+\frac{e^{-W_0\left(e^{\Theta r-1}/Q\right)+\Theta r-1}}{Q\left(W_0\left(\frac{e^{\Theta r-1}}{Q}\right)+1\right)^3\left(W_0\left(e^{W_0(e^{\Theta r-1}/Q)}Q\right)+1\right)}\right).$$

It is easy to see from this expression that f is convex (concave up) on $[0, \infty)$—the expression (2.31) is always strictly positive. We already know that $f(r) = r$ has a unique solution on $[0, \infty)$ (the proof of Theorem 2.3), which we will denote as r^*. A simple illustrative plot of this intersection is given in Fig. 2.3.

These observations tell us that $|f'(x)| < 1$. We can apply the Mean Value Theorem on f over $[0, \bar{r}]$, if $\bar{r}$ is large enough so that $\{f(r) \ : \ r \in [0, \bar{r}]\} \subseteq [0, \bar{r}]$. It is easy to see that such $\bar{r}$ always exists. Particularly, it would be enough for $\bar{r}$ to take values greater or equal to r^*. From here we see that our f is Lipschitz with a constant $L < 1$. In other words, it is a contraction on a large enough closed bounded interval $[0, \bar{r}]$. ☐

We can now prove a statement about the limit of price war sequence in $(r|p)$HCPuPW.

Theorem 2.19. *In* $(r|p)$*HCPuPW, for a particular O–D pair, the corresponding* finite *Bertrand–Nash price equilibrium is a limit of every price war sequence.*

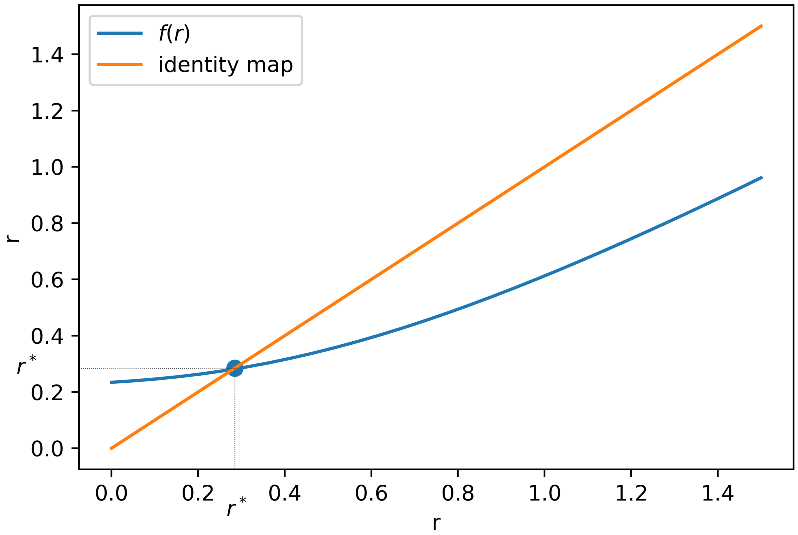


Figure 2.3: The function $f(r)$ from Lemma 2.18 (blue) and identity map (orange), when $\Theta = 7$ and $Q = 0.9$. A point (r^*, r^*) denotes the intersection.

Proof. Take t_0 to be a leader's war starting price. In a way, we want to examine the relationship between t_i and the next leader's price t_{i+1}, which should be the best response to the follower's price q_i—the best response to t_i (for $i \in \mathbb{N}$). Because the corresponding margins preserve relationships (costs are the same), from Eq. (2.22) we have that

$$r_{i+1} = \frac{W_0\left(Qe^{W_0\left(\frac{e^{\Theta r_i - 1}}{Q}\right)}\right) + 1}{\Theta}, \tag{2.32}$$

where r_i and r_{i+1} are margins corresponding to t_i and t_{i+1}, respectively. Here, Q is defined in the same way as in Theorem 2.3. Lemma 2.18 and its proof say that the right hand side of Eq. (2.32) is a contraction on $[0, \bar{r}]$, when $\bar{r} = \max\{r_0, r_L^*\}$ and r_L^* represents the leader's margin in the finite Bertrand–Nash price equilibrium. Now, the statement directly follows from the application of Banach–Caccioppoli fixed point theorem and Corollary 2.11. $\square$

As a consequence of this theorem, we have the following corollary concerning (∞, ∞) case.

Corollary 2.20. *In $(r|p)$HCPuPW, for some particular O–D pair, (∞, ∞) is not a limit of any price war sequence.*

2.3 Existence of Stackelberg Equilibria

Using the Theorem 2.3, we obtain the following result about the existence of Stackelberg equilibrium.

Theorem 2.21 (Čvokić, Kochetov, Plyasunov, Savić in [28])**.** *A safe Stackelberg equilibrium exists for $(r|p)$HCPuPW.*

Proof. The number of possible hub and spoke networks for both players in the market is finite. For each pair of networks, there exists a unique finite Bertrand–Nash price equilibrium. Therefore, $(r|p)$ HCPuPW has a finite/feasible optimal solution. When it comes to the pessimistic leader's expectations about the follower's behavior, we can apply the same reasoning. □

Remark 2.22. Having multiple (safe) Stackelberg equilibria is possible and it does not make the problem ill–posed.

Corollary 2.23. *The unsafe Stackelberg equilibrium exists for* $(r|p)$*HCPuPW.*

Taking into account these results and how demand is dispersed according to the logit model, we have the following corollary.

Corollary 2.24. *For all* $r > 0$*, the entry deterrence is not possible in* $(r|p)$*HCPuPW.*

However, this observation suggest that the deterring of follower can be addressed in a more sophisticated manner. For example, in the $(r|p)$HCPuPW setting, the entry deterrence is, in its nature, close to the minimization of follower's profit. This is also interesting from the naming viewpoint, taking into account our observation from the beginning of chapter.

Additionally, we can provide two more observations.

Theorem 2.25. *In* $(r|p)$*HCPuP, when* $\Theta \to 0$*, the profits of both competitors tend to* ∞*.*

Proof. From Eq. (2.29) we immediately see that $t^*_{ij,kl} \to \infty$ and $q^*_{ij,kl} \to \infty$ (for all $i, j, k, l \in N$). Taking into account constraints (2.6) and (2.14) the limit of market share term in the leader's function is

$$\begin{aligned}
&\lim_{\Theta \to 0} \frac{\rho_{ij,kl} e^{-\Theta t^*_{ij,kl}}}{\sum_{u,v \in N} \rho_{ij,uv} e^{-\Theta t^*_{ij,uv}} + \sum_{u,v \in N} \varsigma^*_{ij,uv} e^{-\Theta q^*_{ij,uv}}} \\
&= \frac{\rho_{ij,kl}}{\sum_{u,v \in N} \rho_{ij,uv} + \sum_{u,v \in N} \zeta^*_{ij,uv}} \\
&= \frac{\rho_{ij,kl}}{2}
\end{aligned}$$

If the route $\rho_{ij,kl}$ is established, the corresponding market share converges to $\frac{1}{2}$. The similar observation holds for the follower. We know that for every O–D pair at least one route must be established. Therefore, the statement follows directly. □

The next proposition addresses the effect of Θ values from the opposite end of its domain.

Theorem 2.26. *In* $(r|p)$*HCPuP, when* $\Theta \to \infty$*, the profits of both competitors converge to zero.*

Proof. From Eq. (2.29) we immediately see that $t^*_{ij,kl} \to c_{ij,kl}$ and $q^*_{ij,kl} \to c_{ij,kl}$ (for all $i, j, k, l \in N$). Now, there are two cases: (1) one competitor has established the route of lower cost and (2) routes of both competitors have the same cost.

In the first case, w.l.o.g., we can assume that the leader has established a lower–cost route. The limit of the corresponding market share term is

$$\lim_{\Theta \to \infty} \frac{\rho_{ij,kl} e^{-\Theta t^*_{ij,kl}}}{\sum_{u,v \in N} \rho_{ij,uv} e^{-\Theta t^*_{ij,uv}} + \sum_{u,v \in N} \varsigma^*_{ij,uv} e^{-\Theta q^*_{ij,uv}}}$$
$$= \lim_{\Theta \to \infty} \frac{1}{1 + \sum_{u,v \in N} \zeta^*_{ij,uv} e^{-\Theta(c_{ij,uv} - c_{ij,kl})}}$$

$$= 1 \qquad \text{(because } c_{ij,uv} > c_{ij,kl} \text{ when } \zeta^*_{ij,uv} = 1\text{)}$$

When it comes to the second case, taking into account constraints (2.6) and (2.14) it is easy to see that the limit of market share term for both players is equal to $\frac{1}{2}$. From here, the statement follows directly in both cases. □

Remark 2.27. These two propositions indicate that, in general, the effect of Θ onto the profit should roughly follow the inverse function.

2.4 Computational Complexity

In this section, we address some questions about the computational complexity of the leader's and follower's problems. We prove that the follower's problem is NP–hard by showing a polynomial reduction from the well known NP–complete decision problem for the r–clique to the standard decision problem of $(r|H_p)$HMPuPW.

Problem 2.1 (The decision problem for the r-clique [88, 51]). Given an undirected graph $G = (N, E)$ and an integer r, determine if G has an r–clique, i.e., that there exists a set of nodes K with $|K| \geq r$ such that for each pair of nodes in K there is an edge in E between them.

Theorem 2.28 (Čvokić, Kochetov, Plyasunov, and Savić in [28]). *The following problems are all NP–hard:*

(1) the $(r|H_p)$HMPuPW.

(2) the auxiliary problem in $(r|p)$HCPuPW.

(3) the follower's problem in $(r|p)$HCPuPW.

Proof. (1) The bi–criteria formulation of the follower requires solving the corresponding $(r|H_p)$HMPuPW. The values of variables $q_{ij,kl}$ $(\forall i, j, k, l \in N)$ can be precomputed, i.e., we can consider them as constants. From the constraint sets (2.12)–(2.14) we know that only one route can be established per O–D pair. If $\varsigma_{ij,kl} = 1$, for some $k, l \in N$, then for all other $k', l' \in N \wedge (k', l') \neq (k, l)$ we have that $\varsigma_{ij,k'l'} = 0$, and vice versa. This means that instead of $\sum_{u,v \in N} \varsigma_{ij,uv} e^{-\Theta q_{ij,uv}}$, we can write just $e^{-\Theta q_{ij,uv}}$ in the denominator, which leads to the following reformulation of the first objective:

$$\max \sum_{i,j,k,l \in N} w_{ij} Q_{ij,kl} \varsigma_{ij,kl}. \tag{2.33}$$

where $Q_{ij,kl}$ is computed as $(q_{ij,kl} - c_{ij,kl}) \frac{e^{-\Theta q_{ij,kl}}}{\sum_{s,t \in N} \rho_{ij,st} e^{-\Theta t_{ij,uv}} + e^{-\Theta q_{ij,kl}}}$. As we have already said, if we show a polynomial reduction from the r-clique decision problem to the standard decision problem of $(r|H_p)$ HMPuPW, then the follower's problem is NP–hard.

Consider an r–clique instance $G = (N, E)$, where N is the set of nodes and E is the set of edges. We can construct a (digraph) network $G' = (N', A')$ where $N' = N$ and $(i, j) \in A' \subseteq N^2$ if $\{i, j\} \in E$. Note that for $(r|H_p)$ HMPuPW, the opponent's network does not need to satisfy the constraints (2.4)–(2.7), nor the route costs have to be computed in the same way. It just needs to be a valid hub and spoke network with non–negative finite values for the route costs.

Now, assume that in the opponent's network all established routes are spokes $i \to i \to j \to j$, for all available O–D pairs $(i, j) \in A'$. Furthermore, take that $\Theta \geq 0$, $\alpha = 1$, and $w_{ij} = \frac{2\Theta}{1+\tau}$ $(\forall i, j \in N)$, where τ is the solution of equation $\tau = W_0(e^{W_0(e^{\tau})})$.

If the r-clique exists, then there is a network (y, ς) in which the hub backbone corresponds to this r–clique. For each inter–hub spoke, both competitors have the same route costs, which implies the same equilibrium prices with margin $\frac{1+\tau}{\Theta}$ and equal market share. Therefore, the profit obtained just on the hub backbone is greater than $\frac{r(r-1)}{2}$.

On the other hand, the solution existence itself implies the existence of r-clique, because of constraint (2.15). If the solution with objective value greater then or equal to $\frac{r(r-1)}{2}$ does not exists, then we have two cases:

· the objective value of every feasible solution is strictly less than $\frac{r(r-1)}{2}$

· the set of feasible solutions is empty.

In the first case, we easily reach the contradiction. In the second case, we infer that there is no r–clique.

Similarly we analyze the situation when $\alpha \in [0, 1)$. We need to assume that the opponent's route costs are all discounted.

(2) We can easily see that almost the same approach can be used to prove that the auxiliary subproblem is also NP–hard. The proof follows basically the same scheme as the proof of previous theorem, except that $Q_{ij,kl}$ is computed as $(t_{ij,kl} - c_{ij,kl})\frac{\rho_{ij,kl}e^{-\Theta t_{ij,kl}}}{\sum_{u,v\in N}\rho_{ij,uv}e^{-\Theta t_{ij,uv}}+e^{-\Theta q_{ij,uv}}}$.

(3) From (1) and (2) we have that (3) follows directly. □

Although, the follower's problem is NP–hard, the corresponding allocation problem is easier to solve.

Theorem 2.29 (Čvokić, Kochetov, Plyasunov, and Savić in [28])**.** *In $(r|p)$HCPuP, the linear relaxation of the corresponding allocation problem for the follower has an integer solution.*

Proof. Regarding the follower's first objective, for an O–D pair $(i, j) \in N^2$, the highest coefficient of variable $\varsigma_{ij,kl}$ (calculated form (2.33)) determines the optimal route. Moreover, a linear relaxation where $\varsigma_{ij,kl} \in [0, 1]$ must have an integer optimal solution. If we assume that the fractional optimal solution exists for the linear relaxation, we can easily see that the corresponding first objective function (2.33) will have a linear deviation.

Regarding the second follower's objective, we know that solution should always be from the set of optimal solutions concerning the first objective. Observe that the higher values of $e^{-\Theta q_{ij,uv}}$ $(u, v \in N)$ (i.e. the lower values of $q_{ij,uv}$) are more preferable, for a given O–D pair $(i, j) \in N^2$. If we assume that the fractional optimal solution exists for this linear relaxation, we can easily see that the corresponding second objective function will have a linear deviation. □

From the proof of this statement, the following two corollaries follow.

Corollary 2.30. *The allocation problem of $(r|H_p)$HMPuPW is polynomially solvable.*

Corollary 2.31. *The allocation problem of auxiliary problem is polynomially solvable.*

Because finding the optimal solution for the leader requires solving the follower's NP–hard problem, we could assess that the leader's problem is NP–hard. The proof of a statement about the leader's computational complexity is based on the vertex cover decision problem.

Problem 2.2 (The decision problem for the vertex cover [88, 51])**.** Given an undirected graph $G = (N, E)$ and an integer p, determine if G has a vertex cover, i.e., if there is a set of vertices C with $|C| \leq p$ such that for each edge $\{i, j\} \in E$, either i or j is in C.

Theorem 2.32 (Čvokić, Kochetov, Plyasunov, and Savić in [28])**.** *The $(r|p)$HCPuPW is NP–hard.*

Proof. Given an instance of vertex cover problem on an undirected graph $G = (N, E)$, we can construct a digraph $G' = (N', A')$ where $N' = N$ and $A' = N' \times N'$. Let

$$w_{ij} = \begin{cases} 1, & \text{if } \{i, j\} \in E \\ 0, & \text{otherwise.} \end{cases} \tag{2.34}$$

We need to show that there exists a vertex cover C with $|C| \leq p$ if and only if exists the set of p nodes H_p, in G', such that the follower's network will coincide with the leader's one on edges (i, j), for which $w_{ij} = 1$. We know from the expression (2.1) what are the margins for both competitors, if their corresponding profits are equal. Therefore, we are able to know exactly the leader's profit, from which we can derive the corresponding standard decision problem.

($\rightarrow$) Assume that the vertex cover problem has a solution $C \subseteq N$ and $|C| \leq p$. We can let $H_p \supseteq C$ and observe that if $i \in H_p$ or $j \in H_p$ then the unit flow w_{ij} may get value 1, depending on their corresponding memberships to E. In all other cases, w_{ij} is always equal to 0. In other words, the pricing can be important only for those O–D pairs (i, j) that have at least i or j in H_p. The leader and follower could only compete for the profit on those routes in which each flow is routed via one (single) spoke with at least one end in C (a subset of H_p). In this situation, the follower can not choose strictly better routes for O–D pairs than those the leader is already using.

($\leftarrow$) Suppose that $H_p \subseteq N'(= N)$ is such that the best follower's response is to copy the leader's solution on all O–D pairs for which $w_{ij} = 1$ and $r > p$. If H_p does not contain as a subset the vertex cover of G, then there exists an edge $\{i, j\} \in E$ and $i \notin H_p$ and $j \notin H_p$ (otherwise H_p would be a vertex cover). Then, on that particular O–D pair (i, j), the follower can profit more than the leader, if his hub backbone includes i or j. In this situation, the follower is offering a non–stop (direct) route to the customers, while the leader has to route the flow through the intermediate hubs. However, this scenario contradicts the assumption that the follower is using the same solution as the leader, as his best one.

Hence, we can conclude that the standard decision problem for the leader is polynomially equivalent to the vertex cover decision problem. $\square$

2.5 Optimal Routes

The function (2.2) describes the best response pricing. Knowing that for every O–D pair each competitor proposes only one price, we realize that (2.2) can be reduced to

$$\lambda_{ij,kl}(N, c, \rho, t, \varsigma) = \varsigma_{ij,kl} \left(c_{ij,kl} + \frac{1}{\Theta} \left(1 + W_0 \left(\frac{e^{-\Theta c_{ij,kl} - 1}}{e^{-\Theta t_{ij,uv}}} \right) \right) \right), \tag{2.35}$$

taking that the opponent has established a route $i \to u \to v \to j$ with its corresponding price $t_{ij,uv}$. Under the assumption that the opponent's price is fixed, we can address the particular income on a given O–D pair as a function of corresponding cost c and derive the following global optimization problem

$$\max_{c \geq 0} \frac{1}{\Theta} \left(1 + W_0(e^{-\Theta(c-t)-1})\right) \frac{e^{-\Theta\left(c+\frac{1}{\Theta}\left(1+W_0\left(e^{-\Theta(c-t)-1}\right)\right)\right)}}{e^{-\Theta t} + e^{-\Theta\left(c+\frac{1}{\Theta}\left(1+W_0\left(e^{-\Theta(c-t)-1}\right)\right)\right)}}. \tag{2.36}$$

Using the identity $e^{W_0(x)} = \frac{x}{W_0(x)}$ this problem can be reduced to

$$\max_{c \geq 0} \frac{1}{\Theta} W_0(e^{-\Theta(c-t)-1}). \tag{2.37}$$

The corresponding first–order condition is $W_0\left(e^{-\Theta(c-t)-1}\right) = 0$ and the solution of this equation does not exist on $\mathbb{R}$. Therefore, no stationary point exists. The objective function in (2.37) is monotone decreasing on $\mathbb{R}_+$, and from this we infer that the lower–cost is more preferred in (2.37). This observation gives us an indication that the lower–cost routes could be better under the Bertrand–Nash price equilibrium.

Because the equilibrium price equations hold for both players, in the following statements we will denote the one considering different route costs – as a competitor and the one which route cost is fixed – as an opponent.

Theorem 2.33. *For every O–D pair $(i, j) \in N^2$, fixed set of hubs H and fixed opponent's route cost, a lower–cost route yields a higher equilibrium price margin, compared to the corresponding one of higher–cost route.*

Proof. W.l.o.g., we can drop the O–D indexes and just focus on the hubs. If $c_{st} < c_{kl}$, then we have that

$$e^{\tau+\Theta(c_{st}-c_{uv})} < e^{\tau+\Theta(c_{kl}-c_{uv})}, \text{ i.e.,} \tag{2.38}$$

$$W_0(e^{\tau+\Theta(c_{st}-c_{uv})}) < W_0(e^{\tau+\Theta(c_{kl}-c_{uv})}), \text{ i.e.,} \tag{2.39}$$

$$\frac{e^{\tau}}{W_0(e^{\tau+\Theta(c_{st}-c_{uv})})} > \frac{e^{\tau}}{W_0(e^{\tau+\Theta(c_{kl}-c_{uv})})}, \text{ i.e.,} \tag{2.40}$$

$$W_0\left(\frac{e^{\tau}}{W_0(e^{\tau+\Theta(c_{st}-c_{uv})})}\right) > W_0\left(\frac{e^{\tau}}{W_0(e^{\tau+\Theta(c_{kl}-c_{uv})})}\right). \tag{2.41}$$

Regrading the Eq. (2.28), the intersection of a linear function with the left–hand side of inequality (2.41) must be higher than the intersection with the right–hand side. Therefore, it is easy to infer from (2.29) that a higher–margin corresponds to a lower–cost route. □

Of course, we can not claim that the profit is higher because of this. The lower–cost route could yield a lower price, regardless of a little bit higher margin. This phenomenon, a counter–intuitive at first sight, is depicted in the following example.

Example 2.1. Consider a setting in which $\Theta = 1$ and the competitors share the same route with transportation cost $c = 1$. The Bertrand-Nash price equilibrium equations (2.28)–(2.29) indicate that for both of them the margin is equal to 2.

Now, assume that a follower can use a route that is half the cost. Incidentally, we can see that in this case, the leader's and follower's equilibrium prices are going to be 1.59 and 1.92, respectively. In other words, the follower's margin is higher than the leader's one, but the relation is opposite for the corresponding prices. ▲

Theorem 2.34. *For every O–D pair $(i, j) \in N^2$, fixed set of hubs H and fixed opponent's route cost, a competitor's lower–cost route yields a lower equilibrium price, compared to the corresponding one of higher–cost.*

Proof. We will ignore route indexes in the interest of clarity. Denote the opponent's route cost as $\bar{c}$. The theorem implies that for two routes with costs c_1 and c_2, when $c_1 \leq c_2$ the following inequality holds

$$c_1 + \frac{\tau_1 + 1}{\Theta} \leq c_2 + \frac{\tau_2 + 1}{\Theta}, \tag{2.42}$$

when τ_1 and τ_2 are computed as

$$\tau_1 = W_0\left(\frac{e^{\tau_1}}{W_0(e^{\tau_1 + \Theta(c_1 - \bar{c})})}\right), \tag{2.43}$$

$$\tau_2 = W_0\left(\frac{e^{\tau_2}}{W_0(e^{\tau_2 + \Theta(c_2 - \bar{c})})}\right). \tag{2.44}$$

The inequality (2.42) is equivalent to

$$\Theta(c_2 - c_1) \geq \tau_1 - \tau_2. \tag{2.45}$$

The Eq. (2.43) can be rewritten as

$$\tau_1 = W_0\left(\frac{e^{\tau_1}}{W_0\left(e^{\tau_1 - \Theta(c_2 - c_1) - \Theta(\bar{c} - c_2)}\right)}\right). \tag{2.46}$$

Again, for the sake of clarity, we can substitute $\Theta(\bar{c} - c_2)$ and $\Theta(c_2 - c_1)$ with a and b, respectively, to obtain the following simpler expressions for (2.43) and (2.44):

$$\tau_1 = W_0\left(\frac{e^{\tau_1}}{W_0(e^{\tau_1 - (a+b)})}\right), \tag{2.47}$$

$$\tau_2 = W_0\left(\frac{e^{\tau_2}}{W_0\left(e^{\tau_2 - a}\right)}\right). \tag{2.48}$$

Also, the inequality (2.45) becomes

$$b + \tau_2 \geq \tau_1. \tag{2.49}$$

The following two equations immediately follow from Eq. (2.47)–(2.48) and identity $W_0(x) = \ln\left(\frac{x}{W_0(x)}\right)$:

$$\tau_1 W_0\left(e^{\tau_1-(a+b)}\right)-1=0, \tag{2.50}$$

$$\tau_2 W_0\left(e^{\tau_2-a}\right)-1=0. \tag{2.51}$$

We can take the left hand side of (2.51) to be a function and translate it to the right for vector $(b, 0)$. This operation will yield a new function

$$T_{a+b}(\tau)=(\tau-b)W_0\left(e^{\tau-(a+b)}\right)-1. \tag{2.52}$$

Obviously, for every $\tau \in \mathbb{R}$ we have that

$$(\tau-b)W_0\left(e^{\tau-(a+b)}\right)-1 \leq \tau W_0\left(e^{\tau-(a+b)}\right)-1, \tag{2.53}$$

where right hand side of (2.53) is practically the left hand side of (2.50). Therefore, the zero of T_{a+b} is larger than the solution of Eq. (2.50). From here we conclude that the inequality (2.49) holds. □

Now, we need to prove that the lowest–cost route generates the highest profit. We remind the reader that a profit is affected by a margin and a market share combined. Furthermore, both of them depend on the equilibrium prices, which are affected by route costs, as it was described in Theorems 2.33 and 2.34.

Theorem 2.35. *For every O–D pair $(i, j) \in N^2$, fixed set of hubs H, and fixed opponent's route cost, the lowest–cost route is the optimal one.*

Proof. As it was done in the proof of previous two theorems, we ignore the route indexes in the interest of clarity. Denote the opponent's route cost as $\bar{c}$. The theorem implies that by lowering the route cost c, i.e., by increasing a difference $\bar{c}-c$, the corresponding profit computed as

$$(t-c)\frac{e^{-\Theta t}}{e^{-\Theta t}+e^{-\Theta q}} \tag{2.54}$$

will increase, too. Here, t and q are the equilibrium prices corresponding to the route costs c and $\bar{c}$, respectively.

From the opponent's viewpoint, the equivalent form of Eq. (2.28) is

$$\bar{\tau} W_0\left(e^{\bar{\tau}+\Theta(\bar{c}-c)}\right)-1=0. \tag{2.55}$$

Increasing $\bar{c}-c$, i.e., decreasing c, leads to the increased value of Lambert W function term. Therefore, the solution of Eq. (2.55) *decreases*. In other words, we have that the lower–cost route yields a lower opponent's equilibrium price, too.

For equilibrium prices t and q we have the following two respective forms of Eq. (2.28): $\tau W_0\left(e^{\tau-\Theta(\bar{c}-c)}\right)-1=0$ and $\bar{\tau} W_0\left(e^{\bar{\tau}+\Theta(\bar{c}-c)}\right)-1=0$. Subtracting these two equations we obtain

$$\tau W_0\left(e^{\tau-\Theta(\bar{c}-c)}\right)=\bar{\tau} W_0\left(e^{\bar{\tau}+\Theta(\bar{c}-c)}\right) \tag{2.56}$$

This ratio of τ and $\bar{\tau}$ can be evaluated from the last equation as

$$\frac{\tau}{\bar{\tau}} = \frac{W_0\left(e^{\bar{\tau}+\Theta(\bar{c}-c)}\right)}{W_0\left(e^{\tau-\Theta(\bar{c}-c)}\right)} \tag{2.57}$$

Increasing $\bar{c}-c$ leads to the increased ratio of $\frac{\tau}{\bar{\tau}}$. Because $e^{-\Theta x}$ is monotone, we have that $\frac{e^{-\Theta\tau}}{e^{-\Theta\bar{\tau}}}$ increases, as denominator decreases faster than numerator. In other words, increasing $\bar{c}-c$, i.e., taking lower–cost routes, leads not only to a larger margin (as the Theorem 2.33 states), but it also leads to a (slightly) larger market share. Therefore, taking into account Eq. (2.54) the theorem holds. □

On Fig. 2.4 we can see an illustration of both price and profit deviations when route cost c changes its value, in a setting where $\bar{c} = 3$ and $\Theta = 3$. At point $c = 0$, i.e., when $\bar{c} - c = 3$, the corresponding profit deviations are maximal. However, taking the route cost c closer to $\bar{c}$ (which always remains the same), leads to a situation in which prices and profits are equal. Furthermore, we can see that decrease in the route cost c yields lower prices for the opponent, as well as for the competitor. This mechanism does not work for the profits. The competitor's profit increases, while the opponent's one decreases.

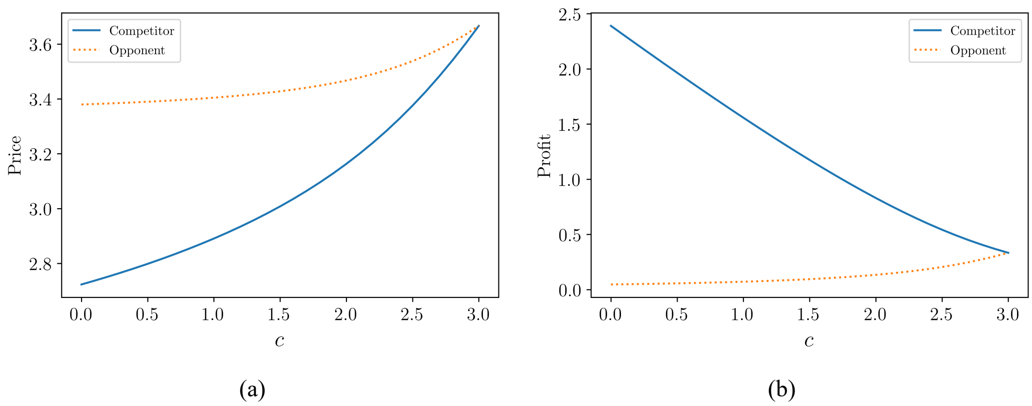


Figure 2.4: Curves of (**a**) price deviations and (**b**) profit deviations. The blue (ordinary) and orange (dotted) lines correspond to the competitor and opponent, respectively.

Remark 2.36. The proof of Theorem 2.35 does not depend substantially on Theorem 2.34.

Remark 2.37. The statements of Theorems 2.33, 2.34, and 2.35 do not assume the role of competitors, because they arise from the equilibrium equations.

From these three theorems, we immediately infer the following two corollaries.

Corollary 2.38. *If both competitors in* $(r|p)$*HCPuPW have their corresponding hub backbones fixed, then the sum of (2.3) and (2.10) is a constant.*

In other words, an allocation problem derived from $(r|p)$HCPuPW when hub backbones are fixed, does not require AM, i.e., a bi–objective formulation of the follower's model.

Corollary 2.39. *In $(r|p)$HCPuPW, the AM solution will not affect the leader's profit if its hub backbone is the same as the one obtained by solving the corresponding $(r|H_p)$HMPuPW.*

2.6 Reformulation of Follower's Model

Having an integer linear programming formulation of the problem allows us to use state-of-the-art solvers. In the following subsection, we show how to reformulate our model for $(r|H_p)$HMPuPW. Subsequently, we also show how to reformulate AM as an integer linear program.

2.6.1 Linear Reformulation of Model for $(r|H_p)$HMPuPW

The entries of (ordered) equilibrium price pairs $(t_{ij,uv}, q_{ij,kl})$ (for all $i, j, u, v, k, l \in N$) can be computed in the preprocessing phase using Eq. (2.28)–(2.29). For every pair we can define a function $a_{ij,uv,kl} : \{1, 2\} \to \{t_{ij,uv}, q_{ij,kl}\}$ as

$$a_{ij,uv,kl}(x) = \begin{cases} t_{ij,uv}, & x = 1 \\ q_{ij,kl}, & x = 2. \end{cases} \tag{2.58}$$

These functions make a foundation for new functions $T_{ij,uv} : N^2 \to \mathbb{R}$ which serve to extract the first entry from the equilibrium price pair

$$T_{ij,uv}(k, l) = a_{ij,uv,kl}(1). \tag{2.59}$$

Similarly, we can introduce functions $Q_{ij,kl} : N^2 \to \mathbb{R}$, defined as

$$Q_{ij,kl}(u, v) = a_{ij,uv,kl}(2), \tag{2.60}$$

to extract the second entry from the equilibrium price pairs. In other words, during the preprocessing phase, we evaluate function values $T_{ij,uv}(k, l)$ and $Q_{ij,kl}(u, v)$ for all $i, j, u, v, k, l \in N$.

Recall that the leader's network is fixed when we are solving the follower's problem. From (2.6) we realize that the numerator in the follower's first objective can be multiplied by $\sum_{u,v\in N} \rho_{ij,uv}$, without affecting its value. Therefore, in the numerator we can write

$$\sum_{u,v\in N} \rho_{ij,uv} w_{ij} \left(Q_{ij,kl}(u, v) - c_{ij,kl}\right) \varsigma_{ij,kl} e^{-\Theta Q_{ij,kl}(u,v)}. \tag{2.61}$$

If $\varsigma_{ij,kl} = 1$ for the particular $k, l \in N$, then for all other $k', l' \in N \wedge (k', l') \neq (k, l)$ constraint (2.14) stipulates that $\varsigma_{ij,k'l'} = 0$. Therefore, instead of $\sum_{u,v\in N} \varsigma_{ij,uv} e^{-\Theta q_{ij,uv}}$, we can just write $e^{-\Theta q_{ij,kl}}$ in the denominator.

Theorem 2.29 states that the linear relaxation of corresponding allocation problem for the follower has an integer solution, i.e., we can relax the route variables $\varsigma_{ij,kl}$ (for all

$i, j, k, l \in N$). Therefore, we have a new way to formulate the model for $(r|H_p)$HMPuPW as

$$\max \sum_{i,j,k,l\in N} w_{ij} \frac{\sum_{u,v\in N} \rho_{ij,uv} \left(Q_{ij,kl}(u,v) - c_{ij,kl}\right) e^{-\Theta Q_{ij,kl}(u,v)}}{\sum_{u,v\in N} \rho_{ij,uv} \left(e^{-\Theta T_{ij,uv}(k,l)} + e^{-\Theta Q_{ij,kl}(u,v)}\right)} \varsigma_{ij,kl} \tag{2.62}$$

$$\text{s.t. } (2.12)\text{–}(2.15), \tag{2.63}$$

$$y_k \in \{0,1\}, \qquad \forall k \in N, \tag{2.64}$$

$$\varsigma_{ij,kl} \in [0,1], \qquad \forall i,j,k,l \in N. \tag{2.65}$$

This formulation has $|N|$ integer and $|N|^4$ continuous variables, compared to $|N|^4+|N|$ and $2|N|^4$ in the original version. The number of constraints is $2|N|^3+|N|^2+1$, compared to $2|N|^4+2|N|^3+|N|^2+1$. Interestingly, the reformulation gives us a model with reduced number of variables and constraints.

Remark 2.40. For this problem, one needs to pay special attention to the preprocessing. In case of one level problems, this issue is rarely addressed, if ever. The preprocessing phase for the lower–level model is considered being part of the algorithm that solves the upper–level one. In our case, the preprocessing consists of computing the Bertrand–Nash price equilibrium for every pair of routes, which has a computational complexity $O(|N|^6)$.

2.6.2 Linear Reformulation of AM

The AM is a non–linear fractional mathematical program. Concerning the leader's routes we have two cases:

(1) $\rho_{ij,kl} = 1$, and we can write $e^{-\Theta t_{ij,kl}}$ instead of $\sum_{u,v\in N} \rho_{ij,uv} e^{-\Theta t_{ij,uv}}$, in the denominator;

(2) $\rho_{ij,kl} = 0$, i.e., it does not matter what is in the denominator.

Therefore, we should focus only on the first case. The denominator can look as

$$e^{-\Theta t_{ij,kl}} + \sum_{u,v\in N} e^{-\Theta q_{ij,uv}}. \tag{2.66}$$

Constraint (2.14) allows us to multiply the left–term with $\sum_{u,v\in N} \varsigma_{ij,uv}$. For a given O–D pair (i,j) we are searching for an adequate follower's route through hubs u and v, when the leader's route $i{\to}k{\to}l{\to}j$ is fixed. Using the extraction functions from the previous subsection, we can write the last expression as

$$\sum_{u,v\in N} \varsigma_{ij,uv} \left(e^{-\Theta T_{ij,kl}(u,v)} + e^{-\Theta Q_{ij,uv}(k,l)}\right). \tag{2.67}$$

Also, because of (2.14) we can multiply the whole fraction with $\sum_{u,v \in N} \varsigma_{ij,uv}$ to obtain the following equivalent expression for the objective

$$\max \sum_{i,j,k,l \in N} \sum_{u,v \in N} \frac{\varsigma_{ij,uv} w_{ij} \left(T_{ij,kl}(u,v) - c_{ij,kl}\right) \rho_{ij,kl} e^{-\Theta T_{ij,kl}(u,v)}}{\sum_{m,n \in N} \varsigma_{ij,mn} \left(e^{-\Theta T_{ij,kl}(m,n)} + e^{-\Theta Q_{ij,mn}(k,l)}\right)}. \tag{2.68}$$

The same analysis for the leader's routes can be applied to those of the follower. Therefore, we can substitute the objective (2.11) with the following linear one

$$\max \sum_{i,j,k,l \in N} \sum_{u,v \in N} \frac{\varsigma_{ij,uv} w_{ij} \left(T_{ij,kl}(u,v) - c_{ij,kl}\right) \rho_{ij,kl} e^{-\Theta T_{ij,kl}(u,v)}}{e^{-\Theta T_{ij,kl}(u,v)} + e^{-\Theta Q_{ij,uv}(k,l)}}. \tag{2.69}$$

The Corollary 2.39 states that we are not interested in follower's networks that share the same hub backbone. Following this theoretical observation, we introduce another linear constraint, as a valid inequality, in order to differentiate obtained solutions from those that share the same hub backbone with $(r|H_p)$HMPuPW

$$\sum_{k \in N} y_k^* y_k \leq r - 1. \tag{2.70}$$

Here, y_k^* represent the optimal hub backbone from the previously solved $(r|H_p)$ HMPuPW. Finally, we provide an integer linear programming formulation of AM:

$$\max \sum_{i,j,k,l \in N} \sum_{u,v \in N} \frac{\varsigma_{ij,uv} w_{ij} \left(T_{ij,kl}(u,v) - c_{ij,kl}\right) \rho_{ij,kl} e^{-\Theta T_{ij,kl}(u,v)}}{\left(e^{-\Theta T_{ij,kl}(u,v)} + e^{-\Theta Q_{ij,uv}(k,l)}\right)} \tag{2.71}$$

s.t. (2.12)–(2.15), (2.70),

$$\sum_{i,j,k,l \in N} w_{ij} \frac{\sum_{u,v \in N} \rho_{ij,uv} \left(Q_{ij,kl}(u,v) - c_{ij,kl}\right) e^{-\Theta Q_{ij,kl}(u,v)}}{\sum_{u,v \in N} \rho_{ij,uv} \left(e^{-\Theta T_{ij,uv}(k,l)} + e^{-\Theta Q_{ij,kl}(u,v)}\right)} \varsigma_{ij,kl} \geq F^*, \tag{2.72}$$

$$y_k \in \{0,1\}, \quad \forall k \in N, \tag{2.73}$$

$$\varsigma_{ij,kl} \in [0,1], \quad \forall i,j,k,l \in N, \tag{2.74}$$

where F^* represents the optimal value of the prior $(r|H_p)$HMPuPW.

This reformulation of AM has $|N|$ integer and $|N|^4$ continuous variables, compared to $|N|^4$+$|N|$ integer and $2|N|^4$ continuous ones in the original formulation. The number of constraints is $2|N|^3$+$|N|^2$+3, while the original program has $2|N|^4$+$2|N|^3$+$|N|^2$+1 of them. Interestingly, the new linear formulation of AM is much more compact than the original one. The preprocessing is not needed, because the equilibrium price pairs are the same as those for the $(r|H_p)$HMPuPW.

Chapter 3

Solution Approach

If you do not have someone to ask, and you do not know what to do, ask a little child and it will give you an answer.

Old Serbian proverb

Previously, we have introduced some relevant properties of $(r|p)$HCPuPW. In this chapter, we describe how this problem can be solved. The goal of solution approach is to find the optimal or near–optimal solution of high quality, hopefully without an unwanted high computational effort. For an optimization problem, a near–optimal solution is a feasible one with an objective function value within a specified range from the (usually unknown) optimal objective function value [100]. The computational effort for optimization problems can be measured as the computational time and space consumed during the execution of solution approach. According to one classification, the solution approaches can be distinguished between two different types:

- general methods;
- specific problem class methods.

The first type can be divided into local and global optimization methods. Except for specific problems, the local optimization methods only provide locally optimal results. However, their computational cost is usually much lower than those of global ones.

On the other hand, solution approaches can also be classified into:

- exact methods — finding the optimal solution is guaranteed;
- approximation methods — sub–optimal algorithms with provable guarantees about the quality of their output solution;
- heuristics — a guarantee of any kind is not implied, but executions on a valid sample of test instances show that high–quality solutions can be easily found.

Usually, an exact optimization method is the solution approach of choice if it can solve an optimization problem with an effort that grows polynomially with the problem size. As we saw in Chapter 2, the linear relaxation of follower's problem, and the corresponding allocation problems belong to the class P. The situation is different if a problem is NP–hard, as the decision about the number of alternatives is highly combinatorial. In those cases, even instances of medium size could be intractable. Moreover, even the approximation methods may not suffice for practical applications.

These issues are usually resolved by using heuristics as they often show excellent performance for many NP–hard problems of practical relevance. In particular, there is a trade–off between a guarantee for the solution quality and a computational effort. Roughly speaking, heuristics can be distinguished as the problem–specific heuristics, metaheuristics, and matheuristics.

The problem–specific heuristics should be based on a robust underlying theory: they are either derived in a top–down manner from the theory about the problem model or based on experimental and real–world instances. Others are just so–called *rules of thumb* based on a real–world observation or experience, sometimes without even a glimpse of theory. The later are therefore exposed to a more significant number of pitfalls.

A metaheuristic is a heuristic framework designed to find, generate, or select a partial search algorithm to find sufficiently good solutions to an optimization problem. They are convenient when the information is imperfect, incomplete, or with limited computational capacities. It samples a too large solution set and may make few assumptions about the optimization problem being solved. These properties make them usable for a variety of problems, hence the origin of name. Here we present a small list of some well–known metaheuristic approaches for combinatorial optimization problems:

- simulated annealing;
- tabu search;
- greedy randomized search procedure;
- variable neighborhood search;
- evolutionary algorithms;
- particle swarm optimization;
- artificial immune systems.

Matheuristics are solution approaches for the optimization problem created by the inter–operation of metaheuristics and mathematical programming techniques. Some features are derived or proved from the corresponding mathematical model of the optimization problem, which are then utilized in the part of (meta)heuristic algorithm for the exploitation. Occasionally, in the literature, they are called *a model–based heuristics*. One should be aware that the use of mathematical programming techniques as heuristics to solve optimization problems, is much older and much more widespread than the matheuristics. However, this is not the case with metaheuristics. Even the very idea of designing mathematical programming methods specifically for a heuristic solution has innovative traits. This viewpoint is different from the one when enough computational resources are not available and thus exact methods turn into approximation methods or heuristics. The merging of mathematical programming techniques and metaheuristics can go two–ways:

(1) mathematical programming is used to improve or design metaheuristics;

(2) metaheuristics are used to improve known mathematical programming techniques.

The first of these two directions is more studied in the literature. On the other hand, the second direction is widely used in commercial solvers for the exact solving of optimization problems.

In the following section of this chapter, we present the Gurobi Optimizer, a current state-of-the-art commercial solver for finding the optimal solution of (integer) linear programs, and how it can be used to find optimal solutions of follower's problem. Afterward, for the leader, an alternating heuristic and variable neighborhood search are presented as matheuristic solution approaches for the leader's problem.

3.1 Exact Solver for Follower's Model

Designing an efficient branch-and-cut algorithm to solve mixed–integer linear program from scratch is far beyond the scope of this dissertation. Following the latest development in this area, one can be quite content with using a commercial optimization solver for making choices regarding preprocessing, choosing branching variables, and whether to branch or cut at a current subproblem. Therefore, the follower's model or the corresponding derived models will be solved by an exact solver.

Gurobi Optimizer [56], the product of Gurobi Optimization, LLC, is a state-of-the-art commercial optimization solver intended to solve linear programs, quadratic programs, quadratically constrained programs, mixed–integer linear programs, mixed–integer quadratic programs, and mixed–integer quadratically constrained program. The company was founded in 2008 and named according to the surnames of its founders [57]: Zonghao <u>Gu</u>, Edward <u>Ro</u>thberg, and Robert <u>Bi</u>xby. Bixby was also the founder of CPLEX (commercial optimization solver of IBM corp.) [58], while Rothberg and Gu were leading its development for almost a decade [59, 60]. According to [56], the Gurobi Optimizer supports a variety of programming and modeling languages including:

- object–oriented interfaces for C++, Java, .NET, and Python;
- matrix–oriented interfaces for C, MATLAB, and R;
- links to standard modeling languages: AIMMS, AMPL, GAMS, and MPL;
- links to Excel through their Analytic Solver and Solver SDK products.

The Gurobi Optimizer also includes several features to support the building of optimization models including:

- flexible prioritization for multi-objective models;
- general constraints such as MIN/MAX, ABS, AND/OR, and indicator constraints which help us to avoid turning commonly occurring constraints in linear equivalents;
- models with convex, piecewise–linear objective functions, to capture certain non–linear problems;
- arbitrary piecewise–linear objective functions, to make it easier to express this common modeling feature;
- distributed tuning, to speed up the exploration of parameter settings;
- cloud deployment and client–server computing.

The solver uses sophisticated branch-and-cut framework with presolve procedures, cutting planes, heuristics, and parallelism. In addition to this, it includes a long list of additional techniques. Some of them are a sophisticated branch variable selection, node presolve, symmetry detection, and disjoint subtree detection.

The process of development and deployment of optimization problems using Python API is vastly simplified using an Anaconda data science platform [3]. This non–commercial software can combine the Gurobi Optimizer with Python, Spyder Integrated Development Environment, and JupyterLab notebook interface. Gurobi Optimization, LLC, allows a full–featured university version of Gurobi Optimizer that can be installed on a single physical machine. This version has no limits on model size, and it is oriented explicitly for use by students, faculty, and staff at a recognized degree–granting academic institutions. The license must be activated while connected through the university network.

In the following codes we will see how mathematical formulations for $(r|H_p)$ HMPuPW (2.62)–(2.65) and auxiliary problem (2.71)–(2.74) can been coded using the corresponding Python API. Python is a well–known computer programming language, so it will not be presented here. The Python functions use the following parameters:

- `route_costs` for a data structure in which the costs of all possible routes are stored;
- `lead_net` for a data structure in which the given network for the leader is stored;
- `n` for the instance size;
- `r` for the number of hubs that follower needs to locate;
- `w` for a data structure in which the demands for all possible O–D pairs are stored;
- `Theta` for the sensitivity parameter Θ;
- `F` for the data structure which represents the follower's optimal objective value.

The Python definition of prototype function for solving $(r|H_p)$HMPuPW by Gurobi, using the Python API, is given in the Code 3.1. The body of this function is self–explanatory, as the appropriate Python comments and docstrings are given.

```
def medianoid(route_costs, lead_net, n, r, w, Theta):
    '''medianoid(np.array, tuple, int, int, np.array, float) ->
    (float, dict)

    Creates a Gurobi Python model for (r|H_p)HMPuPW according to
    its linear reformulation.  Solves the problem for a given
    instance.  Returns the objective value, hub locations, and
    established routes.
    '''

    # Creating the empty model.
    m = gpy.Model( "medianoid" )

    # Creating variables' dictionaries.
    y = {}
    zeta = {}

    # Adding variables.
    for i in range(n):
    y[i] = m.addVar( vtype=gpy.GRB.BINARY, name=f"y_{i}" )
    for i in range(n):
        for j in range(n):
            for k in range(n):
                for l in range(n):
                    zeta[i,j,k,l] = m.addVar(
                        name=f"zeta_{i}_{j}_{k}_{l}"
                    )

    # Adding constraint: the number of hubs is equal to r.
    m.addConstr( gpy.quicksum( y[k] for k in range(n) ) == r )

    # Adding constraints: only one route is allowed.
```

```
            for i in range(n):
                for j in range(n):
                    m.addConstr( gpy.quicksum(
                        zeta[i,j,k,l] for k in range(n) for l in range(n)
                        ) == 1
                    )
                    for l in range(n):
                        m.addConstr(
                            gpy.quicksum(
                                zeta[i,j,k,l] for k in range(n)
                            ) <= y[l]
                        )
                        m.addConstr(
                            gpy.quicksum(
                                zeta[i,j,l,k] for k in range(n)
                            ) <= y[l]
                        )

        # Adding the objective (profit) function.
        profit = gpy.quicksum(
            w[i][j] * \
            equilib_profits(route_costs, lead_net, zeta, i,j,k,l) \
            for i in range(n) for j in range(n) \
            for k in range(n) for l in range(n)
        )

        # Setting the objective.
        m.setObjective( profit, gpy.GRB.MAXIMIZE )

        # Solving the problem (calling the optimizer).
        m.optimize()

        # Extracting the optimal solution.
        objval = m.getObjective().getValue()
        hubs = m.getAttr('x', y)
        routes = m.getAttr('x', zeta)

    # Returning the optimal solution.
    return objval, hubs, routes

```

Code 3.1: Python code for the linear variant of $(r|H_p)$HMPuPW model.

Solving the auxiliary model is specified with a Python prototype function `auxiliary` in a Code 3.2. As we can see, the parameter list is almost the same as in the previous function, except that we have a `foll_net` and `F` parameters, representing the follower's network and its optimal profit, respectively.

```
def auxiliary(route_costs, lead_net, foll_net, n, r, w, Theta, F):
    '''auxiliary(np.array, tuple, tuple, int, int, np.array,
    float) -> (dict, dict)

    Creates a Gurobi Python model for the auxiliary subproblem
    according to its linear reformulation. Solves the problem
    for a given instance. Returns the optimal hub locations and
    established routes.
    '''

    # Extracting medianoid hubs.
```

```
    y_star = foll_net[0]

    # Creating the empty model.
    m = gpy.Model( "auxiliary" )

    # Creating variables' dictionaries.
    y = {}
    zeta = {}

    # Adding variables.
    for i in range(n):
    y[i] = m.addVar( vtype=gpy.GRB.BINARY, name=f"y_{i}" )
    for j in range(n):
        for k in range(n):
            for l in range(n):
                zeta[i,j,k,l] = m.addVar(
                    name="zeta_{i}_{j}_{k}_{l}"
                )

    # Adding constraint: the number of hubs is equal to r.
    m.addConstr( gpy.quicksum( y[k] for k in range(n) ) == r )

    # Adding constraints: only one route is allowed.
    for i in range(n):
        for j in range(n):
            m.addConstr( gpy.quicksum(
                zeta[i,j,k,l] \
                for k in range(n) \
                for l in range(n) ) == 1
            )
            for l in range(n):
                m.addConstr(
                    gpy.quicksum(
                        zeta[i,j,k,l] for k in range(n)
                    ) <= y[l]
                )
                m.addConstr(
                    gpy.quicksum(
                        zeta[i,j,l,k] for k in range(n)
                    ) <= y[l]
                )

    # Adding the auxiliary constraint.
    m.addConstr(
        gpy.quicksum(
            w[i][j] * \
            equilib_profits(lead_net, zeta, i,j,k,l) * \
            zeta[i,j,k,l] \
            for i in range(n) for j in range(n) \
            for k in range(n) for l in range(n)
        ) >= F, name="auxiliary_constraint"
    )

    # Adding the objective (profit) function.
    leader_profit = gpy.quicksum(
        w[i][j] * equilib_profits(
             route_costs, lead_net, zeta, i,j,k,l
          ) \
        for i in range(n) for j in range(n) \
        for k in range(n) for l in range(n)
```

```
    )

    # Ignoring those solutions that share the same hub backbone
    # as medianoid  problem
    m.addConstr(
        gpy.quicksum(
            y_star[k] * y[k] for k in range(n)
        ) <= r - 1
    )

    # Setting the objective.
    # The leader has optimistic expectations, i.e.
    # the follower is behaving benevolently.
    m.setObjective( flow_income, gpy.GRB.MAXIMIZE )

    # Solving the problem (calling the optimizer).
    m.optimize()

    # Extracting the optimal solution.
    hubs = m.getAttr('x', y)
    routes = m.getAttr('x', zeta)

    # Returning the optimal solution.
    return hubs, routes

```

Code 3.2: Python code for the linear variant of auxiliary model.

Remark 3.1. There is no need to return the optimal objective value in `auxiliary` Python function, as it is equal to F.

An example of Gurobi Optimizer log–file is given in the following lines:

```
Gurobi Optimizer version 9.0.2 build v9.0.2rc0 (win64)
Optimize a model with 946 rows, 4510 columns and 14410 nonzeros
Model fingerprint: 0xf13ca6bb
Variable types: 4500 continuous, 10 integer (10 binary)
Coefficient statistics:
Matrix range      [1e+00, 1e+00]
Objective range   [3e-02, 2e+01]
Bounds range      [1e+00, 1e+00]
RHS range         [1e+00, 3e+00]
Presolve time: 0.02s
Presolved: 946 rows, 4510 columns, 14410 nonzeros
Variable types: 4500 continuous, 10 integer (10 binary)
Found heuristic solution: objective 73.3523777
Found heuristic solution: objective 103.3777838


Root relaxation: objective 1.132394e+02, 445 iterations,
```

```
0.00 seconds

Nodes    |    Current Node    |     Objective Bounds      |     Work
Expl Unexpl | Obj  Depth IntInf | Incumbent   BestBd  Gap | It/
Node Time

0    0  113.23935    0    5  103.37778  113.23935  9.54%    -   0s
H    0               0       112.76523  113.23935  0.42%    -   0s
0    0  113.06432    0    4  112.76524  113.06432  0.27%    -   0s
0    0  113.06432    0    5  112.76524  113.06432  0.27%    -   0s

Cutting planes:
Gomory: 3
MIR: 10
Relax-and-lift: 4

Explored 1 nodes (643 simplex iterations) in 0.10 seconds
Thread count was 8 (of 8 available processors)

Solution count 3: 112.765 103.378 73.3524

Optimal solution found (tolerance 1.00e-04)
Best objective 1.127652365437e+02, best bound 1.127652365437e+02,
gap 0.0000%
```

Particularly, a `medianoid` function is called up for a 10–node CAB instance (please, see [103]) with $\alpha = 0.4$, $\Theta = 9$, and the leader has chosen a p–HMLP for her network. For auxiliary model we have basically the same output structure of log–file.

Of course, these Python functions are customizable using different parameter values and computer programming techniques. The objective of this section is only to share a user experience on a tool that allows us to implement quite easily complex formulations.

However, a mathematical optimization computer application might also need the functionality for manipulating data. This "non–modeling" expressiveness of the Python API is called scripting, and it is used in three different situations:

- preprocessing;
- postprocessing;

- flow control.

When it comes to Codes 3.1 and 3.2, the need for preprocessing is seen in a helper function `equilib_profits`. It partially prepares the data which will be used by the Gurobi model, i.e., its role is to perform a part of the computation concerning Bertrand–Nash price equilibrium profits for a given leader's network.

Notably, we need to iterate over the set of all O–D pairs, and for each pair we iterate over the set of all possible routes. The Bertrand–Nash price equilibrium is computed for each route addressed, taking into account that we know the leader's route for that O–D pair. Afterward, equilibria give us the corresponding profit estimations for both competitors.

Remark 3.2. These computations require $O(n^4)$ steps, for instances of n nodes. Computing the Bertrand–Nash price equilibria and their corresponding profits for each leader's network that we examine is a computationally very exhaustive job. This part of the computation is essentially done in the preprocessing phase using parallel computing and efficient data structures that support so–called vectorized operations to reduce the model creation time.

As the above remark has meant to suggest, it is out of the scope to go into details of `equilib_profits` function. Therefore, it will not be presented, i.e., only a proposed accompanying docstring in Code 3.3.

```
equilib_profits(...)
equilib_profits(route_costs, leader_network, zeta, i,j,k,l) ->
(float, float)

Compute a Bertrand--Nash price equilibrium profits for a given
follower's route i->k->l->j, concerning a given leader_network.

```

Code 3.3: Docstring for the equilib_profits function.

The postprocessing is concerned with working on or manipulating obtained solutions and objective function values in an appropriate way so that they can be used by a solution approach computer program written to solve the leader's problem. For example, the optimal objective function value obtained after solving the medianoid model is used in the auxiliary model.

The flow control is used to chain multiple models. Firstly, in developing an optimization computer application for the leader's problem, the auxiliary model must be solved after the medianoid one. Secondly, we have to solve the follower problem multiple times. Creating the model from scratch every time is a daunting task. A better approach could be to have a model, partially created, and on the as–needed basis to complete it by deleting and adding appropriate objective functions and required constraints.

3.2 Alternating Heuristic

The alternating heuristic (AH) was firstly presented in [11] by Bhadury et al. in 2003. Some of the recent applications can be found in [27, 79, 106]. The main idea is built on Hotelling's observation that at equilibrium, on a line, the duopoly facilities tend to cluster together at some central point of the market [65]. While d'Aspermont et al. [42] have shown that this observation does not hold in the case of variable prices, it does hold in the case of fixed and equal prices (please, inspect [11]). After that, it was further analyzed and exploited as a solution approach for many bi–level location problems (the reader is referred to a small survey of [30, 79, 106, 84, 108]).

Essentially, the AH is similar to coordinate descent (ascend). The same underlying idea is presented in several widely popular heuristics: Blahut–Arimoto algorithm in coding theory [14, 7], the expectation–maximization algorithm in statistics [39], the concave–convex procedure in global optimization [127], or k–means in machine learning [85, 50, 87]. The description of AH for our problem is presented in Algorithm 1.

The instance is represented as a pair of finite associative arrays (x, y^*), according to the description in Section 2.1. The size of array is $|N|$. In the algorithm, x^i represents the leader's network, during the i-th iteration. Accordingly, y^i represents the corresponding follower's network.

The AH starts with the initial incumbent hub and spoke topology, which is obtained by solving the well known p–median hub location problem (p–HMLP), as a natural choice, taking into account that the solution of p–HMLP is classified in the same way as the leader's problem when it comes to the hub backbone size and spoke allocation. Furthermore, the p–HMLP is well suited for a monopolistic situation, it is simple enough to implement, and it can be solved in a reasonable amount of time using the state-of-the-art solver like Gurobi Optimizer. Therefore, we can think of AH as a natural approach to estimate the evolution of market competition for this setting.

Remark 3.3. The AH can start from the arbitrary feasible hub and spoke network.

In the i-th iteration, $(r|H_p)$HMPuPW is solved *exactly*, utilizing the proposed reformulations (2.62)–(2.65).

Each *iteration* corresponds to one callback of the exact algorithm for the follower's problem. In the same iteration, the roles are switched, i.e., for the follower's obtained solution, the leader takes his role. This mechanism is portrayed at line 9 of Algorithm 1. Note that the leader's profit is computed after the follower's one.

The termination condition is a cycle detection. It is obvious that this procedure would converge because the number of all semi–feasible solutions is finite. Moreover, this way, the AH could capture a pure Nash equilibrium, if it exists. Because $(r|H_p)$HMPuPW and auxiliary problem can be exactly solved (utilizing the proposed reformulations), AH returns a feasible solution, not just semi–feasible one.

Remark 3.4. For some problem instances, cycles can be very long, and therefore in the literature, it is not unusual to include the upper bound on the number of iterations in the termination criterion.

The AM is not solved during the alternation phase, i.e., inside the while–loop. Solving AM inside the while–loop could lead to undesirable long running times, especially if the execution time of AM is generally long. Therefore, the AM is solved only at the final step of AH. If the time limit exceedance is taken as a termination condition, the solving of AM inside the while–loop could lead to a reduced search, thus affecting the solution quality.

Algorithm 1: `AH — Alternating Heuristic`

Data: instance arguments
Result: a feasible solution (x, y^*)
1 $i \leftarrow 1$
2 $x^i \leftarrow$ the optimal network of p–HMLP
3 **while** $x^i \notin \{x^j \mid j \in \{1, \dots, i-1\}\}$ **do**
4     $H_p \leftarrow$ the hub backbone of temporary leader's network x^i
5     $y^{i*} \leftarrow$ the exact solution of $(r|H_p)$HMPuPW
6     $z^i_{F*} \leftarrow$ the follower's profit
7     $z^i_L \leftarrow$ the leader's profit
8     $i \leftarrow i + 1$
9     $x^i \leftarrow y^{(i-1)*}$
10 $K \leftarrow \arg\max_{j \in \{1,\dots,i\}} z^j_L$
11 pick a random $k \in K$
12 $x \leftarrow x^k$
13 $z_L \leftarrow z^k_L$
14 $z_{F*} \leftarrow z^k_{F*}$
15 $y^* \leftarrow$ the solution of corresponding AM for x and z_{F*}

Remark 3.5. If a semi–feasible solution (x, y^*) constitutes a pure Nash equilibrium, solving AM (line 15) can give a feasible solution that does not need to be a Nash equilibrium.

Remark 3.6. AH does not have an optimizing characteristic. In other words, we can not guarantee that arbitrary extension of any computational resource will eventually help us to find the optimal solution. This observation is true even if the initial leader's network is randomly chosen.

3.3 Variable Neighborhood Search

Variable neighborhood search (VNS), firstly presented by Mladenovic and Hansen in [98], systematically varies neighborhoods and the corresponding change in the landscape during the search for the global optimum. The VNS algorithms showed excellent performance and for some location problems they are currently state-of-the-art, e.g., for the r–allocation hub location problem [122], single and multiple allocation p–hub maximal covering problems [69], or for bi–level location problems like a discrete $(r|p)$–centroid problem [32]. A constructive survey of VNS principles and applications is the recent paper by Hansen et al. [62].

The VNS relies upon the following three observations:

(O1) a local maximum concerning one neighborhood structure is not necessarily a local maximum for another neighborhood structure;

(O2) a global maximum is a local maximum for all possible neighborhood structures;

(O3) for many problems, local maximums for one or several neighborhoods are relatively close to each other.

The first two observations are theoretical and the last one is empirical. For instance, in the case of HLPs, local and global optimal solutions may have several hubs located at the same places, but their indices are usually unknown. Therefore, the systematic exploration of a local optimum neighborhood is required, until a better solution is found (hopefully the global optimum).

In a given VNS approach, outlined in Algorithm 2, the Basic VNS scheme (BVNS) is followed. The main parameters are:

- $t_{\max}$ which represents the time limit;
- $k_{\max}$ which represents the iteration limit for the inner loop (maximal k–neighborhood of shake operation);
- x which is an additional boolean variable that serves as a flag: if $x = \top$, then $(r|H_p)$HMPuPW is solved exactly during the local search (LS), otherwise the corresponding linear programming relaxation (LPR) is solved (similarly to [32]).

The initial solution is obtained by an alternating heuristic (AH), as denoted on line 1 of Algorithm 2. Also, in the beginning, the flag variable x is set to $\perp$, i.e., LPR should be initially solved during LS.

Because there exists a possibility of being stuck in a local optimum, a probabilistic `shake` procedure is utilized.

In order to disseminate the search, BVNS uses neighborhoods of increasing cardinality to find better local optimum during LS. Suppose that s is an arbitrary solution and $\mathcal{N}_k$ ($k \in \{k_{\min}, \ldots, k_{\max}\}$) is a finite sequence of neighborhood structures. Then $\mathcal{N}_k(s)$ is defined as the set of solutions in the kth neighborhood of s.

A solution s'' is better than s if and only if s'' yields a higher profit. The improvement check at line 9 is always done for feasible solutions, i.e., we solved the corresponding

Algorithm 2: `BVNS` — `Basic Variable Neighborhood Search`

Data: instance arguments, $t_{\max}$, $k_{\max}$
Result: a feasible leader's solution s

1 $s \leftarrow$ `AH`
2 $x \leftarrow \bot$
3 **while** $t \leq t_{\max}$ **do**
4     $k \leftarrow 1$
5   **while** $k \leq k_{\max}$ **do**
6         $s' \leftarrow \texttt{shake}(s, x, k)$
7         $s'' \leftarrow \texttt{LS}(s', x, k)$
8         $k \leftarrow k + 1$
9   **if** *s'' is better than s* **then**
10             $s \leftarrow s''$
11             $k \leftarrow 1$
12             $x \leftarrow \bot$
13     $x \leftarrow \neg x$

$(r|H_p)$HMPuPW and AM exactly. The flag x switches from $\top$ to $\bot$ if an improvement is observed (line 12 of Algorithm 2), or if k becomes greater than the limit value $k_{\max}$ (line 13 of Algorithm 2). Solving LPR reduces time and hopefully, relaying on the results in [32], is a helpful in guiding the search process. It is easy to see that switching does not affect the effectiveness.

Following the paradigm "less is more", popularized by Mladenović [99], we formally present the classic k*–swap neighborhood* that Algorithm 2 is based on:

$$\texttt{swap}(k, H) \triangleq \{H' \subset N \ : \ |H'| = |H| \wedge |H' \backslash H| = k\}, \tag{3.1}$$

where $k \in \mathbb{N}$.

Definition 3.1. An ordered pair $(H,\ H \backslash K \cup K')$ is called a k*–swap move* on a non–empty subset $H \subset N$ if $K \subseteq H$, $K' \subset N \backslash H$, $K \cap K' = \emptyset$, and $|K| = |K'| = k$.

A k–swap move can be seen as a ternary operation. Applying it to a subset $H \subset N$, $K \subseteq H$, and $K' \subseteq N \backslash H$ results in a new subset $H' \subset N$, i.e., $H \neq H'$. In fact, there are many ways to formally define/represent a k–swap move, all essentially equivalent. Therefore, it is sensible to also assume and go along with the intuitive understanding of this term.

Definition 3.2. We say that H and H', both subsets of N and $|H \cap H'| < |H| = |H'|$, are in *a* k*–swap relation*, denoting it as $H \bowtie_k H'$, if there exist $K \subseteq H$ and $K' \subseteq H'$, such that $|K| = |K'| = k$, $K \cap K' = \emptyset$ and $H' = H \backslash K \cup K'$, i.e., H' can be obtained from H by a k–swap move.

It is easy to see that this relation is irreflexive, symmetric, and non–transitive.

Definition 3.3. We call the finite sequence of ordered pairs $(H, H'), (H', H''), \ldots, (H^{(m-1)}, H^{(m)})$ *a sequence of* k*–swap moves* (or *a* k*–swap sequence*) of length m if

$H \bowtie_k H' \wedge H' \bowtie_k H'' \wedge \ldots \wedge H^{(m-1)} \bowtie_k H^{(m)}$. H and $H^{(m)}$ are called *the start* and *end* of this sequence, respectively.

Remark 3.7. In a k–swap sequence, non–subsequent subsets can be equal, i.e., $H \neq H'$ and $H' \neq H''$ must hold, but it may happen that $H = H''$.

Definition 3.4. The minimal length k–swap sequence which has H as the start and H' as the end is called *a k–swap distance* between H and H' and it is denoted as $d_k(H, H')$. Formally, we take that:

- $d_k(H, H) = 0$;
- $d_k(H, H') = \infty$, when there exists no k–swap sequence between H and H'.

Remark 3.8. It is easy to see that $d_k : \mathcal{P}(N)^2 \to \mathbb{N}_0$ has the following properties:

- it is positive between two different subsets and it is precisely zero when it maps a subset to itself (identity of indiscernibles);
- it is symmetric;
- it satisfies the triangle inequality (sub-additivity).

In other words, d_k is a metric and (N, d_k) is a metric space.

Proposition 3.9. *Let N be a non–empty set and $|N|< 2p$ for some $p \in \mathbb{N}$. Then, for each pair of subsets $H, H' \subset N$, such that $|H| = |H'| = p$ and $H \neq H'$, there exists some $k \leq \lceil \frac{p}{2} \rceil$ so that $d_k(H, H') \leq 2$.*

Proof. We know that $H \cup H'$ can be represented as a union of disjoint sets $H \cup H' = K \sqcup K^{int} \sqcup K'$, where $K = H \backslash H',\ K^{int} = H \cap H',\ K' = H' \backslash H$. From the proposition conditions we realize that $K^{int} \neq \emptyset$. If $|K|\leq \lceil \frac{p}{2} \rceil$, then $d_k(H, H') = 1$. The rest of this proof, i.e., the case $|K|> \lceil \frac{p}{2} \rceil$, is based on the parity of $|K|$ (or equivalently $|K'|$).

Assume that $|K|$ is even. Let $k = \frac{|K|}{2}$, $K = K_1 \sqcup K_2,\ |K_1|= |K_2|,\ K' = K_1' \sqcup K_2'$, and $|K_1'|= |K_2'|$. We can create a sequence of two k–swap moves $(H, \tilde{H}),\ (\tilde{H}, H')$. For the first move, we can take that $\tilde{H} = H \backslash K_1 \cup K_1'$, and for the second one $H' = \tilde{H} \backslash K_2 \cup K_2'$.

Assume that $|K|$ is odd. Let $k = \lceil \frac{|K|}{2} \rceil$, $K = K_1 \sqcup K_2,\ |K_1|= |K_2|-1,\ K' = K_1' \sqcup K_2'$, and $|K_1'|-1 = |K_2'|$. Again, we can create a sequence of two k–swap moves: $(H, \tilde{H}),\ (\tilde{H}, H')$. For the first move, we can take that $\tilde{H} = H \backslash (K_1 \cup \{h\}) \cup K_1'$ and for the second one $H' = \tilde{H} \backslash K_2 \cup (K_2' \cup \{h\})$, where $h \in K^{int}$. It is easy to see that these set operations define the valid k–swap moves.

In both cases discussed above, k was less than $\lceil \frac{p}{2} \rceil$ and we composed k–swap sequences from two moves, at most. $\square$

Next, we address the case when $N = 2p$.

Proposition 3.10. *Let N be a non–empty set and $|N|= 2p$ for some even $p \in \mathbb{N}$. Then, for each pair of subsets $H, H' \subset N$, such that $|H| = |H'| = p$ and $H \neq H'$, there exists some $k \leq \frac{p}{2}$ so that $d_k(H, H') \leq 2$.*

Proof. If $|H \cup H'|< 2p$ then we can apply the previous proposition. Therefore, we only consider the case when H and H' are disjoint. Nevertheless, in this case, it is easy to see that for $k = \frac{p}{2}$ we can create a proper k–swap sequence of length two at most. □

Remark 3.11. Obviously, in the previous theorem k is also less then or equal to $\lceil \frac{p}{2} \rceil$.

Unfortunately, a similar statement to the previous one does not hold when p is odd.

Proposition 3.12. *Let N be a non–empty set and $|N|= 2p$ for some odd $p \in \mathbb{N}$ and $p \geq 3$. Then, for each pair of disjoint subsets $H, H' \subset N$, such that $|H|= |H'|= p$, their k–swap distance $d_k(H, H')$ is greater than 2, for all $k \leq \lceil \frac{p}{2} \rceil$.*

Proof. Assume the opposite. In the first k–swap move, a two disjoint subsets $K \subset H$ and $K' \subset N'$ of cardinality $\lceil \frac{p}{2} \rceil - 1$ are swapped, i.e., we obtain two new disjoint subsets $\tilde{H} = H \backslash K \cup K'$ and $\tilde{H}' = H' \backslash K' \cup K$. In the next and last move, because we assume the opposite, we need to swap $H \backslash K$ (a subset of $\tilde{H}$) with $H' \backslash K'$ (a subset of $\tilde{H}'$), which have $\lceil \frac{p}{2} \rceil$ elements each. If $k \leq \lceil \frac{p}{2} \rceil - 1$, $|H \backslash K|= |H' \backslash K'| \geq \lceil \frac{p}{2} \rceil$, but we are allowed only to swap $k \leq \lceil \frac{p}{2} \rceil - 1$, which is insufficient. Otherwise, $|H \backslash K|= |H' \backslash K'|= \lceil \frac{p}{2} \rceil - 1$, but we have to swap $k \leq \lceil \frac{p}{2} \rceil$, which is too much. □

For $|N|$ larger than $2p$, again, we prove the statement similar to Propositions 3.9 and 3.10.

Proposition 3.13. *Let N be a non–empty set and $|N|> 2p$ for some $p \in \mathbb{N}$. Then, for each pair of subsets $H, H' \subset N$, such that $|H|= |H'|= p$ and $H \neq H'$, there exists some $k \leq \lceil \frac{p}{2} \rceil$ so that $d_k(H, H') \leq 2$.*

Proof. If $|H \cup H'|< 2p$, we can apply Proposition 3.9 on $N' = H \cup H'$. To prove the statement for $|H \cup H'|= 2p$, we focus our attention to the parity of p. For even p we can apply Proposition 3.10, again on a $N' = H \cup H'$. However, if p is odd, let $k = \lceil \frac{p}{2} \rceil$ and $H = K_1 \sqcup K_2,\ |K_1|= |K_2|+1,\ H' = K'_1 \sqcup K'_2,\ |K'_1|+1 = |K'_2|$. Because $|N|> 2p$ we have at least one element $i \in N$ that satisfies the following: $i \notin H$ and $i \notin H'$. Now, we can easily create a sequence of two k–swap moves: $(H, \tilde{H}),\ (\tilde{H}, H')$. For the first move, we can take that $\tilde{H} = H \backslash K_1 \cup (K'_1 \cup \{i\})$. Regarding the second one, we have that $H' = \tilde{H} \backslash (K_2 \cup \{i\}) \cup K'_2$. □

In a way, these propositions provide a theoretical ground for the utilization of k–swap neighborhood in Algorithm 2. With `shake` and LS, we can make two consecutive k–swap moves. If p is not odd, for $k_{\max}$ we can take $\lceil \frac{p}{2} \rceil$. Otherwise, we need to take into account what Proposition 3.12 states.

In reality, we do not expect that the number of non–hub nodes is equal or smaller than the hub backbone, i.e., usually, we have that $p \ll |N|$. Therefore, in our Algorithm 2, $k_{\min} = 1$ (line 4) and $k_{\max} = \lceil \frac{p}{2} \rceil$. However, if needed, it is easy to extend the set of considered neighborhoods by allowing a p–swap move, although, **(O3)** suggests that we could also ignore that. Another reason for $k_{\max} = \lceil \frac{p}{2} \rceil$ as an upper counter bound lies in a theoretical suggestion, which states that the currently inspected neighborhood should be larger than the previous one. Of course, when p is odd, the last neighborhood does not have to be greater than the previous one, but taking the floor instead of the ceiling would result in missing to inspect the largest neighborhood.

3.3.1 Objective Function

The reason behind solving the LPR is that the evaluation of the leader's objective function is an NP–hard problem, according to Theorem 2.28. In situations like this, the objective function estimation is a plausible substitute. A natural way to address that is to exercise the corresponding LPR of the follower's model. Differently, to approach in [32], the estimation of leader's profit is affected by the logit model and equilibrium pricing. Moreover, if the follower establishes a different route for some O–D pair, the leader's equilibrium price changes accordingly.

Example 3.1. Assume that for an O–D pair (i, j) we have following fractional values for the follower's routes in the corresponding LPR: $\varsigma_{ij,33} = 0.5$ and $\varsigma_{ij,44} = 0.5$. The respective route costs are $c^F_{ij,33} = 2$ and $c^F_{ij,44} = 1$. At the same time, the leader's network is fixed and $\rho_{ij,24} = 1$, where the corresponding route cost is $c^L_{ij,24} = 3$. The price sensitivity is $\Theta = 3$.

The equilibrium price Eq. (2.28)–(2.29) yield that the corresponding prices are:

- $q^*_{ij,33} \approx 3.16$ and $t^*_{ij,24} \approx 3.47$;
- $q^*_{ij,44} \approx 2.89$ and $t^*_{ij,24} \approx 3.40$.

As we can see, for the same route $i \to 2 \to 4 \to j$ we have two different prices regarding $t^*_{ij,24}$. ▲

This example tells us that we can not directly use (2.3) to estimate the leader's profit by the LPR of $(r|H_p)$HMPuPW, i.e., we can not expect to have one leader's price as the universal best response to different follower's prices. It is not immediately clear what would be a suitable substitute for evaluation. Here, the approach is to build it cumulatively.

After preprocessing, we know for each O–D pair (i, j) the equilibrium prices $(t_{ij,kl}, q_{ij,st})$ of corresponding route pairs $(\rho_{ij,kl}, \zeta_{ij,st})$. In order to build the substitute, we introduce a function $T_{ij,kl}(s,t) : N^2 \to \mathbb{R}$ (for all $i, j, k, l \in N$)

$$T_{ij,kl}(u,v) = \begin{cases} t_{ij,kl}, & \text{if } q_{ij,uv} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

If the relaxed follower's solution has a partially established route $\tilde{\zeta}_{ij,uv}$, then the corresponding equilibrium leader's price should be taken into account. Constraint (2.14) stipulates that $\sum_{u,v} \tilde{\zeta}_{ij,uv} = 1$. Therefore, our substitute for the leader's equilibrium price is given as

$$\tilde{t}_{ij,kl} = \sum_{u,v \in N} \tilde{\zeta}_{ij,uv} T_{ij,kl}(u,v). \tag{3.3}$$

When we have the LPR solution, the effect of follower onto the corresponding leader's market share can be written as

$$\tilde{\gamma}_{ij} = \sum_{u,v \in N} \tilde{\zeta}_{ij,uv} e^{-\Theta q_{ij,uv}} \tag{3.4}$$

Finally, the estimation of leader's profit is given as

$$\sum_{i,j,k,l \in N} w_{ij}(\tilde{t}_{ij,kl} - c_{ij,kl}) \frac{1}{1 + \tilde{\gamma}_{ij} e^{\Theta \tilde{t}_{ij,kl}}} \rho_{ij,kl}. \tag{3.5}$$

We will use this expression in our local search algorithm to estimate the leader's total profit for neighboring solutions.

The downside of the estimation is that it can be misleading. For example, we could inspect two leader's networks s_1 and s_2 and the LPR–based estimation could favor s_1, while actually s_2 is a better one. Therefore, if the improvement is not observed for some time (e.g., during the whole inner loop), it is natural to switch from estimations to evaluations. On the other hand, the exact solution of a follower's problem is quite costly, so switching back should be done after a while. This mechanism is implemented on the line 13 of Algorithm 2.

The AM is always solved exactly, as it is unclear how its linear programming relaxation would provide any useful information. On the other hand, solving the AM can significantly affect the algorithm execution time. In our design, the AM is solved only at the end of AH and at the end of local search. This approach enables a push, from time to time, towards the optimistic leader's expectation and, at the same time, resources are not vastly spent on the AM.

3.3.2 Shaking

At each iteration, the VNS algorithm applies to the leader's current solution s a probabilistic procedure shake(s, x, k). This procedure replaces k randomly chosen hubs of the leader by some other randomly chosen nodes, as presented in Algorithm 3.

Algorithm 3: `shake`

```
Data: s, x, k
Result: a (semi—)feasible solution s'
H_p ← a hub backbone of leader's network in solution s
H'_p ← a hub backbone from swap(k, H_p) selected at random
ρ' ← allocate(H'_p)
if ¬x then
    solve the corresponding LPR of (r|H'_p)HMPuPW
    estimate the leader's profit
else
    solve exactly the corresponding (r|H'_p)HMPuPW
    compute the leader's profit
compose a new (semi—)feasible solution s'
```

The allocation procedure allocate(H'_p) is applied to the new set of leader's hubs H'_p to compose a new leader's solution network ρ', according to the statement of Theorem 2.35. At the end of shaking, the $(r|H_p)$HMPuPW of its corresponding LPR is solved for this new leader's network in order to compute or estimate the leader's profit. The decision about solving type is made on the basis of flag x value. Only for $x = \top$ the obtained solution s' will be feasible. It must be noted that for $x = \perp$ the solution is not feasible, nor semi–feasible.

The allocation is done following the indications about the preferred routes. Therefore, for every O–D pair (i, j), the established routes are those that have the lowest cost. The pseudo–code of procedure is given in Algorithm 4.

Remark 3.14. The decision about solving AM is shifted to LS procedure. This is why the composed solution can be semi–feasible.

Algorithm 4: `allocate`

```
Data: H
Result: a hub and spoke network
for every O–D pair (i, j) do
    establish the lowest cost route using hubs from H
```

3.3.3 Local Search

The LS procedure is described in Algorithm 5. It is based on the swap neighborhood with one major difference — the optimal policy for the well–known secretary problem is used as a stopping rule. In the literature, this policy is also called a $\frac{1}{e}$–stopping rule or the

$\frac{1}{e}$–law of best choice. The reason for choosing this rule is based on the following three observations:

- the best improvement strategy is time expensive for discrete bi–level optimization problems;
- if the starting solution is of low quality (e.g., after the `shake` procedure), the first improvement strategy could easily find a better solution, but still a low–quality one;
- if the local search starts with a near (locally) optimal solution, the first improvement strategy will require almost an equivalent amount of time as the best improvement.

On the other hand, we can think about the hubs as rankable applicants in the secretary problem. Thus, to maximize the probability of selecting the best applicant, i.e., the best hub for swap, the $\frac{1}{e}$–law of best choice yields the optimal policy. This means that only $\lceil \frac{p}{e} \rceil$ hubs, randomly chosen, should be considered for swapping with non–hub nodes. Particularly, we have that if $p \in \{1, 2\}$ then $\lceil \frac{p}{e} \rceil = 1$, if $p \in \{3, 4, 5\}$ then $\lceil \frac{p}{e} \rceil = 2$, if $p \in \{6, 7, 8\}$ then $\lceil \frac{p}{e} \rceil = 3$, and so on. The set of randomly chosen $\lceil \frac{p}{e} \rceil$ hubs is denoted as O (line 2 in Algorithm 5). Obviously, we have that $O \subseteq H_p$.

Algorithm 5: `LS` — `Local Search`

Data: a (semi–)feasible solution s', x, k
Result: a locally optimal feasible solution s''
1 $H'_p \leftarrow$ the leader's hub backbone in s'
2 $O \leftarrow \lceil \frac{p}{e} \rceil$ randomly chosen hubs from H'_p
3 **for each** $H''_p \in$ `coreswap`(k, H'_p, O) **do**
4 $\rho'' \leftarrow$ `allocate`(H''_p)
5 **if** $\neg x$ **then**
6   solve the corresponding LPR of $(r|H''_p)$HMPuPW
7   estimate the leader's profit for ρ''
8 **else**
9   solve exactly the corresponding $(r|H''_p)$HMPuPW
10   compute the leader's profit for ρ''
11 $\overline{\rho} \leftarrow$ the profit–best leader's network among all that were examined
12 $\overline{H_p} \leftarrow$ the hub backbone of $\overline{\rho}$
13 **if** $\neg x$ **then**
14 solve exactly $(r|\overline{H_p})$HMPuPW
15 solve the corresponding AM

It is easy to see that O yields a specific subset of swap–neighborhood, which is, in fact, the target neighborhood of our local search. In other words, O can be seen as a core of the new neighborhood, which we formally define as

$$\texttt{coreswap}(k, H, O) \triangleq \{H' \subseteq \texttt{swap}(k, H) \mid |H' \backslash O| = |H \backslash O| + 1\}. \tag{3.6}$$

After finding the best solution, according to the value of flag x, we check if the real improvement is observed. This means that for leader's network $\tilde{s}$ we solve the corresponding $(r|H_p')$HMPuPW exactly. Thus, as we already pointed out, in Algorithm 2 the improvement check at line 9 is done for feasible solutions.

Remark 3.15. Obviously, one way to quickly improve the best solution time of BVNS is to skip the calls for solving the corresponding AM in the LS procedure, especially if it is unlikely that AM solutions will differ from the corresponding $(r|H_p)$HMPuPW ones. However, this bald assumption does not reflect the true nature of our bi–level problem. Our goal is not to thirstily design a solution approach that will be fast on some particular instance sets. Instead, our goal is to design an algorithm that will be fast, systematic in its search process, and which will address our bi-level problem's nature in the right way. Solving AM solely at the end of BVNS could be seen as a cosmetic instruction and not a very useful circumvention of the issue with the follower's behavior. In our opinion, it is better from time to time to solve AM (i.e., in the LS phase), to check if the AM solution differs from $(r|H_p)$HMPuPW one, while also being reasonably confident that the search is pushed in the right direction.

Chapter 4

Computational Experiments

«Нет области математики, как бы абстрактна она ни была, которая однажды не смогла бы быть применена к явлениям реального мира.»

Николай Иванович Лобачевский

The computational experiments are conducted using the instances generated from the well–known CAB dataset, composed and published by O'Kelly in 1987 [103]. All mathematical programs were implemented in Python 3.8 (as a part of Anaconda package). Gurobi Optimizer 9.0 was used as an integer linear programming solver, installed on the Windows 10 operating system. The hardware platform for computation was Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz with 24.0 GiB of DDR4 RAM. All figures in this work were produced using a Python package matplotlib [66]. The basic figure concerning the USA map corresponding to CAB data set is given in the Appendix.

When it comes to the size of instance, three batches were considered: 15, 20, and 25 node instances. The discount factors χ and δ were set to 1, and α took values from the set $\{0.2, 0.4, 0.6, 0.8\}$, usual for the CAB data set. The price sensitivity factor Θ took values from a set $\{3, 6, 9, 12, 15\}$, which are modified values from [86]. The difference is that the values are integer and equally separated in range from 3 to 15.

When it comes to the hub backbone size, the number of hubs for both competitors is equal (i.e., $r = p$). The time limit $t_{\max}$ was set as:

- when $|N|= 15$, $t_{\max} = 45$ minutes;
- when $|N|= 20$, $t_{\max} = 1.5$ hours;
- when $|N|= 25$, $t_{\max} = 2$ hours.

In total, 240 instances were examined. For every one of them, our VNS algorithm has been executed 10 times, i.e., 2400 tests were performed. The result concerning running times and objective values, are given in Tables C.1–C.3, in the Appendix.

Besides this, a situation in which the leader ignores the follower is also addressed. Two classic hub–location networks are considered as the leader's strategic option: p–HMLP and p–HCLP. For these networks, the VNS algorithm has not been executed. Only the optimal follower's solution was determined. Notably, 480 tests were performed.

The case $r \neq p$ is usually not presented in the research literature. Therefore, it is not addressed in this thesis. However, for illustration purposes, some computational experiments are done, and the results can be found in Harvard's Dataverse repository [31].

In the sections that follow, we are going to investigate, in terms of descriptive statistics, the results concerning the stability of VNS approach, comparison of VNS and AH solutions, the effect of particular instance parameters, how AM affects the leader, and what happens when the leader ignores the follower.

Remark 4.1. The result of computational experiments are not values of random variables.

4.1 Stability of VNS Solution Approach

We address the stability of VNS solution approach by two measures of dispersion on the instance level: range and mean absolute deviation (MAD).

The summary of range data is given in Table 4.1. The first column ($|N|$) represents the instance size — a batch. In the following three columns, the average range, the average range percentage, and maximal range are provided, respectively. We can see that all three measures increase gradually, as $|N|$ gets larger, and the average range percentage indicates good stability of our solution approach. Moreover, the ratio of average range percentages for $|N|= 25$ and $|N|= 20$ is around 1.25, which is much less than the corresponding ratios of average range (2.05) and maximal range (6.51).

Table 4.1: The stability of solution approach regarding range.

|N|	**avg. range**	**avg. range percentage (%)**	**max range**
15	1.23	0.003	50.47
20	11.97	0.009	215.21
25	24.56	0.012	1401.21

The summary of MAD data is given in Table 4.2. Here, we address the average of mean absolute deviation (average MAD) on the level of instance and maximal absolute deviation. The first column ($|N|$) represents the instance size, as in the previous table. In the following three columns, the average MAD (avg. MAD), the percentage of this average MAD compared to the smallest objective value returned by VNS for a given instance batch (avg. MAD percentage (%)), and the maximal absolute deviation (max abs. deviation), are presented, respectively. As we can see, although the maximal absolute deviation can be significant, the average MAD values are quite small. Also, we do not have gradual increase of values, like in the case of range.

Taking all this into account, we can say that our VNS algorithm shows/exercises a good stability.

Table 4.2: The stability of solution approach concerning MAD.

$\|N\|$	avg. MAD	avg. MAD percentage (%)	max abs. deviation
15	1.44	0.71	113.11
20	5.78	0.86	203.48
25	5.23	0.58	147.75

4.2 Comparison of AH and VNS Solutions

Because our VNS approach uses AH to obtain an initial solution, and AH can be used as a standalone method, it is reasonable to compare the quality of AH solutions with the VNS ones. For that, we need to calculate the differences in objective values between the VNS and AH solutions. For every difference, its corresponding percentage of VNS obtained value is computed. Because for every instance we did 10 runs of our VNS algorithm, the best value among them is taken for the percentage computation.

The histogram of these values is presented on Fig. 4.1. As we can see, the histogram is right–skewed.

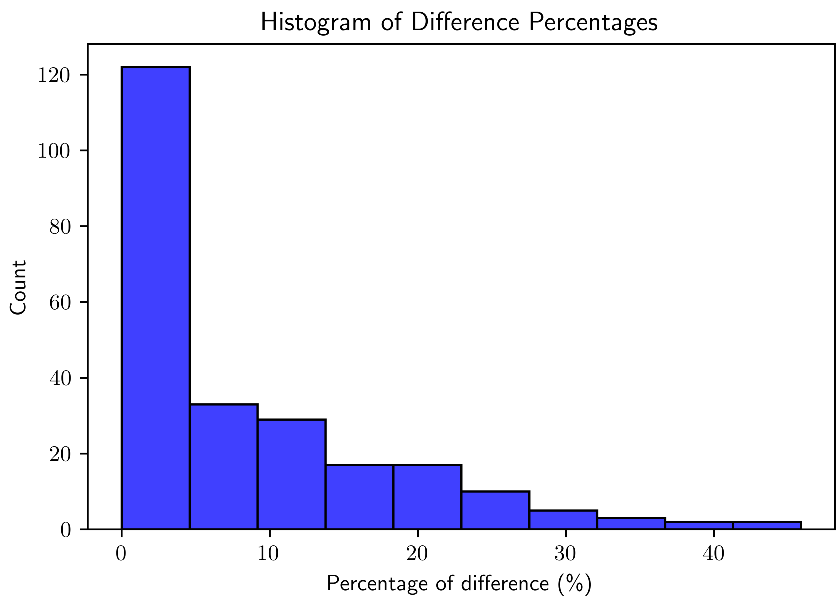


Figure 4.1: The distribution of percentages concerning the difference of leader's profits obtained by VNS and AH, compared to the best VNS leader's profit per instance. The number of bins is 10. The total amount of data processed is 240.

As histogram can be biased, on Fig. 4.2 we can see an estimation of CDF, based on the empirical CDF (ECDF).

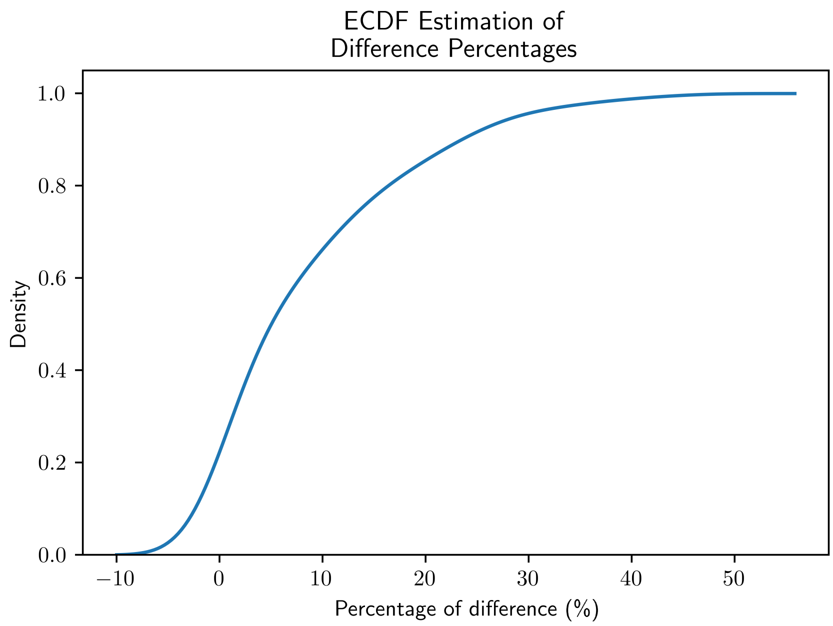


Figure 4.2: An estimation of CDF based on ECDF, regarding percentages of differences between leader's profits obtained by VNS and AH, compared to the best VNS leader's profit per instance.

In most cases, the difference between AH and VNS solutions is negligible in terms of the leader's profit. Thus, it is sensible to say that the AH represents a reasonably good heuristic approach for solving this problem, and a well–suited method for obtaining the initial solution in the VNS algorithm. Furthermore, from Tables C.1-C.3 we can see that the number of AH iterations per instance is not very large. The corresponding bar plot is given in Fig. 4.3. Mod and median of AH iterations number is 3, while mean is 5.058. The number of AH iterations is in range from 3 to 19, although it never happened that AH ended after 17 iterations.

Remark 4.2. Three iterations means that the initial leader's network was the best one that AH has found.

4.3 Effect of Θ and α Parameters

This part of the computational investigation is focused on the effects of Θ, α and hub backbone size to the following attributes of resulting solution:

- the leader's profit;
- the ratio of leader's and followers' profit;
- the leader's market share;
- the leader's net profit margin (NPM).

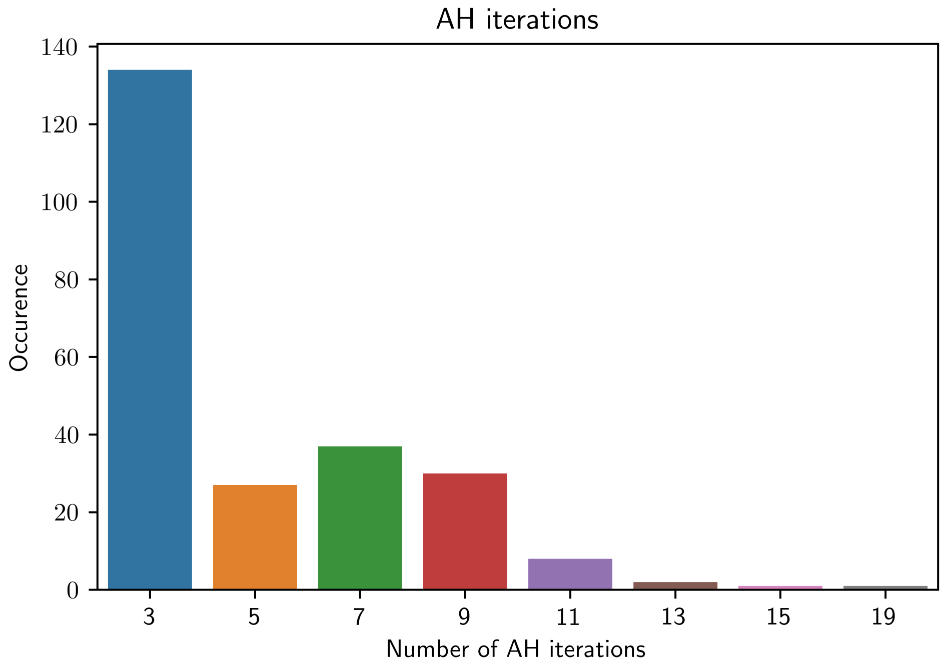


Figure 4.3: The bar plot of AH iterations.

The NPM is calculated in the following way

$$\text{NPM} = \frac{\text{objective value}}{\text{revenue}} \times 100\%. \tag{4.1}$$

The revenue is computed simply by ignoring the transportation costs

$$\text{revenue} = \sum_{i,j,k,l \in N} w_{ij} t^*_{ij,kl} \frac{\rho_{ij,kl} e^{-\Theta t^*_{ij,kl}}}{\sum_{u,v \in N} \rho_{ij,uv} e^{-\Theta t^*_{ij,uv}} + \sum_{u,v \in N} \varsigma^*_{ij,uv} e^{-\Theta q^*_{ij,uv}}}. \tag{4.2}$$

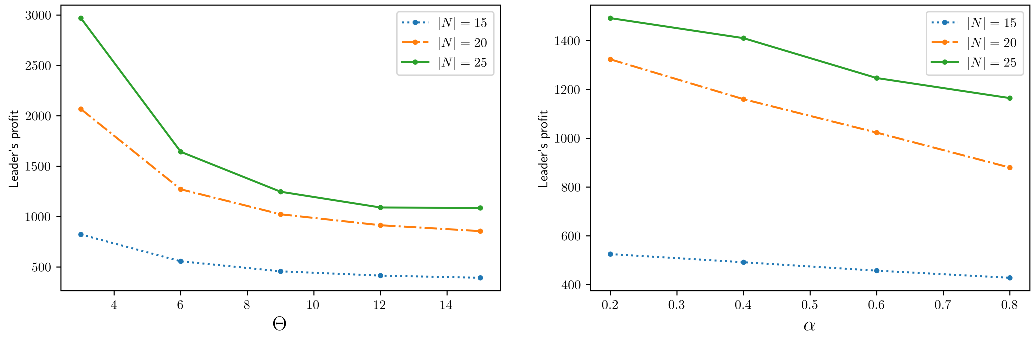
Except the leader's profit, the rest of attributes should not be strongly affected by a instance size. Therefore, for them, the data are aggregated according to the parameter values that are being inspected.

The effects of price sensitivity parameter Θ and inter–hub transportation discount α happen to be the most discernible. For some attributes, it is possible to observe almost a solid pattern, in all three instance batches. On the other hand, no solid patterns have been observed regarding the effect of hub backbone size.

Figure 4.4 is an illustrative example of observed patterns, i.e., how parameters Θ and α affect the leader's profit. For both sub–plots, the hub backbone size is taken to be 4.

On Fig. 4.5–4.9 we can see the depiction of how Θ affects the aforementioned attributes. Fig. 4.5 presents three point plots corresponding to our instance batches. This plot indicates that, in general, the leader can expect to make less profit if Θ gets increased. In a way, this indication and the curvatures are roughly predicted by Theorems 2.25 and 2.26. Also, as instances get larger, so to speak, the dispersion increases, too.

When it comes to the effect of Θ to the leader's market share, the box plot on Fig. 4.6 suggests that increase of Θ value results in lower expected leader's market share. Also, the



(a) (b)

Figure 4.4: The illustration of how parameters affect the leader's profit: (**a**) Θ and (**b**) α. Blue (dotted), orange (dash–dotted), and green (ordinary) lines correspond to instances of 15, 20, and 25 nodes, respectively. The size of hub backbone is 4. For the left sub–figure $\alpha = 0.6$, while for the right one $\Theta = 9$.

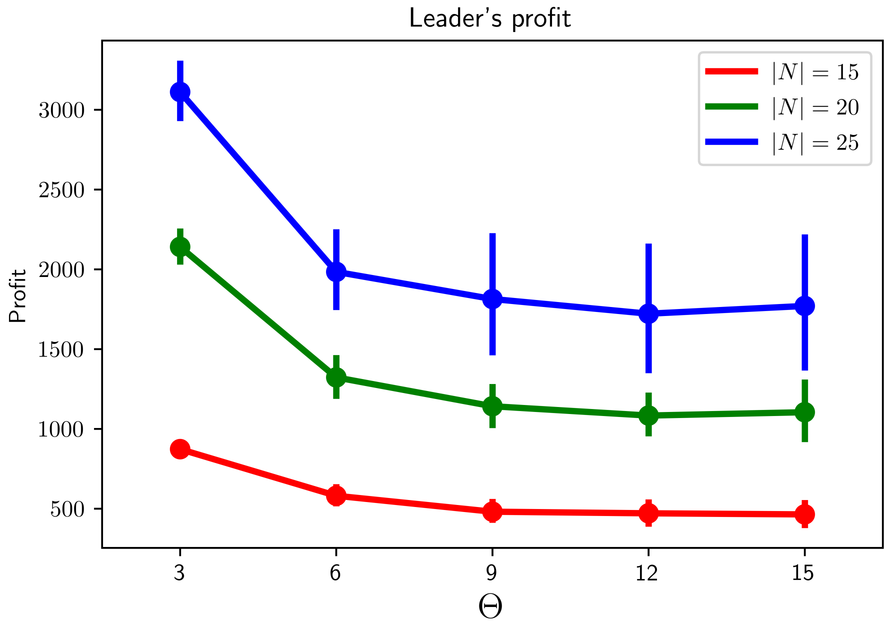


Figure 4.5: The effect of Θ on the leader's profit.

corresponding variability is increased. It can be seen that a low $\Theta = 3$ means the leader's market share will be close to 50%. On the other hand, it looks like that increment of Θ value leads to less predictable outcomes, i.e., the market share values are getting more dispersed. Besides that, we can observe that the leader can not take more than slightly above the 50% of market share, regardless of the Θ value. It seems that in this Stackelberg competition, an increased price sensitivity implies a shift towards the non–conservative behavior of customers. This observation is interesting, as it is basically a mathematical/ computational implication, not sociological, not economical, nor it is psychological.

On the other hand, by visual inspection of Fig. 4.7 and Fig. 4.8, it appears that for higher Θ we can expect larger ratios of leader's and follower's profits and increased variability, i.e., increased sensitivity to price differences is affecting more "the maneuvering space" left to the follower, compared with the leader. Interestingly, on some occasions,

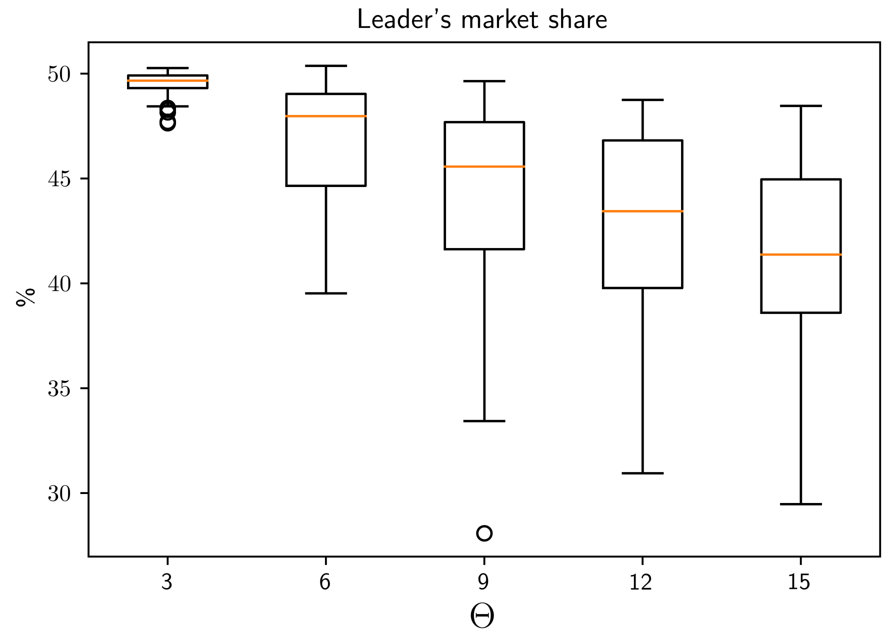


Figure 4.6: The effect of Θ on the leader's market share.

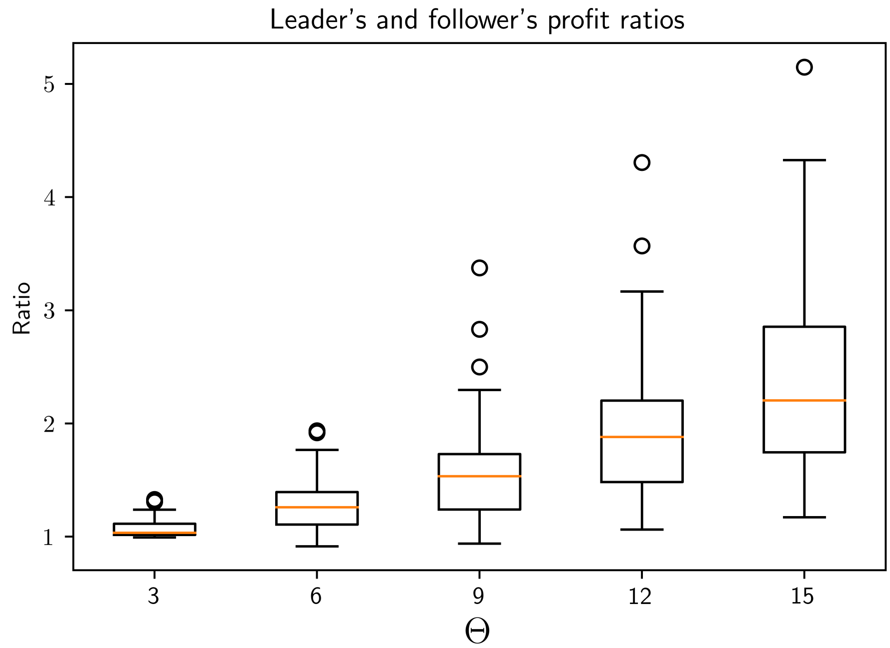


Figure 4.7: The effect of Θ on the ratio of leader's and follower's profits.

the leader can profit even five times more than the follower. We could say that if customers are highly sensitive to the price differences, then this situation is worse to the follower.

On Fig. 4.9 the observed pattern is not so solid, but it seems that the smaller NPM and increased variability correspond to a higher Θ.

Fig. 4.10–4.13 depict the effect of α to the aforementioned attributes. In many cases, increase in α has led to a smaller profit ratio. Fig. 4.10 presents three point plots corresponding to our instance batches. It indicates that a larger α corresponds to a lower ex-

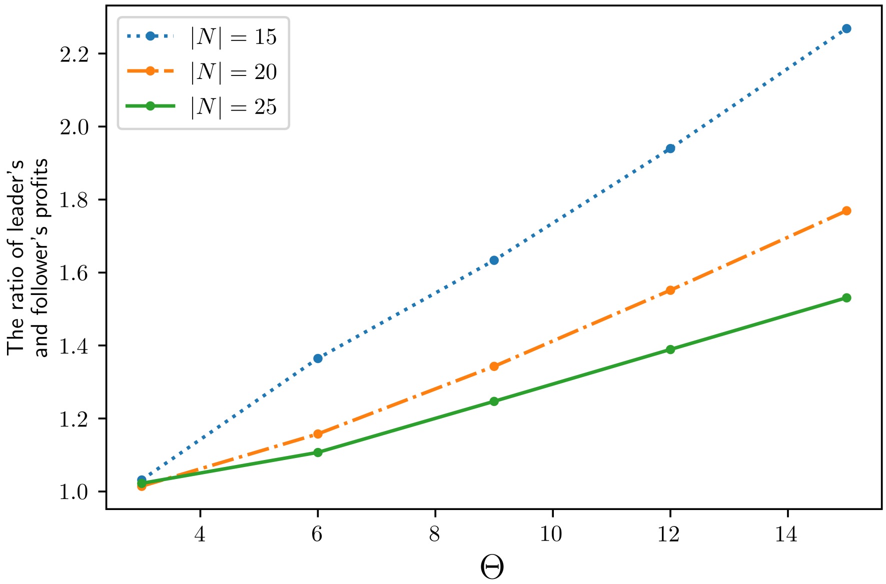


Figure 4.8: The effect of Θ on the ratio of leader's and follower's profits when $\alpha = 0.6$ and $r = p = 4$. The blue (dotted), orange (dash–dotted), and green (ordinary) lines correspond to instances of 15, 20, and 25 nodes, respectively.

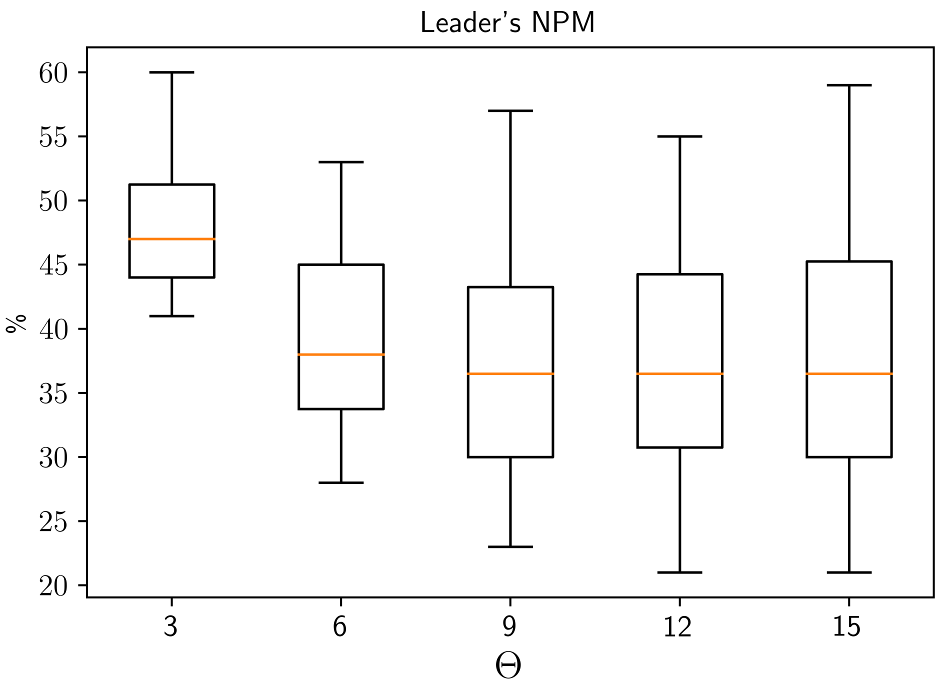


Figure 4.9: The effect of Θ on the leader's NPM.

pected profit, in general. Also, as instances get larger, the "dispersion" of values around the mean estimate increases for a given set of profits.

When it comes to the effect of α to the leader's market share, we can see that violin plot on Fig. 4.11 does not suggest any solid relationship. However, here, too, we can see that the leader usually takes less than 50% of market share.

In a way, these observations about the market share are in contrast to the naïve in-

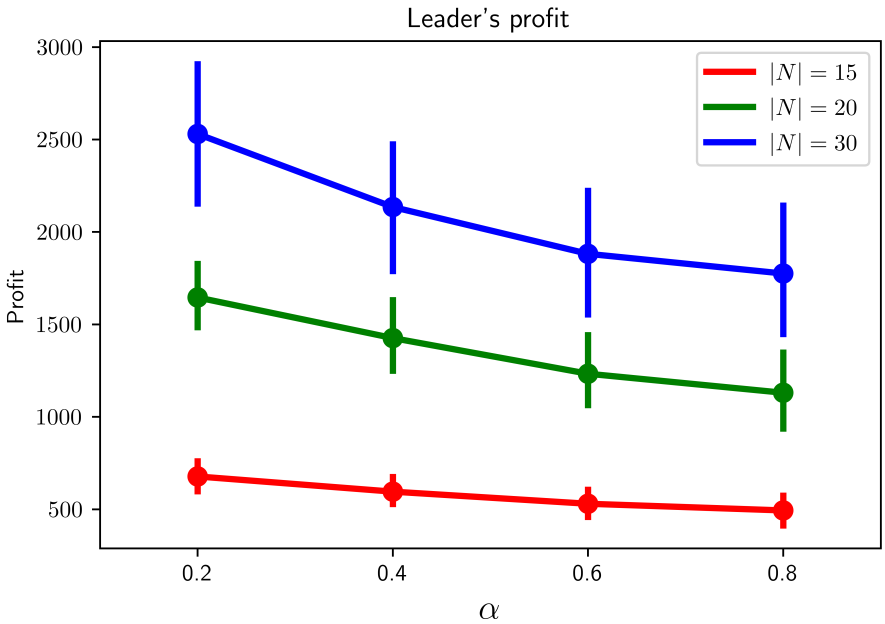


Figure 4.10: The effect of α on the leader's profit.

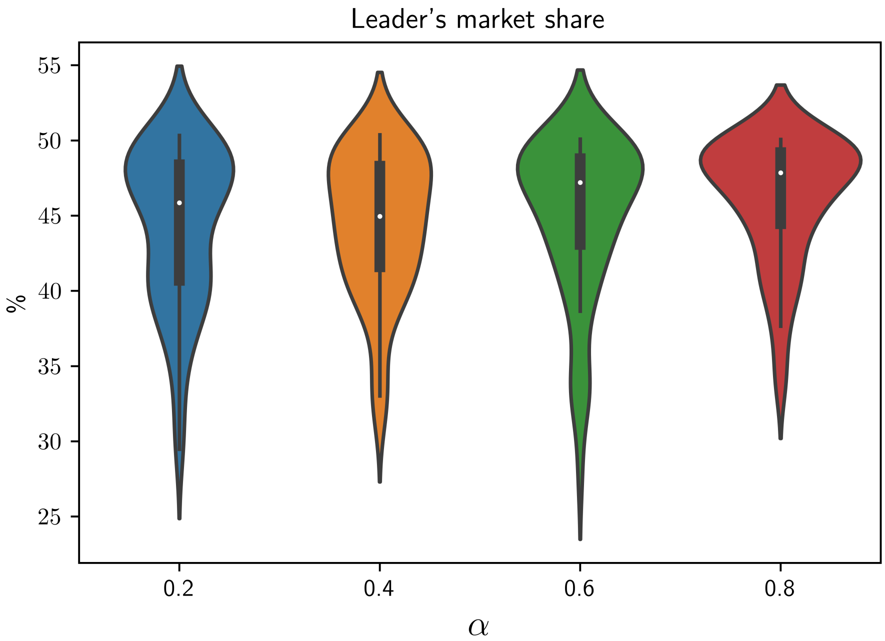


Figure 4.11: The effect of α on the leader's market share.

terpretation of Theorem 2.34 and the other observations in the literature concerning the competitive hub location (in which pricing is *not* taken into account). A possible explanation for this could be that the leader focuses on gaining profit on high demand O–D pairs. We expect that it is not always optimal for the follower to copycat the leader's network. Therefore, it could be that the leader is usually getting the larger cut on more important O–D pairs, while the follower's market share can, in general, be often larger than the leader's one as some kind of compensation.

By visual inspection of Fig. 4.12, it appears that the variability of profit ratios is smaller

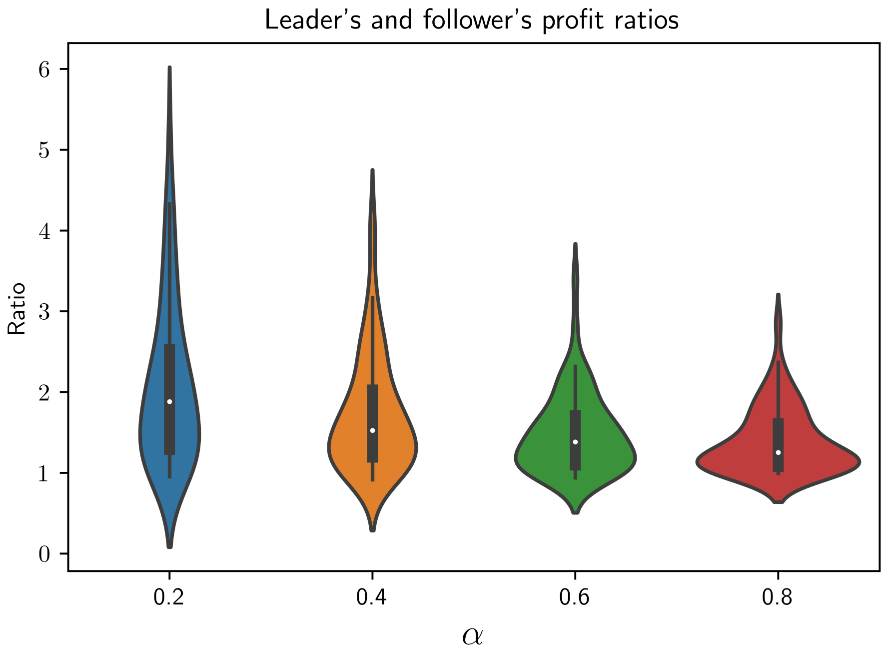


Figure 4.12: The violin plot depicting the effect of α on the ratio of leader's and follower's profits.

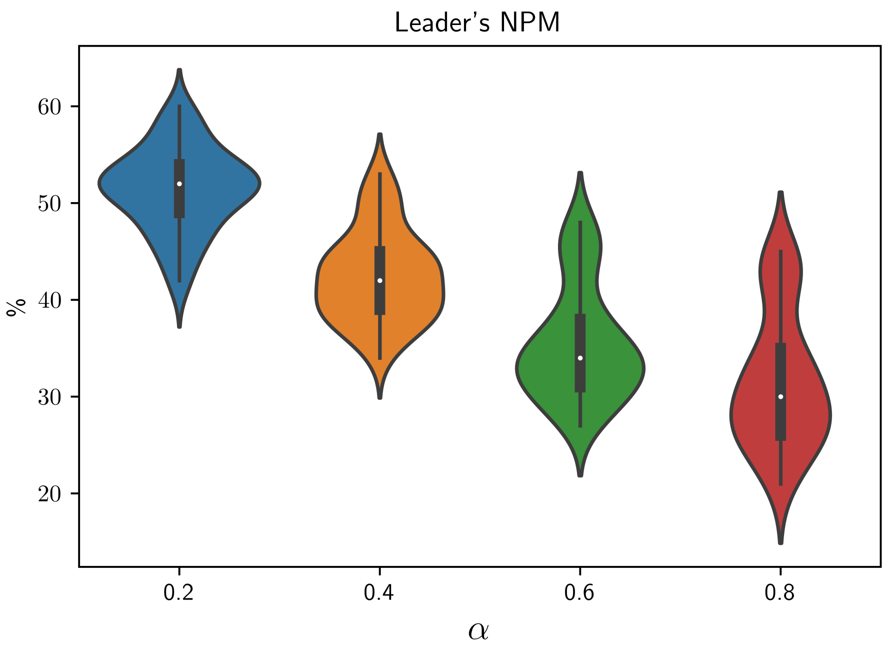


Figure 4.13: The effect of α on the leader's NPM.

for larger α. Recall that when it comes to the leader's profit the effects of α and Θ are similar. Thus, we may assume that the parameters α and Θ have the opposite effects (up to some degree) when it comes to the follower's profit.

On Fig. 4.13 we can see quite a solid pattern. Interestingly, it seems that the smaller α corresponds to the larger NPM. This observation is similar to what we have seen in the case of Θ.

4.4 Similarities Between Hub Backbones of Competitors

A similarity between leader's and follower's hub backbones can be measured with Jaccard similarity index $J : \mathcal{P}(N) \times \mathcal{P}(N) \to [0, 1]$, computed as

$$J(H_p, H_r) = \frac{|H_p \cap H_r|}{|H_p \cup H_r|}, \tag{4.3}$$

where 0 corresponds to no similarity at all ($H_p \cap H_r = \emptyset$), and 1 represents the perfect match ($H_p = H_r$).

The similarity results are presented in Table 4.3. In the first column, the instance batches are given ($|N|$). The second column presents the MAD of Jaccard similarity index values for a given instance batch (MAD of Jaccard). Minimal and maximal values of the Jaccard similarity index are given in the columns 'min Jaccard' and 'max Jaccard', respectively. In the last column, 'avg. Jaccard', Table 4.3 provides mean values of Jaccard similarity indexes, per given instance batch. Although there are cases when the leader and follower completely share the hub backbone, in general, that is not the case, as the average Jaccard similarity index and corresponding MADs are small.

Table 4.3: The Jaccard similarity index values.

|N|	MAD of Jaccard	min Jaccard	max Jaccard	avg. Jaccard
15	0.13	0.0	0.67	0.22
20	0.18	0.0	1.00	0.29
25	0.18	0.0	1.00	0.22

Another similarity measure which we will consider, refers to the well known Szymkiewicz–Simpson overlapping coefficient $SS : \mathcal{P}(N) \times \mathcal{P}(N) \to [0, 1]$. It is related to the Jaccard similarity index and it can be computed as

$$SS(H_r, H_p) = \frac{|H_r \cap H_p|}{\min(|H_r|, |H_p|)} \tag{4.4}$$

where 0 corresponds to no overlapping at all ($H_r \cap H_p = \emptyset$), and 1 represents a situation in which $H_r \subseteq H_p$ or $H_p \subseteq H_r$.

Table 4.4: The Szymkiewicz–Simpson overlapping coefficient values.

|N|	MAD of SS	min SS	max SS	avg. SS
15	0.19	0.0	0.80	0.36
20	0.23	0.0	1.00	0.40
25	0.25	0.0	1.00	0.34

The results of computing the Szymkiewicz–Simpson overlapping coefficient are presented in Table 4.4. This table is organized similarly as Table 4.3, except that instead of Jaccard, we consider the value of SS function. Although there are cases when the leader's and follower's hub backbone are overlapping, in general, that is not the case, as the average Szymkiewicz–Simpson overlapping coefficient and corresponding MAD are small.

4.5 Effect of Solving Auxiliary Model

To address the effect of solving AM, each time the AM was solved, we have compared its solution with $(r|H_p)$HMPuPW one. Table 4.5 presents, in percentages, how many times the hub backbone was different (column 'Different hub backbones'), and how many times the new solution affected the leader's profit (column 'Different leader's profits'). $|N|$ represents the instance size, i.e., a batch.

Table 4.5: The effect of AM

$\|N\|$	**Different hub backbones (%)**	**Different leader's profits (%)**
15	0.25	0
20	0.25	0
25	2.75	0

It never happened that the leader's profit was affected by the AM solution. A reason for this could lie in the fact that the CAB data set does not have enough symmetries for such a phenomenon to occur.

4.6 When Leader Ignores Follower

We can compare the results of our computational experiments, concerning $(r|p)$HCPuPW, with the situations when the leader ignores the follower. For this comparison, naturally, we take the corresponding solutions of two classical HLPs: p–HMLP and p–HCLP. Particularly, we are interested in finding the corresponding relative profit deviations

$$\frac{\mathtt{profit}_{(r|p)HCPuPW} - \mathtt{profit}_{ignore}}{\mathtt{profit}_{(r|p)HCPuPW}} \tag{4.5}$$

where $\mathtt{profit}_{(r|p)HCPuPW}$ represents the leader's profit in $(r|p)$HCPuPW (obtained by VNS algorithm) and $\mathtt{profit}_{ignore}$ represents the leader's profit in situation in which she ignores the competition. This formula assumes that the profit achieved by VNS approach is always greater than the one when the leader ignores the follower's best response.

Example 4.1. Consider an instance $(|N|, \alpha, \Theta, r, p) = (25, 0.2, 15, 3, 3)$. If the leader sets her hubs taking into account the follower's best response, then her profit will be 3315.45. However, if the leader sets her hubs according to the optimal solution of p–HMLP, then her profit will be 3192.86 and she will lose 122.63 (around 3.7%). The optimal solution of p–HCLP is even worse alternative. The leader's profit in this situation is 568.67, i.e.,

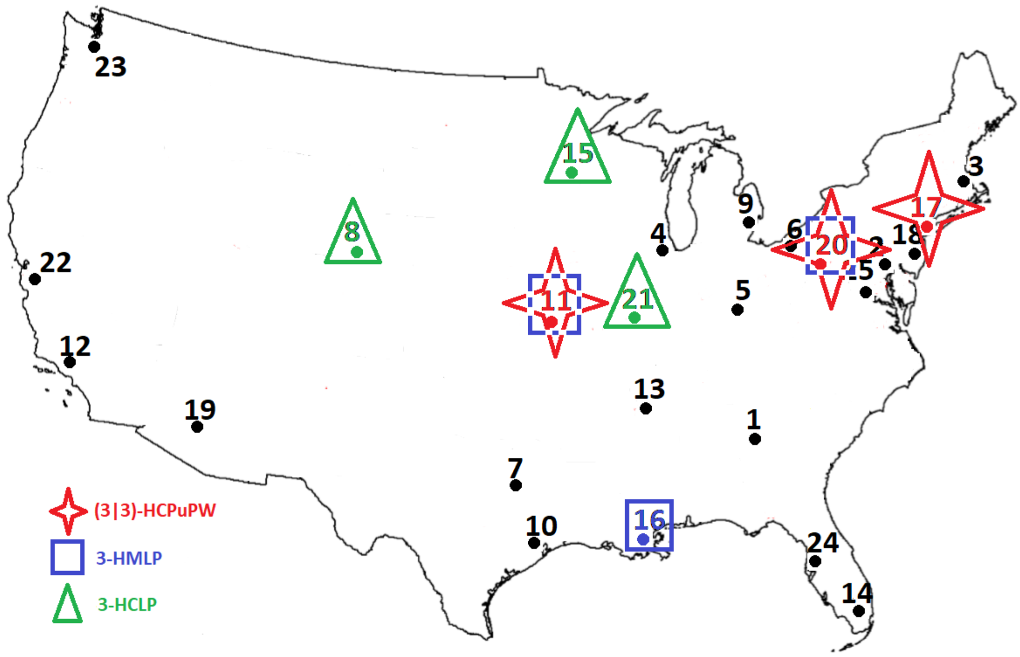


Figure 4.14: The leader's hub location in case of (3|3)HCPuPW (red), 3–HMLP (blue), and 3–HCLP (green), for CAB instance $(|N|, \alpha, \Theta, r, p) = (25, 0.2, 15, 3, 3)$.

she will lose 2746.82 (around 82.8%). Fig. 4.14 presents the leader's hub location in case of these three solutions. ▲

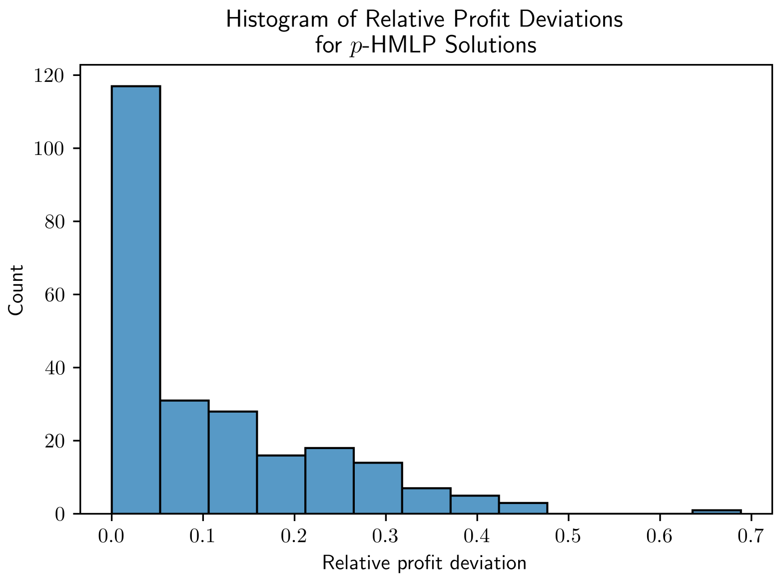


Figure 4.15: The distribution of relative profit deviations when the leader chooses p–HMLP as her strategy.

On Fig. 4.15 and Fig. 4.16 we can see the histogram and estimated ECDF corresponding to the relative profit deviations, when the leader chooses p–HMLP as her strategy.

Fig. 4.17 and Fig. 4.18 provide a comparison between p–HMLP and p–HCLP strategies. As we can see, p–HMLP is much better option than p–HCLP, concerning profit deviations. It could be that this difference is a consequence of objectives. The p–HMLP

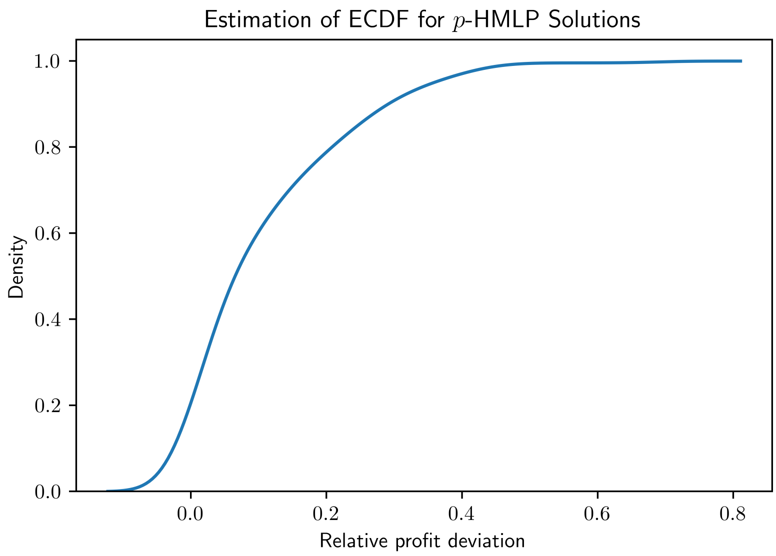


Figure 4.16: The estimation of ECDF for relative profit deviations when the leader chooses p–HMLP as her strategy.

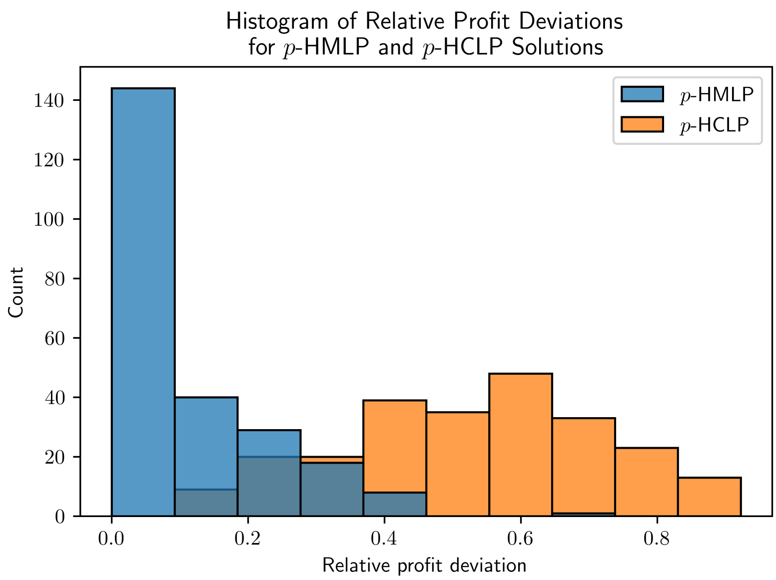


Figure 4.17: Two histograms of relative profit deviations when the leader chooses p–HMLP (orange) and p–HCLP (blue) as her strategies.

minimizes the weighted sum of variable costs for all O–D pairs, while p–HCLP minimizes only maximal corresponding weighted sum. The “narrow focus” of p–HCLP is not well suited in competitive environment, as follower has a lot of maneuvering space for its best response. Taking into account all O–D pairs puts the leader in better position.

From these plots we can realize that p–HMLP solution is a good choice to start with, either in AH or VNS, in order to address $(r|p)$HCPuPW. As we can see the relative profit deviation is reasonably small in most cases (Fig. 4.18).

The investigation of how parameters affect the relative profit deviation revealed that the effects of α and hub backbone size do not have a solid pattern. However, it seems that

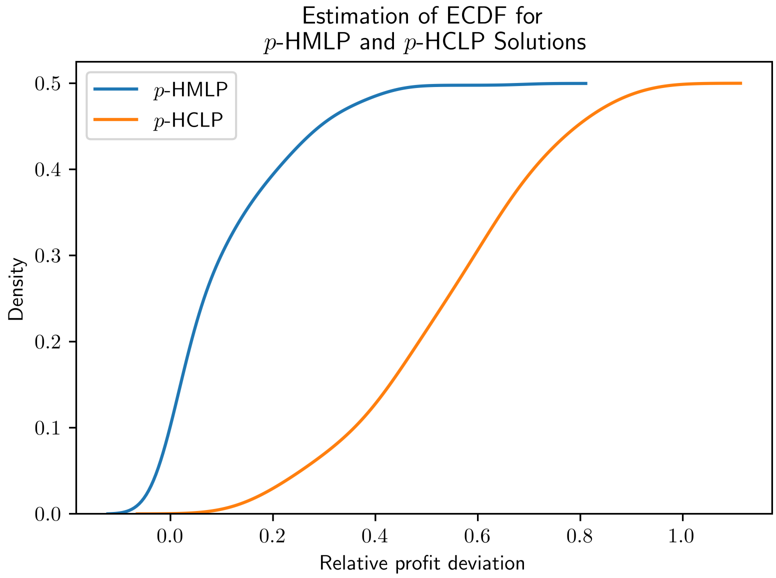


Figure 4.18: Two estimations of ECDF corresponding to the relative profit deviations when the leader chooses p–HMLP (orange) and p–HCLP (blue) as her strategy.

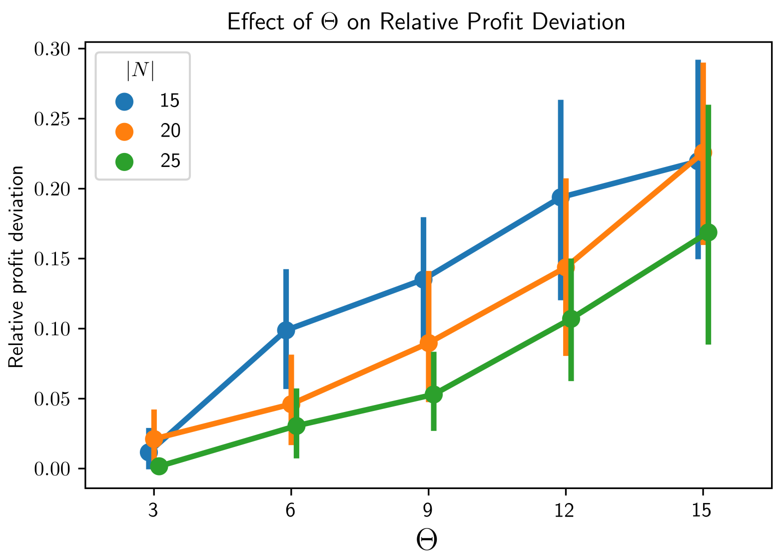


Figure 4.19: The point plots representing the effect of Θ on relative profit deviation. Different colors correspond to different instance batches.

higher Θ means also a larger relative profit deviation. The results for Θ parameter are shown as point plots for all three instance batches in Fig. 4.19.

Chapter 5

Conclusion

You don't have to be a mathematician to have a feel for numbers.

John Nash

This study introduces an intermediate variant of hub location and pricing problem in which competitors are sequentially entering the market (a leader–follower scenario), but the pricing is resolved as in the Bertrand price game. Involving pricing is a more realistic scenario than relying solely on costs and demands. The leader and the follower intend to locate p and r hubs, respectively. The setting for hub location and route opening is derived from the classic uncapacitated multiple allocation hub location problem. Multiple allocations are allowed, and there are no limits on hub capacities. Only one route can be established per O–D pair. The demand is perfectly inelastic and split between the competitors in accordance with the logit model, which is also a more realistic assumption. The objective for both companies is profit maximization, contrary to the usual viewpoint in which the company is interested in minimizing its costs. The problem is called the $(r|p)$ hub–centroid problem under the price war. Compared to some other bi–level problems in the literature, we need to define the behavior of follower accurately.

The existence of finite Bertrand–Nash price equilibrium for perfectly inelastic demand is shown, which further implied the existence of an optimal solution to the problem itself, i.e., the Stackelberg equilibrium. The new price equations are proposed for the follower, and they could be seen as a game–theoretic generalization of the expression given by Bitran and Ferrer [13]. Interestingly, the logit model and possibly different route costs yield a Bertrand–Nash price equilibrium that is not a perfect competition. Besides the pricing related statements, we have addressed the computational complexity of the leader's and follower's problems. It is shown that the follower's problem is NP–hard, but on the other hand, the derived allocation problem is, in fact, polynomially solvable. As one could assess, the leader's problem is NP–hard, too.

In this study, we showed how the corresponding follower's model could be transformed linearly. These reformulations allow the usage of commercial solvers in order to solve them. Besides that, it is shown that the optimal routes are those with the lowest costs.

As a solution approach, the AH and VNS algorithms are proposed. Different to other implementations, the initial solution is obtained by AH. Periodically, instead of evaluations of the objective functions, the estimations are used. They are build–up using the LPR solution of $(r|H_p)$HMPuPW. The $\frac{1}{e}$–law is used as a stopping rule in the local search. Computational experiments showed that our VNS algorithm is stable, and the AH is a good choice for generating the initial solution. Its cycle length is relatively small. In most cases, the difference between the best objective values returned by VNS and AH is less than a few percentages, compared to the VNS value.

The effects of price sensitivity parameters Θ and α were the most discernible. Roughly, the leader could expect to make less profit for larger Θ. A similar observation holds for the discount parameter α. On the other hand, it looks like that Θ and α have opposite effects (up to some degree) on the ratio of competitors' profits. Particularly, for lager Θ the ratio of leader's and follower's profits is also larger, while for lager α this ratio is smaller. Besides these observations, it is improbable that the leader's market share will be larger than 50%. In most cases, it will be a little bit smaller than the follower's cut. It seems that in this Stackelberg competition, we can not expect from customers to manifest a conservative behavior, in general. On the other hand, it looks like that the leader has the first–mover advantage. It is rarely optimal for the follower to copycat the leader. Finally, it is interesting that the leader's profit was never affected by the auxiliary model's solution during the computational investigation.

An exciting new research direction is investigation of relationships concerning the polynomial and approximation hierarchies, similarly as it was done in [80]. A different line of research could address a robust variant of $(r|p)$HCPuPW.

The unique finite Bertrand–Nash price equilibrium indicates that it would be reasonable to consider the cooperative price game with transferable utilities. In this setting, the leader announces the strategy profile in terms of prices and has an incentive to find the market position that would ensure the best credible threat strategy or the *status quo* point on whether the side payment is allowed or not.

A new solution approach is another possible line of future research. It is a bald move to dive into designing the exact solution method for the bi–level optimization problems with non–linear objectives and where the follower's behavior must be defined appropriately. Even constructing useful heuristic approaches for this kind of problem is a decent challenge.

5.1 Contributions

The main results obtained during the conducted research in this thesis are:

- a new competitive hub location problem is introduced in which pricing is considered;
- the bi–level non–linear mixed–integer mathematical program is formulated as a model for $(r|p)$HCPuPW;
- the existence of Bertrand–Nash price equilibrium is proved for fixed networks of competitors;
- the existence of safe Stackelberg equilibria is proved;
- the new equilibrium price equations are provided;
- the extremal behavior of optimal objective values is analyzed;
- it is proved that follower's problem is NP–hard;
- it is proved that the corresponding derived allocation problem is in class P;
- it is proved that the leader's problem is NP–hard;
- it is proved that in this setting the optimal routes for a particular O–D pair are the lowest–cost ones;
- it is shown how $(r|H_p)$HMPuPW can be formulated as a mix–integer linear program;
- it is shown how the auxiliary model can be reformulated as a mix–integer linear program;
- it is shown how to implement AH as a solution approach for the leader;
- it is shown how to design a substitute for the leader's objective function using the LPR of follower's model;
- a new coreswap neighborhood is introduced and defined;
- it is shown how to implement the VNS–based solution approach for the leader.

In total, one novel problem is introduced, for which a bi–level mathematical optimization model is formulated, 12 theorems addressing this problem are formulated and proved, accompanied with five propositions, one lemma, and nine corollaries. Besides that, two reformulations are presented, and two solution approaches are designed.

As one can see from the results of conducted computational experiments, the proposed combinatorial approach composed of theoretical and algorithmic contributions can be considered successful in addressing $(r|p)$HCPuPW. Considering all of the aforementioned, this research presents contributions to combinatorial optimization, location theory, and matheuristics. Some results and presented ideas are already published in the proceedings of international conferences and refereed international scientific journals. Other parts of this research are already in the publication process.

Bibliography

[1] Abou-Kandil, H., Bertrand, P.: Government – private sector relations as a Stackelberg game: A degenerate case. Journal of Economical Dynamics and Control 11, 513–-517 (1987). https://doi.org/10.1016/S0165-1889(87)80004-0

[2] An, Y., Zhang, Y., Zeng, B.: The reliable hub-and-spoke design problem: Models and algorithms. Transportation Research Part B: Methodological 77, 103–122 (2015). https://doi.org/10.1016/j.trb.2015.02.006.

[3] Anaconda Inc.: Anaconda Documentation, 22-07-2020. [Online]. Available: https://docs.anaconda.com/. [Accessed: 22-July-2020]

[4] Alekseeva, E., Kochetov, Yu.: Matheuristics and exact methods for the discrete $(r|p)$-centroid problem. In: El-Ghazali, T. (ed) Metaheuristics for Bi-level Optimization, pp. 189–219. Springer-Verlag, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-37838-6_7

[5] Alumur, S., Kara, B.: Network hub location problems: the state of the art. European journal of Operational Research 190(1), 1–21 (2008). https://doi.org/10.1016/j.ejor.2007.06.008

[6] Alumur, S. A., Kara, B. Y., Karasan, O. E.: The design of single allocation incomplete hub networks. Transportation Research Part B: Methodological 43, 936–951 (2009). https://doi.org/10.1016/j.trb.2009.04.004

[7] Arimoto, S.: An algorithm for computing the capacity of arbitrary discrete memoryless channels. IEEE Transactions on Information Theory, 18(1), 14–20, (1972). https://doi.org/10.1109/TIT.1972.1054753

[8] Airline Deregulation Act of 1978, Pub. L. No. 95-504, 92 Stat. 1705 (1978) (to be codified in 49 U.S.C. §§ 1301-1302, 1305-1308, 1324, 1341, 1371-1373, 1374, 1376, 1378-1382, 1386, 1389, 1461, 1471, 1473, 1482, 1490-1504)

[9] Aziz, N., Chauhan, S., Vidyarthi, N.: The impact of hub failure in hub-and-spoke networks: mathematical formulations and solution techniques. Comput Oper Res 65, 174–188 (2016). https://doi.org/10.1016/j.cor.2014.05.012

[10] Baran, P.: On distributed communications. RAND Series Reports (1964) https://doi.org/10.7249/RM3420

[11] Bhadury, J., Eiselt, H.A., Jaramillo, J.H.: An alternating heuristic for medianoid and centroid problems in the plane. Comput. Oper. Res. 30, 553–565 (2003). https://doi.org/10.1016/S0305-0548(02)00024-2

[12] Biegler, L.T., Grossmann, I.E., Westerberg, A.W.: Systematic methods of chemical process design. Prentice-Hall, New Jersey (1997)

[13] Bitran, G., Ferrer, J.C.: On Pricing and Composition of Bundles. Production and Operations Management 16, 93–108 (2007). https://doi.org/10.1111/j.1937-5956.2007.tb00168.x

[14] Blahut, R.: Computation of channel capacity and rate-distortion functions. IEEE Transactions on Information Theory, 18(4), 460–473 (1972). https://doi.org/10.1109/TIT.1972.1054855.

[15] Bollapragada, R., Li, Y., Rao, U.S.: Budget-constrained, capacitated hub location to maximize expected demand coverage in fixed-wireless telecommunication networks. INFORMS Journal on Computing, 18(4), 422–-432 (2006). https://doi.org/10.1287/ijoc.1050.0143

[16] Brimberg, J., Mladenovic, N., Todosijevic, R., Uroševic, D.: General variable neighborhood search for the uncapacitated single allocation p-hub center problem. Optimization Letters, 11(2), 377–388 (2017). https://doi.org/10.1007/s11590-016-1004-x

[17] Brown, G., Carlyle, M., Salmerón, J., Wood K.: Defending critical infrastructure. Interfaces, 36(6), 530–544 (2006). https://doi.org/10.1287/inte.1060.0252

[18] Camargo, R.S., Miranda, G., Lokketagen, A. A new formulation and an exact approach for the many-to-many hub location-routing problem. Appl Math Model 37, 12-–13 (2013). https://doi.org/10.1016/j.apm.2013.02.035

[19] Campbell, J.F.: Integer programming formulations of discrete hub location problems. European Journal of Operational Research, 72(2), 387–405 (1994). https://doi.org/10.1016/0377-2217(94)90318-2.

[20] Campbell, J., O'Kelly, M.: Twenty–five years of hub location research. Transportation Science 46(2), 153–169 (2012). https://doi.org/10.1287/trsc.1120.0410

[21] Chang, Y.J., Sahinidis, N.V.: Optimization of metabolic pathways under stability considerations. Comput Chem Eng 29, 467–479 (2005). https://doi.org/10.1016/j.compchemeng.2004.08.013

[22] Contreras, I.: Hub Location Problems. In: Laporte, G., Nickel, S., Saldanha da Gama, F. (eds) Location Science. pp. 311–344. Springer, Cham (2015)

[23] Contreras, I., Cordeau, J. F., & Laporte, G.: Stochastic uncapacitated hub location. European Journal of Operational Research, 212, 518–528 (2011). https://doi.org/10.1016/j.ejor.2011.02.018

[24] Contreras, I., Cordeau, J. F., Laporte, G.: Benders decomposition for large scale uncapacitated hub location. Operations Research, 59(6), 1477–1490 (2011). https://doi.org/10.1287/opre.1110.0965

[25] Contreras, I., Cordeau, J.F., Laporte, G.: Exact solution of large-scale hub location problems with multiple capacity levels. Transp Sci 46, 439–459 (2012). https://doi.org/10.1287/trsc.1110.0398

[26] Contreras, I., Tanash, M., Vidyarthi, N.: Exact and heuristic approaches for the cycle hub location problem. Annals of Operations Research, 258(2), 655–677 (2017). https://doi.org/10.1007/s10479-015-2091-2

[27] Čvokić, D.D., Kochetov, Y.A., Plyasunov, A.V.: A leader-follower hub location and pricing problem under fixed markups. In: Y.A. Kochetov et al. (eds.) DOOR 2016, LNCS 9869, pp. 350–363, Springer, Cham (2016). https://doi.org/10.1007/978-3-319-44914-2_2

[28] Čvokić, D.D., Kochetov, Y.A., Plyasunov, A.V., Savić, A.: The competitive hub location under the price war. In: Khachay M., Kochetov Y., Pardalos P. (eds) Mathematical Optimization Theory and Operations Research. MOTOR 2019. Lecture Notes in Computer Science, vol 11548, pp. 133–146. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22629-9_10

[29] Čvokić, D.D., Stanimirovic, Z.: A single allocation hub location and pricing problem. Comp. Appl. Math. 39, 40 (2020). https://doi.org/10.1007/s40314-019-1025-z

[30] Čvokić, D.D.: A leader–follower single allocation hub location problem under fixed markups. Filomat. 34(8), 2463–2484 (2020). doi: https://doi.org/10.2298/FIL2008463C

[31] Čvokić, D.: Results of computational experiments: solving the (r|p)HCPuPW using a BVNS-based solution approach. Harvard Dataverse, V1, UNF:6:ijSp3ImnrCPWdOHo9VPeAA== [fileUNF] (2020). https://doi.org/10.7910/DVN/UQHFXW

[32] Davydov, I.A., Kochetov, Yu.A., Mladenovic, N., Urosevic, D.: Fast metaheuristics for the discrete $(r|p)$-centroid problem. Automation and Remote Control 75(4), 677–687 (2014). https://doi.org/10.1134/S0005117914040080

[33] de Araújo, A.C.A, Roboredo, M.C., Pessoa, A.A., Pereira, V.: Exact methods for the discrete multiple allocation (r|p) hub-centroid problem, Computers & Operations Research, 116, 104870 (2020) https://doi.org/10.1016/j.cor.2019.104870.

[34] de Camargo, R. S., Miranda, G., Jr., Luna, H. P. L.: Benders decomposition for the uncapacitated multiple allocation hub location problem. Computers and Operations Research, 35(4), 1047–1064 (2008). https://doi.org/10.1016/j.cor.2006.07.002

[35] de Camargo, R. S., Miranda, G.: Single allocation hub location problem under congestion: Network owner and user perspectives. Expert Systems with Applications, 39(3), 3385–3391 (2012). https://doi.org/10.1016/j.eswa.2011.09.026

[36] de Fermat P.: Methodus ad disquirendam maximam et minimam (1636). "Method for the study of maxima and minima" Translated from the Oeuvres de Fermat by Jason Ross, available at science.larouchepac.com/fermat/fermat-maxmin.pdf.

[37] Delta Airlines: Delta First in the Air. [Online]. Available: https://www.deltamuseum.org/exhibits/delta-history/first-in-the-air [Accessed: 22-July-2020]

[38] Dempe, S.: Foundations of Bilevel Programming. Kluwer Academic, Dordrecht, Netherlands (2002). https://doi.org/10.1007/b101970

[39] Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. With discussion. J. Roy. Statist. Soc. Ser. B 39(1), 1–38 (1977). https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

[40] Domencich, T.A., McFadden, D.: Urban Travel Demand – A Behavioral Analysis. American Elsevier Publishing Company, Inc, New York, USA (1975)

[41] Du, DZ., Pardalos, P.M., Wu, W.: History of Optimization. In: Floudas C., Pardalos P. (eds) Encyclopedia of Optimization. Springer, Boston, MA (2008). https://doi.org/10.1007/978-0-387-74759-0_268

[42] d'Aspremont, J., Gabszewicz, J.J., Thisse, J.F.: On Hotelling's stability in competition, Econometrica 49, 1145–1150 (1979). https://doi.org/10.2307/1911955

[43] Ebery, J.: Solving large single allocation p-hub problems with two or three hubs. European Journal of Operational Research, 128, 447–458 (2001)

[44] Ebery, J., Krishnamoorthy, M., Ernst, A., Boland, N.: The capacitated multiple allocation hub location problem: Formulations and algorithms, European Journal of Operational Research, 120(3), 614–631, (2000). https://doi.org/10.1016/S0377-2217(98)00395-6

[45] Elhedhli, S., Wu, H.: A Lagrangean heuristic for hub-and-spoke system design with capacity selection and congestion. INFORMS J Comput 22, 282-–296 (2010). https://doi.org/10.1287/ijoc.1090.0335

[46] Ernst, A.T., Mamacher, H., Jiang, H., Krishnamoorthy, M., Woeginger, G.: Uncapacitated single and multiple allocation p-hub center problems, Computers & Operations Research, 36(7), 2230–2241 (2009). https://doi.org/10.1016/j.cor.2008.08.021.

[47] Euclid of Alexandria, Heath, T.L., Densmore, D.: Euclid's Elements: all thirteen books complete in one volume: the Thomas L. Heath translation. Green Lion Press, Santa Fe, NM, USA (2017).

[48] Farahani, R.Z., Hekmatfar, M., Arabani, A.B., Nikbakhsh, E.: Hub location problems: a review of models, classification, solution techniques, and applications. Computers & Industrial Engineering 64(4), 1096–1109 (2013). https://doi.org/10.1016/j.cie.2013.01.012

[49] Filipovic, V., Kratica, J., Tosic, D., Dugosija, D.: GA inspired heuristic for uncapacitated single allocation hub location problem. In J. Mehnen et al. (Eds.), Applications of soft computing. AISC Vol. 58, pp. 149–158, Berlin, Heidelberg: Springer-Verlag. (2009). https://doi.org/10.1007/978-3-540-89619-7_15

[50] Forgy, E.W.: Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics. 21 (3), 768–769 (1965) JSTOR 2528559

[51] Garey, M.R., Johnson, D.S.: Computers and intractability: a guide to the theory of NP-completeness. W.H. Freeman and Company, New York, USA (2009).

[52] Garrow, L.A.: Discrete choice modelling and air travel demand: theory and applications. Routledge, London, UK (2010).

[53] Gavriliouk, E.O.: Aggregation in hub location problems. Computers and Operations Research, 36, 3136–3142 (2009). https://doi.org/10.1016/j.cor.2009.01.010

[54] Goldman, A.J.: Optimal location for centers in a network. Transportation Science 3, 352–360 (1969)

[55] Gümüş Z.H., Sahin K., Ciric A.: Bilevel Programming: Applications in Engineering. In: Floudas C., Pardalos P. (eds) Encyclopedia of Optimization. Springer, Boston, MA (2008)

[56] Gurobi Optimization, L.L.C.: Gurobi Optimizer Reference Manual. [Online]. Available: https://www.gurobi.com/documentation/9.0/refman/index.html. [Accessed: 12-Jun-2020]

[57] Gurobi Optimization, LLC: About Gurobi. [Online]. Available: https://www.gurobi.com/company/about-gurobi/. [Accessed: 12-Jun-2020]

[58] Gurobi Optimization, LLC: Dr. Robert Bixby. [Online]. Available: https://www.gurobi.com/profile/robert-bixby/. [Accessed: 12-Jun-2020]

[59] Gurobi Optimization, LLC: Dr. Edward Rothberg. [Online]. Available: https://www.gurobi.com/profile/edward-rothberg/. [Accessed: 12-Jun-2020]

[60] Gurobi Optimization, LLC: Dr. Zonghao Gu. [Online]. Available: https://www.gurobi.com/profile/zonghao-gu/. [Accessed: 12-Jun-2020]

[61] Gurobi Optimization, L.L.C.: Mixed-Integer Programming (MIP) - A Primer on the Basics. [Online]. Available: https://www.gurobi.com/resource/mip-basics/. [Accessed: 12-Jun-2020]

[62] Hansen, P., Mladenović, N., Brimberg, J., Moreno Pérez, J.A.: Variable Neighborhood Search. In: Gendreau, M, Potvin, J.Y. (eds.), Handbook of Metaheuristics, International Series in Operations Research & Management Science 272, pp. 57–97. Springer, Cham (2019)

[63] Hasanzadeh, H., Bashiri, M., Amiri, A.: A new approach to optimize a hub covering location problem with a queue estimation component using genetic programming. Soft Comput 22, 949–961 (2018). https://doi.org/10.1007/s00500-016-2398-1

[64] Hekmatfar, M., Pishvaee, M.: Hub Location Problem. In: Zanjirani, R., Hekmatfar, M. (eds) Facility Location, pp. 243–270. Physica-Verlag HD, Heidelberg (2009) https://doi.org/10.1007/978-3-7908-2151-2_11

[65] Hotelling, H.: Stability in Competition, The Economic Journal 39(153), 41–57 (1929). https://doi.org/10.2307/2224214

[66] Hunter, J.D.: Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering 9(3), 90–95 (2007) https://doi.org/10.1109/MCSE.2007.55

[67] Ilić, A., Urošević, D., Brimberg, J., Mladenovic, N.: A general variable neighborhood search for solving the uncapacitated single allocation p-hub median problem. European Journal of Operational Research, 206, 289-–300 (2010). https://doi.org/10.1016/j.ejor.2010.02.022

[68] Ishfaq, R., Sox, C.R.: Design of intermodal logistics networks with hub delays. European Journal of Operational Research, 220(3), 629–641, (2012). https://doi.org/10.1016/j.ejor.2012.03.010

[69] Jankovic, O., Miskovic, S., Stanimirovic, Z., Todosijevic, R.: Novel formulations and VNS–based heuristics for single and multiple allocation p-hub maximal covering problems. Annals OR 259(1-2), 191–216 (2017). https://doi.org/10.1007/s10479-017-2508-1

[70] Kantorovich, L.V.: Mathematical Methods in the Organization and Planning of Production. Publication House, Leningrad State University: Leningrad, RU. (1939)

[71] Kalashnikov, V.V., Ríos-Mercado, R.Z.: A natural gas cash-out problem: A bilevel programming framework and a penalty function method. Optim Eng 7, 403-–420 (2006). https://doi.org/10.1007/s11081-006-0347-z

[72] Kalashnykova, N.I., Kalashnikov, V.V., Maldonado, R.C.H.: Bilevel Toll Optimization Problems: A Heuristic Algorithm Based Upon Sensitivity Analysis. In: Watada J., Watanabe T., Phillips-Wren G., Howlett R., Jain L. (eds) Intelligent Decision Technologies. Smart Innovation, Systems and Technologies, vol 15. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-29977-3_14

[73] Kara, B.Y., Taner, M.R.: Hub location problems: the location of interacting facilities. In: Eiselt, H.A., Marianov, V. (eds) Foundations of Location Analysis, 1st edn., pp. 273–288. Springer (2011) https://doi.org/10.1007/978-1-4419-7572-0_12

[74] Kara, B.Y., Tansel, B. C.: On the single–assignment p-hub center problem. European Journal of Operational Research, 125(3), 648-–655 (2000). https://doi.org/10.1016/S0377-2217(99)00274-X

[75] Kim, H., O'Kelly, M. E.: Reliable p-hub location problems in telecommunication networks. Geographical Analysis, 41(3), 283–306 (2009). https://doi.org/10.1111/j.1538-4632.2009.00755.x

[76] Kimms, A.: Economies of scale in hub and spoke network design: we have it all wrong. In: Morlock M, Schwindt C, Trautmann N, Zimmermann J (eds) Perspectives on operations research, pp. 293–317, Weisbaden, Germany (2006). https://doi.org/10.1007/978-3-8350-9064-4_17

[77] Klincewicz, J.G.: Avoiding local optima in the p-hub location problem using tabu-search and GRASP. Annals of Operational Research, 40, 283-–302 (1992). https://doi.org/10.1007/BF02060483

[78] Klincewicz, J. G.: Enumeration and search procedures for a hub location problem with economies of scale. Annals of Operations Research, 110(1–4), 107-–122 (2002). https://doi.org/10.1023/A:1020715517162

[79] Kochetov, Y., Panin, A., Plyasunov, A.: Comparison of metaheuristics for the bilevel facility location and mill pricing problem. Journal of Applied and Industrial Mathematics 9(3), 392–401 (2015). https://doi.org/10.1134/S1990478915030102

[80] Kononov, A.V., Panin, A.A. Plyasunov, A.V.: A Bilevel Competitive Location and Pricing Model with Nonuniform Split of Demand. J. Appl. Ind. Math. 13, 500–510 (2019). https://doi.org/10.1134/S1990478919030104

[81] Kratica, J., Stanimirovic, Z., Tosic, D., Filipovic, V.: Two genetic algorithms for solving the uncapacitated single allocation p-hub median problem. European Journal of Operational Research, 182, 15–28, (2007)

[82] Kratica, J., Stanimirović, Z., Filipović, V., Tošić, D.: Evolutivni pristup za rešavanje hab–lokacijskih problema. Matematički fakultet Univerziteta u Beogradu, Beograd, Srbija (2009).

[83] Kratica, J., Milanovic, M., Stanimirovic, Z., Tosic, D.: An evolutionary-based approach for solving a capacitated hub location problem. Applied Soft Computing, 11, 1858–1866 (2011). https://doi.org/10.1016/j.asoc.2010.05.035

[84] Küçükaydın, H., Aras, N., Altınel, İ.K.: A leader–follower game in competitive facility location. Computers & Operations Research 39(2), 437–448 (2012). https://doi.org/10.1016/j.cor.2011.05.007

[85] Lloyd, S.P.: Least square quantization in PCM. Bell Telephone Laboratories Paper, (1957). Published in journal much later: Lloyd, S.P. Least squares quantization in PCM. IEEE Transactions on Information Theory. 28(2), 129–137 (1982) https://doi.org/10.1109/TIT.1982.1056489

[86] Lüer-Villagra, A., Marianov, V.: A competitive hub location and pricing problem. European Journal of Operational Research 231(3), 734–744 (2013). https://doi.org/10.1016/j.ejor.2013.06.006

[87] MacQueen, J.B.: Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical

Statistics and Probability. 1. pp. 281–297, University of California Press (1967) https://projecteuclid.org/ euclid.bsmsp/1200512992

[88] Mahmutogullari, I., Kara, B.: Hub location under competition. European Journal of Operational Research 250(1), 214–225 (2016) https://doi.org/10.1016/j.ejor.2015.09.008

[89] Marianov, V., Serra, D.: Location models for airline hubs behaving as M/D/c queues. Computers and Operations Research, 30, 983–1003 (2003). https://doi.org/10.1016/S0305-0548(02)00052-7

[90] Marić, M., Stanimirović, Z., Stanojević, P.: An efficient memetic algorithm for the uncapacitated single allocation hub location problem. Soft Comput 17, 445–466 (2013). https://doi.org/10.1007/s00500-012-0919-0

[91] Marín, A.: Formulating and solving splittable capacitated multiple allocation hub location problems. Computers and Operations Research, 32, 3093–3109 (2005). https://doi.org/10.1016/j.cor.2004.04.008

[92] Mayer, G., Wagner, B.: Hub Locator: an exact solution method for the multiple allocation hub location problem, Computers & Operations Research, 29(6), 715-739 (2002) https://doi.org/10.1016/S0305-0548(01)00080-6

[93] McFadden, D.: Conditional logit analysis of qualitative choice behavior, In: P. Zarembka (eds.) Frontiers in Econometrics. pp. 105–142, New York: Academic Press (1974)

[94] Meyer, T., Ernst, A. T., Krishnamoorthy, M.: A 2-phase algorithm for solving the single allocation p-hub center problem. Computers and Operations Research, 36, 3143–3151 (2009). https://doi.org/10.1016/j.cor.2008.07.011

[95] Migdalas, A.: Bilevel programming in traffic planning: Models, methods and challenge. J Glob Optim 7, 381–405 (1995). https://doi.org/10.1007/BF01099649

[96] Mirzaghafour, F.: Modular hub location problems. Msc thesis, Concordia University, Montreal, Canada (2013)

[97] Mišković, S., Stanimirović, Z.: A hybrid metaheuristic method for the deterministic and robust uncapacitated multiple allocation p-hub centre problem. European J. of Industrial Engineering (EJIE), 11(5), 631–662 (2017). https://doi.org/10.1504/EJIE.2017.087705

[98] Mladenović, N., Hansen, P.: Variable neighborhood search. Computers & Operations Research 24, 1097–1100 (1997). https://doi.org/10.1016/S0305-0548(97)00031-2

[99] Mladenović, N., Todosijević, R., Urošević, D.: Less is more. Information Sciences 326, 160–171 (2016). https://doi.org/10.1016/j.ins.2015.07.044

[100] Gass S.I., Harris C.M.: Near-optimal solution. In: Gass S.I., Harris C.M. (eds) Encyclopedia of Operations Research and Management Science. Springer, New York, NY (2001). https://doi.org/10.1007/1-4020-0611-X_660

[101] Nunes, J.C., Boatwright, P.: Incidental Prices and Their Effect on Willingness to Pay. Journal of Marketing Research. 41(4), 457–466 (2004). doi:10.1509/jmkr.41.4.457.47014

[102] O'Kelly, M. E.: The location of interacting hub facilities. Transportation Science, 20, 92–106. (1986). https://doi.org/10.1287/trsc.20.2.92

[103] O'Kelly, M.E.: A quadratic integer program for the location of interacting hub facilities. EJOR 32, 393–404 (1987). https://doi.org/10.1016/S0377-2217(87)80007-3

[104] O'Kelly, M. E.: Rectilinear minimax hub location problems. Journal of Geographical Systems, 11, 227–241 (2009). https://doi.org/10.1007/s10109-009-0091-y

[105] O'Kelly, M.E., Luna, H.P.L., de Camargo, R.S., Miranda, G.: Hub location problems with price sensitive demands. Netw. Spat. Econ. 15(4), 917–945 (2015). https://doi.org/10.1007/s11067-014-9276-0

[106] Panin, A., Pashchenko, M., Plyasunov, A.: Bilevel competitive facility location and pricing problems. Automation and Remote Control, 75(4), 715–727 (2014) https://doi.org/10.1134/S0005117914040110

[107] Qu, B., Weng, K.: Path relinking approach for multiple allocation hub maximal covering problem. Computers and Mathematics with Applications, 57, 1890–1894 (2009). https://doi.org/10.1016/j.camwa.2008.10.004

[108] Redondo, J.L., Fernández, J., García, I., Ortigosa P.M.: Heuristics for the facility location and design $(1|1)$-centroid problem on the plane. Comput Optim Appl 45(1), 111–141 (2010) https://doi.org/10.1007/s10589-008-9170-0

[109] Rieck, J., Ehrenberg, C., Zimmermann, J.: Many-to-many location-routing with inter-hub transport and multi-commodity pickup-and-delivery. Eur J Oper Res 236, 863–878 (2009). https://doi.org/10.1016/j.ejor.2013.12.021

[110] Rodrigue, J.: The Geography of Transport Systems 5th Edition Routledge (2020) ISBN 9780367364632

[111] Rodríguez-Martín, I., Salazar-González, J., Yaman, H.: A branch-and-cut algorithm for the hub location and routing problem. Comput Oper Res 50, 161–174 (2014). https://doi.org/10.1016/j.cor.2014.04.014

[112] Sasaki, M.: Hub network design model in a competitive environment with flow threshold. Journal of the Operations Research Society of Japan 48, 158-–71 (2005)

[113] Sasaki, M., Campbell. J.F., Ernst, A.T., Krishnamoorthy, M.: Hub Arc Location with Competition. Technical Report of the Nanzan Academic Society Information Sciences and Engineering. (2009)

[114] Sasaki, M., Fukushima, M.: Stackelberg hub location problem. Journal of the Operations Research Society of Japan 44(4), 390–405 (2001) doi: 10.1016/S0453-4514(01)80019-3. https://doi.org/10.15807/jorsj.48.158

[115] Sasaki, M., Suzuki, A., Drezner, Z.: On the selection of hub airports for an airline hub-and-spoke system. Computers and Operations Research, 26, 1411–1422 (1999). https://doi.org/10.1016/S0305-0548(99)00043-X

[116] Scaparra, M.P., Church, R. L.: A bilevel mixed-integer program for critical infrastructure protection planning. Computers & Operations Research, 35(6, 1905–1923 (2008). https://doi.org/10.1016/j.cor.2006.09.019

[117] Shahabi, M., Unnikrishnan, A.: Robust hub network design problem. Transport. Res. E-Log. 70, 356-373 (2014). https://doi.org/10.1016/j.tre.2014.08.003

[118] Skorin-Kapov, D., Skorin-Kapov, J., O'Kelly, M.: Tight linear programming relaxations of uncapacitated p-hub median problems. European Journal of Operational Research, 94(3), 582–593 (1996) https://doi.org/10.1016/0377-2217(95)00100-X

[119] Smith, R., Krishnamoorthy, M., Palaniswami, M.: Neural versus traditional approaches to the location of interacting hub facilities. Location Science, 4(3), 155–171 (1996). https://doi.org/10.1016/S0966-8349(96)00017-4

[120] Stackelberg, H.von: Marktform und Gleichgewicht. Verl. Wirtschaft und Finanzen, Düsseldorf, Germany (1934).

[121] Tanash, M., Contreras, I., Vidyarthi N.: An exact algorithm for the modular hub location problem with single assignments. Computers & Operations Research 85, 32–44 (2017). https://doi.org/10.1016/j.cor.2017.03.006

[122] Todosijevic, R., Urosevic, D., Mladenovic, N., Hanafi, S.: A general variable neighborhood search for solving the uncapacitated r-allocation p-hub median problem. Optim Lett 11, 1109-–1121 (2017). https://doi.org/10.1007/s11590-015-0867-6

[123] Topcuoglu, H., Corut, F., Ermis, M., Yilmaz, G.: Solving the uncapacitated hub location using genetic algorithms. Computers and Operational Research, 32, 467–984 (2005). https://doi.org/10.1016/j.cor.2003.09.008

[124] Train, K.: Discrete Choice Methods with Simulation. Cambridge University Press (2009)

[125] Ungureanu, V.: Solution Principles for Mixtures of Simultaneous and Sequential Games. In: Pareto-Nash-Stackelberg Game and Control Theory. Smart Innovation, Systems and Technologies, vol 89. Springer, Cham (2018) http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-319-75151-1_9

[126] Yaman, H.: Star p-hub median problem with modular arc capacities. Computers and Operations Research, 35(9), 3009–3019 (2008). https://doi.org/10.1016/j.cor.2007.01.014

[127] Yuille, A.L., Rangarajan, A.: The Concave–Convex Procedure. Neural Computation 15:4, 915–936 (2003). https://doi.org/10.1162/08997660360581958

Appendices

Appendix A

CAB Dataset

This data file has been contributed by M.E. O'Kelly

This is the 25 node data set that has been
used extensively with hub and spoke location models. It is
sometimes referred to as the CAB data set. Full references
to the source and prior results for these data can be found
in several previous paper, including

O'Kelly, M.E. 1987, "A quadratic integer program for the
location of interacting hub facilities", EJOR, 32, 393-404.

The list of cities in the 25 node data set is given below.

```
# 25 node hub data with unrounded costs #
## These new costs correspond to
O'Kelly, Bryan, Skorin-Kapov, Skorin-Kapov, O'Kelly paper
in Location Science volume 4(3), October 1996
# multiple and single assignment hub model #
# for use with mult.mod, single.mod or comb.mod #
# last revised 7 September 1994 #
```

The data show i,j, Wij and Cij for a 25x25 system. The
10, 15 and 20 node subsets are found by taking the top 10x10 15x15
and 20x20 submatrices

```
param:       W          C    :=
1  1        0.         0
1  2     6469.   576.9631
1  3     7629.   946.4954
1  4    20036.   597.5972
1  5     4690.   373.8127
1  6     6194.   559.7673
1  7    11688.   709.0215
1  8     2243.   1208.328
1  9     8857.   603.6477
1 10     7248.    695.208
1 11     3559.    680.709
1 12     9221.   1936.572
1 13    10099.   332.4644
1 14    22866.   592.5679
1 15     3388.   908.7715
1 16     9986.   426.1877
1 17    46618.   756.1987
1 18    11639.   672.5906
1 19     1380.   1590.224
1 20     5261.   527.3008
1 21     5985.   483.4673
```

1	22	6731.	2140.978
1	23	2704.	2184.402
1	24	12250.	408.1648
1	25	16132.	540.7388
2	1	6469.	576.9631
2	2	0.	0
2	3	12999.	369.5327
2	4	13692.	613.0386
2	5	3322.	429.1079
2	6	5576.	312.8831
2	7	3878.	1196.489
2	8	3202.	1502.14
2	9	6699.	405.8975
2	10	4198.	1241.961
2	11	2454.	960.3459
2	12	7975.	2318.076
2	13	1186.	786.5959
2	14	7443.	949.5669
2	15	1162.	938.7461
2	16	5105.	999.5005
2	17	24817.	179.2426
2	18	6532.	96.2744
2	19	806.	1999.584
2	20	8184.	210.7656
2	21	3896.	736.3755
2	22	7333.	2456.263
2	23	3719.	2339.509
2	24	2015.	844.1663
2	25	565.	36.4947
3	1	7629.	946.4954
3	2	12999.	369.5327
3	3	0.	0
3	4	35135.	858.3308
3	5	5956.	749.6018
3	6	14121.	556.0706
3	7	5951.	1541.273
3	8	5768.	1764.791
3	9	16578.	621.3306
3	10	4242.	1603.165
3	11	3365.	1250.962
3	12	22254.	2600.078
3	13	1841.	1137.335
3	14	23665.	1266.851
3	15	6517.	1124.778
3	16	3541.	1368.267
3	17	205088.	190.3157
3	18	37669.	274.3105
3	19	2885.	2299.429
3	20	13200.	494.2224
3	21	7116.	1043.484
3	22	17165.	2703.402
3	23	4284.	2503.828
3	24	8085.	1188.549
3	25	51895.	405.7886
4	1	20036.	597.5972
4	2	13692.	613.0386
4	3	35135.	858.3308
4	4	0.	0
4	5	19094.	255.0303
4	6	35119.	311.3071
4	7	21423.	790.1213
4	8	27342.	907.4331
4	9	51341.	237.0703
4	10	15826.	932.2173
4	11	28537.	406.3386

4	12	65387.	1741.873
4	13	12980.	485.5564
4	14	44097.	1186.858
4	15	51525.	345.8738
4	16	14354.	830.3635
4	17	172895.	720.4687
4	18	37305.	675.3437
4	19	15418.	1447.104
4	20	26221.	403.8657
4	21	42303.	255.8823
4	22	35303.	1853.617
4	23	13618.	1733.132
4	24	17580.	1005.761
4	25	40708.	592.0278
5	1	4690.	373.8127
5	2	3322.	429.1079
5	3	5956.	749.6018
5	4	19094.	255.0303
5	5	0.	0
5	6	7284.	225.8954
5	7	3102.	794.1726
5	8	1562.	1080.374
5	9	7180.	238.944
5	10	1917.	879.5647
5	11	2253.	533.156
5	12	5951.	1889.528
5	13	1890.	402.3291
5	14	7097.	947.3188
5	15	2009.	598.541
5	16	1340.	700.4368
5	17	25303.	578.3286
5	18	6031.	512.3965
5	19	1041.	1570.725
5	20	4128.	255.6551
5	21	5452.	307.3289
5	22	3344.	2036.128
5	23	1067.	1967.256
5	24	4608.	775.239
5	25	7050.	399.2253
6	1	6194.	559.7673
6	2	5576.	312.8831
6	3	14121.	556.0706
6	4	35119.	311.3071
6	5	7284.	225.8954
6	6	0.	0
6	7	5023.	1009.689
6	8	3512.	1216.868
6	9	10419.	94.2588
6	10	3543.	1104.574
6	11	2752.	694.9153
6	12	14412.	2047.122
6	13	2043.	627.115
6	14	15642.	1084.5
6	15	5014.	626.1548
6	16	2016.	922.3181
6	17	62034.	409.3542
6	18	15385.	365.6853
6	19	2957.	1743.432
6	20	5035.	104.6478
6	21	7482.	491.1125
6	22	6758.	2164.855
6	23	2191.	2027.319
6	24	6599.	933.196
6	25	14181.	298.8486
7	1	11688.	709.0215

7	2	3878.	1196.489
7	3	5951.	1541.273
7	4	21423.	790.1213
7	5	3102.	794.1726
7	6	5023.	1009.689
7	7	0.	0
7	8	11557.	663.8762
7	9	6479.	982.7378
7	10	34261.	221.422
7	11	10134.	447.8044
7	12	27350.	1249.763
7	13	6929.	411.1133
7	14	7961.	1097.608
7	15	4678.	851.8228
7	16	13511.	423.7053
7	17	29801.	1362.874
7	18	7549.	1288.966
7	19	5550.	895.0908
7	20	3089.	1049.266
7	21	9958.	537.6206
7	22	14110.	1493.843
7	23	4911.	1686.675
7	24	2722.	912.2104
7	25	10802.	1161.676
8	1	2243.	1208.328
8	2	3202.	1502.14
8	3	5768.	1764.791
8	4	27342.	907.4331
8	5	1562.	1080.374
8	6	3512.	1216.868
8	7	11557.	663.8762
8	8	0.	0
8	9	5615.	1143.791
8	10	7095.	874.5181
8	11	10753.	551.6299
8	12	30362.	841.624
8	13	1783.	880.0728
8	14	3437.	1714.651
8	15	8897.	694.0088
8	16	2509.	1066.563
8	17	23273.	1625.87
8	18	5160.	1574.822
8	19	8750.	593.4216
8	20	2583.	1301.511
8	21	7288.	780.9512
8	22	17481.	955.802
8	23	7930.	1024.566
8	24	1278.	1519.174
8	25	8447.	1475.479
9	1	8857.	603.6477
9	2	6699.	405.8975
9	3	16578.	621.3306
9	4	51341.	237.0703
9	5	7180.	238.944
9	6	10419.	94.2588
9	7	6479.	982.7378
9	8	5615.	1143.791
9	9	0.	0
9	10	4448.	1094.906
9	11	5076.	636.9045
9	12	22463.	1978.943
9	13	4783.	620.488
9	14	24609.	1151.868
9	15	9969.	535.0244
9	16	4224.	936.2502

9	17	79945.	489.5645
9	18	20001.	453.2583
9	19	4291.	1682.489
9	20	10604.	198.9058
9	21	11925.	450.2585
9	22	13091.	2086.845
9	23	4172.	1936.304
9	24	12891.	992.3379
9	25	19500.	392.9045
10	1	7248.	695.208
10	2	4198.	1241.961
10	3	4242.	1603.165
10	4	15826.	932.2173
10	5	1917.	879.5647
10	6	3543.	1104.574
10	7	34261.	221.422
10	8	7095.	874.5181
10	9	4448.	1094.906
10	10	0.	0
10	11	4370.	642.2092
10	12	17267.	1375.635
10	13	3929.	477.459
10	14	8602.	963.7202
10	15	2753.	1046.119
10	16	20013.	305.3132
10	17	28080.	1417.072
10	18	5971.	1337.648
10	19	2131.	1017.332
10	20	3579.	1125.041
10	21	6809.	677.0608
10	22	8455.	1649.619
10	23	2868.	1891.166
10	24	2336.	795.2136
10	25	5616.	1205.747
11	1	3559.	680.709
11	2	2454.	960.3459
11	3	3365.	1250.962
11	4	28537.	406.3386
11	5	2253.	533.156
11	6	2752.	694.9153
11	7	10134.	447.8044
11	8	10753.	551.6299
11	9	5076.	636.9045
11	10	4370.	642.2092
11	11	0.	0
11	12	15287.	1358.213
11	13	3083.	378.5906
11	14	4092.	1236.192
11	15	7701.	405.0906
11	16	2809.	674.479
11	17	17291.	1096.712
11	18	4462.	1038.645
11	19	3239.	1048.539
11	20	2309.	768.1641
11	21	16003.	229.4867
11	22	8381.	1506.451
11	23	3033.	1503.794
11	24	1755.	1038.624
11	25	7266.	931.7148
12	1	9221.	1936.572
12	2	7975.	2318.076
12	3	22254.	2600.078
12	4	65387.	1741.873
12	5	5951.	1889.528
12	6	14412.	2047.122

12	7	27350.	1249.763
12	8	30362.	841.624
12	9	22463.	1978.943
12	10	17267.	1375.635
12	11	15287.	1358.213
12	12	0.	0
12	13	5454.	1608.082
12	14	15011.	2335.816
12	15	17714.	1530.57
12	16	10037.	1661.778
12	17	105507.	2453.352
12	18	20040.	2396.794
12	19	31780.	358.3762
12	20	10822.	2125.512
12	21	16450.	1582.369
12	22	92083.	361.5388
12	23	32908.	986.8149
12	24	3865.	2157.517
12	25	24583.	2288.748
13	1	10099.	332.4644
13	2	1186.	786.5959
13	3	1841.	1137.335
13	4	12980.	485.5564
13	5	1890.	402.3291
13	6	2043.	627.115
13	7	6929.	411.1133
13	8	1783.	880.0728
13	9	4783.	620.488
13	10	3929.	477.459
13	11	3083.	378.5906
13	12	5454.	1608.082
13	13	0.	0
13	14	3251.	858.251
13	15	1126.	700.8213
13	16	5926.	348.2725
13	17	10653.	955.6191
13	18	3062.	879.9795
13	19	759.	1265.573
13	20	1255.	651.1179
13	21	6173.	254.9977
13	22	2974.	1808.52
13	23	1056.	1872.696
13	24	1504.	660.5173
13	25	4588.	751.4614
14	1	22866.	592.5679
14	2	7443.	949.5669
14	3	23665.	1266.851
14	4	44097.	1186.858
14	5	7097.	947.3188
14	6	15642.	1084.5
14	7	7961.	1097.608
14	8	3437.	1714.651
14	9	24609.	1151.868
14	10	8602.	963.7202
14	11	4092.	1236.192
14	12	15011.	2335.816
14	13	3251.	858.251
14	14	0.	0
14	15	5550.	1500.774
14	16	9473.	675.7505
14	17	169397.	1098.282
14	18	25073.	1021.611
14	19	1170.	1977.613
14	20	14272.	1015.165
14	21	8543.	1065.599

14	22	8064.	2591.447
14	23	1840.	2725.79
14	24	20618.	197.8015
14	25	20937.	923.2229
15	1	3388.	908.7715
15	2	1162.	938.7461
15	3	6517.	1124.778
15	4	51525.	345.8738
15	5	2009.	598.541
15	6	5014.	626.1548
15	7	4678.	851.8228
15	8	8897.	694.0088
15	9	9969.	535.0244
15	10	2753.	1046.119
15	11	7701.	405.0906
15	12	17714.	1530.57
15	13	1126.	700.8213
15	14	5550.	1500.774
15	15	0.	0
15	16	2152.	1039.77
15	17	26816.	1018.399
15	18	6931.	987.8645
15	19	4947.	1280.737
15	20	2676.	728.3743
15	21	8033.	450.3982
15	22	12692.	1589.835
15	23	6157.	1401.321
15	24	3065.	1311.21
15	25	12044.	922.3145
16	1	9986.	426.1877
16	2	5105.	999.5005
16	3	3541.	1368.267
16	4	14354.	830.3635
16	5	1340.	700.4368
16	6	2016.	922.3181
16	7	13511.	423.7053
16	8	2509.	1066.563
16	9	4224.	936.2502
16	10	20013.	305.3132
16	11	2809.	674.479
16	12	10037.	1661.778
16	13	5926.	348.2725
16	14	9473.	675.7505
16	15	2152.	1039.77
16	16	0.	0
16	17	21806.	1178.439
16	18	4519.	1095.657
16	19	886.	1304.043
16	20	1742.	918.5615
16	21	4782.	601.9917
16	22	6453.	1916.578
16	23	2022.	2090.089
16	24	3546.	496.4224
16	25	5065.	963.0435
17	1	46618.	756.1987
17	2	24817.	179.2426
17	3	205088.	190.3157
17	4	172895.	720.4687
17	5	25303.	578.3286
17	6	62034.	409.3542
17	7	29801.	1362.874
17	8	23273.	1625.87
17	9	79945.	489.5645
17	10	28080.	1417.072
17	11	17291.	1096.712

17	12	105507.	2453.352
17	13	10653.	955.6191
17	14	169397.	1098.282
17	15	26816.	1018.399
17	16	21806.	1178.439
17	17	0.	0
17	18	9040.	84.3365
17	19	11139.	2143.565
17	20	63153.	328.7515
17	21	34092.	880.5469
17	22	70935.	2574.082
17	23	14957.	2415.489
17	24	28398.	1008.2
17	25	166694.	215.561
18	1	11639.	672.5906
18	2	6532.	96.2744
18	3	37669.	274.3105
18	4	37305.	675.3437
18	5	6031.	512.3965
18	6	15385.	365.6853
18	7	7549.	1288.966
18	8	5160.	1574.822
18	9	20001.	453.2583
18	10	5971.	1337.648
18	11	4462.	1038.645
18	12	20040.	2396.794
18	13	3062.	879.9795
18	14	25073.	1021.611
18	15	6931.	987.8645
18	16	4519.	1095.657
18	17	9040.	84.3365
18	18	0.	0
18	19	2802.	2082.316
18	20	30224.	273.4106
18	21	7982.	818.1228
18	22	14964.	2526.562
18	23	4589.	2388.689
18	24	6227.	926.6267
18	25	12359.	132.7684
19	1	1380.	1590.224
19	2	806.	1999.584
19	3	2885.	2299.429
19	4	15418.	1447.104
19	5	1041.	1570.725
19	6	2957.	1743.432
19	7	5550.	895.0908
19	8	8750.	593.4216
19	9	4291.	1682.489
19	10	2131.	1017.332
19	11	3239.	1048.539
19	12	31780.	358.3762
19	13	759.	1265.573
19	14	1170.	1977.613
19	15	4947.	1280.737
19	16	886.	1304.043
19	17	11139.	2143.565
19	18	2802.	2082.316
19	19	0.	0
19	20	1869.	1814.83
19	21	3716.	1264.193
19	22	11510.	661.6543
19	23	3519.	1129.327
19	24	569.	1800.098
19	25	3520.	1968.689
20	1	5261.	527.3008

20	2	8184.	210.7656
20	3	13200.	494.2224
20	4	26221.	403.8657
20	5	4128.	255.6551
20	6	5035.	104.6478
20	7	3089.	1049.266
20	8	2583.	1301.511
20	9	10604.	198.9058
20	10	3579.	1125.041
20	11	2309.	768.1641
20	12	10822.	2125.512
20	13	1255.	651.1179
20	14	14272.	1015.165
20	15	2676.	728.3743
20	16	1742.	918.5615
20	17	63153.	328.7515
20	18	30224.	273.4106
20	19	1869.	1814.83
20	20	0.	0
20	21	5020.	552.4229
20	22	6610.	2253.211
20	23	2139.	2128.828
20	24	5431.	875.2542
20	25	13541.	194.5945
21	1	5985.	483.4673
21	2	3896.	736.3755
21	3	7116.	1043.484
21	4	42303.	255.8823
21	5	5452.	307.3289
21	6	7482.	491.1125
21	7	9958.	537.6206
21	8	7288.	780.9512
21	9	11925.	450.2585
21	10	6809.	677.0608
21	11	16003.	229.4867
21	12	16450.	1582.369
21	13	6173.	254.9977
21	14	8543.	1065.599
21	15	8033.	450.3982
21	16	4782.	601.9917
21	17	34092.	880.5469
21	18	7982.	818.1228
21	19	3716.	1264.193
21	20	5020.	552.4229
21	21	0.	0
21	22	9942.	1735.937
21	23	3276.	1712.136
21	24	3820.	871.6396
21	25	11799.	706.5024
22	1	6731.	2140.978
22	2	7333.	2456.263
22	3	17165.	2703.402
22	4	35303.	1853.617
22	5	3344.	2036.128
22	6	6758.	2164.855
22	7	14110.	1493.843
22	8	17481.	955.802
22	9	13091.	2086.845
22	10	8455.	1649.619
22	11	8381.	1506.451
22	12	92083.	361.5388
22	13	2974.	1808.52
22	14	8064.	2591.447
22	15	12692.	1589.835
22	16	6453.	1916.578

22	17	70935.	2574.082
22	18	14964.	2526.562
22	19	11510.	661.6543
22	20	6610.	2253.211
22	21	9942.	1735.937
22	22	0.	0
22	23	35285.	694.9363
22	24	2566.	2404.839
22	25	19926.	2430.269
23	1	2704.	2184.402
23	2	3719.	2339.509
23	3	4284.	2503.828
23	4	13618.	1733.132
23	5	1067.	1967.256
23	6	2191.	2027.319
23	7	4911.	1686.675
23	8	7930.	1024.566
23	9	4172.	1936.304
23	10	2868.	1891.166
23	11	3033.	1503.794
23	12	32908.	986.8149
23	13	1056.	1872.696
23	14	1840.	2725.79
23	15	6157.	1401.321
23	16	2022.	2090.089
23	17	14957.	2415.489
23	18	4589.	2388.689
23	19	3519.	1129.327
23	20	2139.	2128.828
23	21	3276.	1712.136
23	22	35285.	694.9363
23	23	0.	0
23	24	940.	2528.479
23	25	4951.	2321.873
24	1	12250.	408.1648
24	2	2015.	844.1663
24	3	8085.	1188.549
24	4	17580.	1005.761
24	5	4608.	775.239
24	6	6599.	933.196
24	7	2722.	912.2104
24	8	1278.	1519.174
24	9	12891.	992.3379
24	10	2336.	795.2136
24	11	1755.	1038.624
24	12	3865.	2157.517
24	13	1504.	660.5173
24	14	20618.	197.8015
24	15	3065.	1311.21
24	16	3546.	496.4224
24	17	28398.	1008.2
24	18	6227.	926.6267
24	19	569.	1800.098
24	20	5431.	875.2542
24	21	3820.	871.6396
24	22	2566.	2404.839
24	23	940.	2528.479
24	24	0.	0
24	25	6237.	813.5513
25	1	16132.	540.7388
25	2	565.	36.4947
25	3	51895.	405.7886
25	4	40708.	592.0278
25	5	7050.	399.2253
25	6	14181.	298.8486

25	7	10802.	1161.676
25	8	8447.	1475.479
25	9	19500.	392.9045
25	10	5616.	1205.747
25	11	7266.	931.7148
25	12	24583.	2288.748
25	13	4588.	751.4614
25	14	20937.	923.2229
25	15	12044.	922.3145
25	16	5065.	963.0435
25	17	166694.	215.561
25	18	12359.	132.7684
25	19	3520.	1968.689
25	20	13541.	194.5945
25	21	11799.	706.5024
25	22	19926.	2430.269
25	23	4951.	2321.873
25	24	6237.	813.5513
25	25	0.	0

Appendix B

USA Map for CAB Dataset

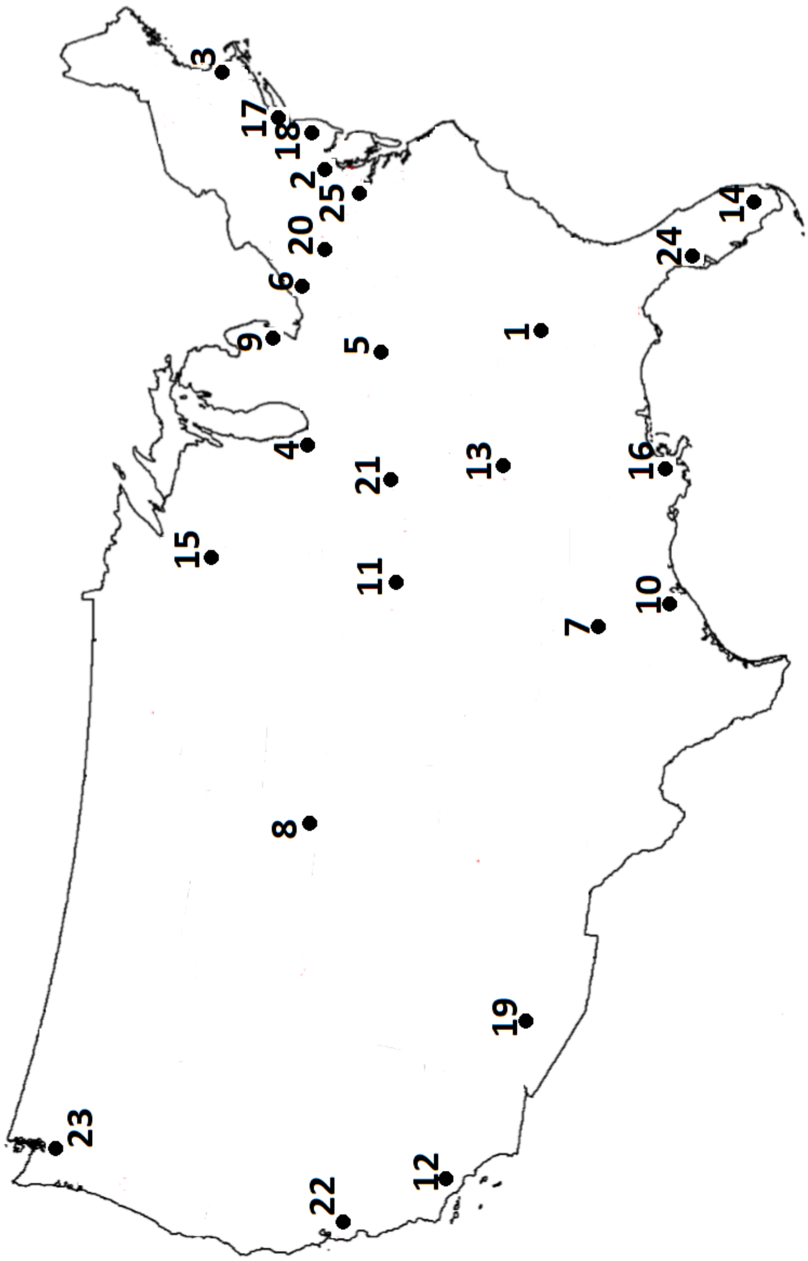


Figure 2.1: The USA map for CAB dataset. Cities are denoted with numbers, as in [103].

Appendix C

Results of Computational Experiments

Tables C.1, C.2, and C.3 present the computational results concerning the profit values, running times and number of AH iterations. In the first three columns, the instance values for the inter–hub discount factor α, sensitivity parameter Θ, and hub backbone size ('# of hubs') are given, respectively. For every instance, the VNS algorithm was executed for 10 times. The best leader's profit value found among all 10 executions is given in the column 'best profit'. The mean value of best leader's profits regarding these 10 executions is presented in the column 'avg. profit'. Similarly, the mean value of times (in seconds) for which the VNS has found best solution is given in the column 'avg. best time (s)'. The mean total time (in seconds) of VNS executions is presented in the column 'avg. run time (s)'. In the last column '# of AH iterations', the number of AH iterations is given.

Table C.1: Computational results when $|N|= 15$.

α	Θ	# of hubs	best profit	avg. profit	avg. best time (s)	avg. run time (s)	# of AH iterations
0.2	3	2	1008.68	1008.68	7.37	2701.43	3
		3	1021.98	1021.98	415.92	2725.23	3
		4	946.24	946.24	148.21	2719.89	3
		5	804.17	804.17	7.39	2713.04	3
	6	2	837.28	837.28	11.81	2706.55	3
		3	756.67	756.67	906.91	2734.38	9
		4	608.39	608.39	1174.16	2723.04	7
		5	509.41	509.41	1041.85	2733.93	3
	9	2	739.01	739.01	19.95	2707.16	9
		3	614.04	614.04	534.24	2737.11	9
		4	525.31	525.31	1752.32	2726.73	7
		5	386.17	379.48	1361.94	2707.04	5
	12	2	722.07	722.07	23.02.20	2706.96	9
		3	744.10	744.10	1037.82	2738.47	7
		4	519.34	519.34	1516.31	2724.66	9
		5	395.57	395.57	1472.95	2728.43	5
	15	2	717.93	717.93	178.18	2721.96	5
		3	740.95	740.95	446.70	2747.28	9
		4	539.74	523.70	1491.13	2723.84	11
		5	415.79	409.39	1360.25	2717.48	9
0.4	3	2	911.91	911.91	7.50	2708.19	3
		3	918.66	918.66	206.73	2714.38	3
		4	848.07	848.07	7.45	2719.31	3
		5	805.08	805.08	7.57	2726.63	3
	6	2	658.67	658.67	7.54	2710.47	3
		3	623.00	623.00	684.99	2734.70	3
		4	582.57	582.57	135.58	2718.42	3

		5	520.59	520.59	654.52	2723.08	3
	9	2	614.20	614.20	121.16	2708.28	3
		3	518.68	518.68	623.30	2739.68	3
		4	491.94	491.94	290.89	2716.48	7
		5	367.19	367.19	859.72	2719.83	3
	12	2	590.33	590.33	19.65	2704.23	11
		3	640.46	640.46	787.79	2724.73	9
		4	454.28	454.28	1083.14	2738.21	9
		5	316.00	316.00	1254.92	2738.69	5
	15	2	581.80	581.80	31.19	2707.61	15
		3	633.64	633.64	1048.57	2732.25	9
		4	435.92	435.92	191.16	2740.76	9
		5	395.12	395.12	1342.81	2731.55	7
0.6	3	2	867.31	867.31	7.49	2704.90	3
		3	872.17	872.17	391.38	2725.24	3
		4	822.93	822.93	7.49	2718.95	3
		5	817.29	817.29	7.75	2711.54	3
	6	2	601.80	601.80	7.41	2707.23	3
		3	537.99	537.99	515.77	2723.46	3
		4	556.87	556.87	883.00	2730.12	3
		5	470.08	470.08	585.92	2721.22	3
	9	2	509.74	509.74	239.10	2705.76	3
		3	444.99	444.99	770.30	2721.83	3
		4	457.36	457.36	1078.24	2721.28	3
		5	378.45	378.45	374.47	2713.49	3
	12	2	479.07	479.07	208.37	2707.65	7
		3	441.00	441.00	798.56	2715.34	3
		4	414.76	414.76	17.38	2726.97	9
		5	340.60	340.60	851.81	2719.16	3
	15	2	466.14	466.14	140.00	2703.34	7
		3	425.65	425.65	253.93	2741.01	3
		4	393.23	393.23	19.38	2726.14	9
		5	302.05	302.05	1003.34	2721.18	11
0.8	3	2	872.62	872.62	7.64	2706.18	3
		3	837.81	837.81	7.33	2723.70	3
		4	811.85	811.85	7.29	2728.31	3
		5	804.65	804.65	7.25	2729.18	3
	6	2	537.99	537.99	151.82	2708.15	7
		3	561.92	561.92	100.55	2714.95	3
		4	513.44	513.44	783.51	2732.28	3
		5	420.14	420.14	7.47	2714.43	3
	9	2	456.01	456.01	133.33	2707.30	7
		3	469.15	469.15	256.63	2734.88	3
		4	428.22	428.22	7.82	2734.78	3
		5	298.70	298.70	403.05	2716.90	3
	12	2	425.58	425.58	54.11	2707.26	5
		3	429.00	429.00	170.97	2741.23	5
		4	380.06	380.06	13.16	2719.05	7
		5	246.74	246.74	1370.75	2726.22	3
	15	2	411.46	411.46	177.88	2705.13	5
		3	378.99	378.99	1026.28	2727.48	7
		4	355.46	355.46	371.41	2708.87	7
		5	236.09	236.09	1138.23	2729.96	3

Table C.2: Computational results when $|N|= 20$.

α	Θ	# of hubs	best profit	avg. profit	avg. best time (s)	avg. run time (s)	# of AH iterations
0.2	3	2	2360.79	2360.79	28.35	5574.83	3
		3	2607.31	2607.31	26.27	5713.41	3
		4	2302.71	2302.71	2204.03	5778.54	5
		5	1918.20	1918.20	24.32	5711.27	3
	6	2	1533.65	1533.65	65.63	5681.98	7
		3	1909.06	1909.06	1298.85	5797.70	7

		4	1526.17	1526.17	30.91	5835.92	5
		5	1323.32	1323.32	30.30	5785.08	3
	9	2	1482.13	1482.13	649.37	5593.85	3
		3	1722.73	1713.98	3972.49	5818.33	9
		4	1323.47	1323.47	885.26	5730.39	3
		5	1100.93	1095.77	3691.11	8324.00	7
	12	2	1463.47	1463.47	1553.89	5659.00	3
		3	1548.44	1548.44	1703.04	5961.20	9
		4	1370.97	1370.97	93.64	5661.87	11
		5	1200.88	1044.48	2351.25	6437.43	7
	15	2	1452.65	1452.65	1978.44	5594.24	3
		3	1748.54	1708.35	2553.09	6362.46	9
		4	1843.25	1683.40	3546.66	6142.89	9
		5	1177.65	1088.88	4590.68	7250.63	11
0.4	3	2	2310.87	2310.87	29.05.20	5684.43	3
		3	2347.05	2347.05	1722.20	5820.91	3
		4	2127.53	2126.53	1854.67	5935.58	5
		5	1918.20	1918.20	24.56	5816.71	3
	6	2	1347.99	1347.99	354.88	5558.66	3
		3	1672.47	1672.47	814.42	5792.69	3
		4	1385.23	1385.23	25.85	5848.23	3
		5	1262.84	1262.84	33.45	5885.28	3
	9	2	1210.84	1210.84	354.85	5546.78	3
		3	1488.36	1452.25	1858.32	5774.65	9
		4	1160.27	1160.27	42.89	5909.29	7
		5	1002.35	1002.35	25.90	5918.69	3
	12	2	1215.26	1215.26	1193.83	5698.25	3
		3	1384.41	1384.41	2682.93	5792.69	9
		4	1092.21	1092.21	62.61	5639.06	7
		5	888.46	887.97	817.80	5916.16	3
	15	2	1287.87	1287.87	735.55	5685.90	3
		3	1280.06	1280.06	1438.25	5744.66	7
		4	1300.01	1244.81	2318.94	5789.43	19
		5	838.46	813.31	1695.22	5980.33	13
0.6	3	2	2192.65	2192.65	251.21	5482.96	3
		3	2096.33	2096.33	25.30	5690.49	3
		4	2067.53	2067.53	823.65	5781.63	3
		5	1918.20	1918.20	24.16	5869.63	3
	6	2	1227.69	1227.69	836.33	5637.18	3
		3	1235.50	1235.50	26.10	5701.78	3
		4	1271.09	1271.09	950.01	5858.88	3
		5	965.33	964.74	2880.93	6072.91	3
	9	2	1055.38	1055.38	1953.66	5562.54	3
		3	1291.64	1291.64	2029.81	5878.83	7
		4	1023.25	1023.25	1620.63	5932.41	3
		5	874.88	849.52	1193.54	5932.39	3
	12	2	980.31	980.31	751.74	5603.32	3
		3	1099.46	1099.46	48.54	5864.22	9
		4	914.63	914.63	2372.54	5814.55	3
		5	816.40	816.40	2577.21	7101.87	3
	15	2	963.16	963.16	262.77	5567.88	3
		3	1059.03	1059.03	630.02	5741.19	7
		4	856.47	856.47	47.23	5824.35	9
		5	755.78	750.82	2153.71	6035.49	3
0.8	3	2	2150.07	2150.07	2210.31	5630.77	5
		3	2025.62	2025.62	24.48	5581.20	3
		4	1992.32	1992.32	24.08	5708.42	3
		5	1918.20	1918.20	24.84	5761.72	3
	6	2	1215.81	1215.81	513.18	5528.42	3
		3	1136.95	1136.95	625.12	5520.09	3
		4	1154.81	1141.63	2787.33	5516.87	3
		5	998.10	998.10	23.96	5523.76	3
	9	2	990.46	990.46	363.41	5567.28	5
		3	929.04	929.04	1768.25	5738.89	3
		4	879.91	879.91	967.83	5941.36	3

		5	738.81	735.32	1899.85	5724.01	3
	12	2	900.90	900.90	270.99	5478.82	5
		3	951.20	951.20	1338.27	5792.23	5
		4	772.04	772.04	2250.44	5950.53	3
		5	741.38	737.78	1824.40	6202.32	3
	15	2	857.08	857.08	452.63	5529.52	5
		3	882.82	882.82	51.07	5689.52	9
		4	701.26	701.26	2336.32	5985.80	3
		5	671.45	669.43	1461.02	7038.01	3

Table C.3: Computational results when $|N|= 25$.

α	Θ	**# of hubs**	**best profit**	**avg. profit**	**avg. best time (s)**	**avg. run time (s)**	**# of AH iterations**
0.2	3	2	3964.75	3964.75	76.62	7425.40	3
		3	2897.21	2897.21	73.16	7690.15	3
		4	2987.53	2987.53	76.88	8670.82	5
		5	2846.67	2846.67	62.75	18610.56	3
	6	2	3211.90	3211.90	98.53	7931.10	3
		3	1830.48	1830.48	3001.18	8262.13	3
		4	1856.14	1856.14	5518.06	8823.74	5
		5	1885.94	1885.94	106.68	18020.28	3
	9	2	3372.23	3372.23	163.67	7432.44	7
		3	3416.31	3416.31	196.10	9009.26	3
		4	1492.47	1492.47	4301.89	8558.04	7
		5	1585.74	1585.12	4722.33	17809.18	7
	12	2	3335.27	3335.27	309.79	8364.31	7
		3	3350.35	3350.35	298.74	10168.58	3
		4	1542.38	1542.38	5262.72	11399.40	7
		5	1426.69	1426.69	262.23	18302.15	13
	15	2	3326.70	3326.70	507.14	9483.55	9
		3	3315.50	3315.50	519.40	11463.31	11
		4	1570.85	1534.29	8830.86	13089.41	7
		5	1389.04	1365.94	1350.31	19418.49	11
0.4	3	2	3744.82	3744.82	77.49	7832.25	3
		3	2917.84	2917.84	71.04	7810.14	3
		4	3072.10	3072.10	76.70	8717.19	5
		5	2846.67	2846.67	61.99	17570.74	3
	6	2	2877.37	2877.37	86.85	7593.11	3
		3	1701.17	1701.17	2118.08	9333.17	3
		4	1784.25	1784.25	4284.25	10238.06	3
		5	1772.93	1772.93	96.12	18799.29	3
	9	2	2917.80	2917.80	1457.19	7734.01	7
		3	1420.49	1420.49	2271.12	7909.20	3
		4	1410.40	1410.40	6143.86	12128.14	3
		5	1398.26	1394.73	5917.25	17787.95	3
	12	2	2853.64	2853.64	1215.19	7976.16	7
		3	1314.40	1314.40	4602.14	8660.29	3
		4	1372.37	1310.73	5306.78	9564.79	7
		5	1248.94	1242.72	11761.19	20554.21	11
	15	2	2850.34	2850.34	219.19	7506.48	5
		3	2677.09	1556.12	4093.62	10413.41	3
		4	1348.53	1348.53	8719.52	11920.27	7
		5	1171.55	1169.06	27105.49	45130.45	9
0.6	3	2	3581.94	3581.94	71.81	7294.51	3
		3	2918.16	2918.16	3504.92	8935.26	3
		4	2969.62	2969.62	80.61	9073.43	5
		5	2856.18	2856.18	66.36	19255.00	3
	6	2	2632.74	2632.74	145.30	7489.43	7
		3	1601.90	1601.90	78.07	7979.25	3
		4	1643.08	1643.08	74.47	9037.13	5
		5	1678.25	1678.25	3861.36	21822.10	3
	9	2	2378.74	2378.74	1411.26	7711.97	7
		3	1279.68	1279.68	5773.61	8339.32	3

		4	1246.93	1246.93	4350.02	9618.32	5
		5	1271.00	1271.00	93.28	19816.06	3
	12	2	2516.95	2516.95	3432.92	8098.90	9
		3	1179.43	1179.43	4597.74	9033.84	3
		4	1090.86	1074.97	6547.52	11024.49	7
		5	1100.48	1087.48	661.21	20273.23	7
	15	2	2480.57	2480.57	3737.54	8054.97	9
		3	1127.98	1127.98	5428.63	8545.31	3
		4	1086.14	1059.50	9765.78	13832.73	7
		5	995.43	981.60	19112.10	39044.33	9
0.8	3	2	3470.91	3470.91	70.99	8371.32	3
		3	2893.48	2893.48	69.80	7969.14	3
		4	2933.44	2933.44	65.59	8914.90	3
		5	2885.94	2878.06	1810.40	19832.70	5
	6	2	2499.39	2499.39	1203.62	7408.57	5
		3	1564.99	1564.99	70.92	8137.43	3
		4	1566.54	1566.54	69.11	9129.86	3
		5	1642.03	1642.03	2516.96	19991.80	5
	9	2	2222.45	2222.45	2484.28	7464.88	9
		3	1221.78	1221.78	5355.40	8116.63	3
		4	1164.71	1164.71	4971.40	10462.57	3
		5	1226.53	1220.94	295.55	20113.28	5
	12	2	2113.27	2113.27	1826.9	7395.91	9
		3	1073.13	1073.13	5335.07	8914.37	3
		4	991.87	987.23	7153.58	10365.20	3
		5	1044.61	1044.61	10090.73	28191.64	5
	15	2	2061.04	2061.04	1553.67	7487.22	9
		3	1012.37	1012.37	3522.70	8409.51	3
		4	902.61	898.62	3526.68	9893.87	3
		5	1015.85	987.20	28373.89	44492.36	7

Table C.4: A leader's profit when she ignores the follower's reaction.

$\lvert N\rvert$	α	Θ	# of hubs	Profit achieved by p-HMLP	Profit achieved by p-HCLP
		3	2	1008.68	643.47
			3	916.42	535.67
			4	907.36	548.13
			5	804.17	495.88
		6	2	837.28	351.21
			3	530.81	272.70
			4	523.34	276.35
			5	493.05	193.71
	0.2	9	2	710.04	281.97
			3	437.17	165.73
			4	421.36	221.25
			5	305.13	127.63
		12	2	685.52	255.36
			3	449.09	133.73
			4	377.42	202.32
			5	249.64	104.18
		15	2	675.37	242.72
			3	430.24	121.76
			4	372.61	194.00
			5	235.90	93.57
		3	2	911.91	669.80
			3	902.25	536.78
			4	848.07	517.93
			5	805.08	610.73
		6	2	658.67	368.13
			3	520.59	278.39
			4	484.38	208.51
			5	462.93	302.42
			2	584.22	298.14

15	0.4	9	3	421.45	223.85
			4	375.60	136.50
			5	351.32	202.12
		12	2	570.95	271.78
			3	378.34	203.90
			4	326.80	108.97
			5	300.65	166.77
		15	2	555.33	334.46
			3	412.30	132.43
			4	311.82	95.54
			5	272.63	148.66
	0.6	3	2	867.31	704.00
			3	851.70	450.17
			4	822.93	475.56
			5	817.29	563.55
		6	2	601.80	346.26
			3	496.38	216.69
			4	451.13	195.06
			5	439.20	256.86
		9	2	473.56	444.62
			3	390.78	166.73
			4	336.81	131.75
			5	320.78	185.25
		12	2	423.20	419.01
			3	343.27	148.77
			4	283.86	108.98
			5	265.42	150.07
		15	2	406.21	407.61
			3	317.34	140.55
			4	254.22	98.55
			5	234.32	131.48
	0.8	3	2	872.62	491.28
			3	837.81	514.82
			4	811.85	451.70
			5	804.65	446.65
		6	2	519.16	194.68
			3	480.86	283.85
			4	433.84	192.78
			5	420.14	180.16
		9	2	413.88	103.11
			3	374.51	233.61
			4	428.22	137.41
			5	296.37	119.40
		12	2	403.31	74.78
			3	312.22	214.66
			4	380.06	116.95
			5	236.66	95.53
		15	2	383.31	59.68
			3	331.68	205.72
			4	353.79	107.08
			5	202.08	83.75
		3	2	2360.79	1927.32
			3	2607.31	1489.45
			4	2040.51	1075.52
			5	1918.20	1213.78
		6	2	1186.58	1311.31
			3	1813.20	803.60
			4	1522.28	455.38
			5	1323.32	463.52
		9	2	944.60	1190.83
			3	1622.02	652.77
			4	1312.68	231.02
			5	1051.20	294.57
			2	840.68	1155.30
			3	1546.44	599.48

0.2				
	12	4	1370.97	181.85
		5	931.71	232.72
		2	786.46	1146.11
	15	3	1509.11	576.24
		4	1141.34	159.20
		5	905.93	203.68
		2	2310.87	1476.94
	3	3	2177.9	1461.68
		4	1996.30	1500.84
		5	1918.20	1335.55
		2	1184.61	826.83
	6	3	1617.54	726.74
		4	1385.23	798.90
		5	1262.84	548.36
		2	935.82	681.96
0.4	9	3	1398.07	560.35
		4	1160.27	636.38
		5	1002.35	355.38
		2	826.40	634.22
	12	3	1306.04	476.42
		4	1065.37	577.76
		5	887.23	278.23
		2	768.21	618.33
	15	3	1258.88	445.43
		4	1015.61	551.71
		5	751.13	238.87
		2	2167.84	1594.71
	3	3	2096.33	1556.62
		4	2046.12	1329.00
		5	1918.20	1477.64
		2	1152.61	994.55
	6	3	1235.50	705.79
		4	1162.98	576.99
		5	959.10	666.41
		2	895.46	845.00
0.6	9	3	1208.32	525.88
		4	896.65	392.74
		5	811.48	442.09
		2	779.30	796.47
	12	3	1099.46	452.67
		4	773.88	323.14
		5	671.25	344.38
		2	715.81	776.39
	15	3	1040.63	416.14
		4	593.86	290.65
		5	591.31	291.18
		2	2000.41	1538.43
	3	3	2025.62	1725.3
		4	1992.32	1568.11
		5	1918.2	1128.85
		2	1118.59	768.62
	6	3	1136.54	934.90
		4	1072.55	823.98
		5	998.10	432.12
		2	854.11	573.80
	9	3	861.12	714.29
		4	852.98	618.24
		5	724.4	266.43
		2	733.20	494.68
	12	3	732.23	615.89
		4	717.80	398.99
		5	575.61	202.54
		2	666.13	455.30
		3	632.85	561.92
		4	640.89	348.62

	15			
		5	488.83	170.39
0.2	3	2	3964.75	3318.62
		3	2897.21	2084.10
		4	2968.28	1844.93
		5	2846.67	2354.66
	6	2	3211.9	2368.23
		3	1527.28	1082.63
		4	1635.91	869.82
		5	1885.94	1100.75
	9	2	3372.23	2182.14
		3	3416.31	831.53
		4	1324.9	473.59
		5	1582.63	761.25
	12	2	3335.27	2127.21
		3	3350.35	611.59
		4	1141.86	393.91
		5	1417.66	618.84
	15	2	3326.7	2114.51
		3	3192.87	568.67
		4	1039.73	356.81
		5	1331.28	544.75
0.4	3	2	3744.82	1029.50
		3	2917.84	2124.09
		4	3062.24	1725.04
		5	2846.67	2088.45
	6	2	2877.37	348.45
		3	1554.85	929.05
		4	1763.09	739.87
		5	1772.93	983.05
	9	2	2611.28	259.56
		3	1133.72	636.38
		4	1232.04	533.43
		5	1371.25	696.90
	12	2	2538.79	232.40
		3	940.59	487.40
		4	1214.33	343.20
		5	1087.39	700.33
	15	2	2850.34	220.96
		3	834.26	431.76
		4	1120.96	302.62
		5	975.97	645.92
0.6	3	2	3581.94	1136.59
		3	2912.81	1392.89
		4	2947.22	2180.6
		5	2856.18	2099.88
	6	2	2632.74	409.19
		3	1601.9	662.54
		4	1597.58	994.83
		5	1661.88	971.11
	9	2	2329.93	304.93
		3	1191.72	526.78
		4	1173.95	664.89
		5	1271.00	691.48
	12	2	2234.13	271.27
		3	1006.94	475.38
		4	820.19	543.20
		5	1078.81	555.07
	15	2	2169.2	256.60
		3	814.49	450.98
		4	694.24	481.12
		5	972.37	497.84
	3	2	3470.91	1158.58
		3	2893.48	1950.10
		4	2933.44	1498.91
		5	2865.12	1957.62

0.8	6	2	2347.13	623.66
		3	1564.99	911.69
		4	1566.54	616.64
		5	1628.15	902.87
	9	2	2177.10	507.96
		3	1144.97	636.44
		4	1128.14	402.43
		5	1204.50	618.71
	12	2	2063.63	464.80
		3	952.26	558.41
		4	915.98	317.13
		5	973.30	480.58
	15	2	2010.62	444.90
		3	845.62	493.60
		4	792.39	274.07
		5	852.55	416.27

Alphabetical Index

Name Index

Curriculum Vitae

Education

Master of Sciences in Mathematics, University of Belgrade, 2012

Bachelor of Sciences in Mathematics and Computer Science, University of Banja Luka, 2010

Academic Employment

Senior Teaching Assistant, University of Banja Luka, 2013–present.

Teaching Assistant, University of Banja Luka, 2010–2013

Publications

Journal Papers

Čvokić, D.D.: A leader–follower single allocation hub location problem under fixed markups. Filomat. 34(8), 2463–2484 (2020). doi: https://doi.org/10.2298/FIL2008463C (WoS journal, $IF_{2019} = 0.848$)

Čvokić, D.D.: Cutting testing costs by the pooling design. MTC. 68(4), 743–759 (2020). https://doi.org/10.5937/vojtehg68-28078

Čvokić, D.D., Stanimirović, Z.: A single allocation hub location and pricing problem. Comp. Appl. Math. 39 (2020). https://doi.org/10.1007/s40314-019-1025-z (WoS journal, $IF_{2019} = 1.883$)

Vidić, S., Kremenović, D., Čvokić, D.D.: Bio-Linux in the light of modern bioinformatics‘ tendencies. Info M. 16(61), 26–34 (2017).

Conference Papers

Čvokić D.D., Kochetov Y.A., Plyasunov A.V., Savić A. (2019) The Competitive Hub Location Under the Price War. In: Khachay M., Kochetov Y., Pardalos P. (eds) Mathematical Optimization Theory and Operations Research. MOTOR 2019. Lecture Notes in Computer Science, vol 11548. Springer, Cham

Čvokić D.D., Kochetov Y.A., Plyasunov A.V. (2017) The Existence of Equilibria in the Leader–Follower Hub Location and Pricing Problem. In: Dörner K., Ljubic I., Pflug G., Tragler G. (eds) Operations Research Proceedings 2015. Operations Research Proceedings (GOR (Gesellschaft für Operations Research e.V.)). Springer

Čvokić D.D., Kochetov Y.A., Plyasunov A.V. (2016) A Leader–Follower Hub Location Problem Under Fixed Markups. In: Kochetov Y., Khachay M., Beresnev V., Nurminski E., Pardalos P. (eds) Discrete Optimization and Operations Research. DOOR 2016. Lecture Notes in Computer Science, vol 9869. Springer

Čvokić D.D. (2015) Elastic Demand Pricing in Competitive Single Allocation Hub Location and Pricing Problem. In: Mladenović N., Urošević D., Stanimirović Z. (eds) Proceedings SYM-OP-IS 2015. XLII International Symposium on Operational Research. Mathematical Institute of SANU

Образац -3

УНИВЕРЗИТЕТ У БАЊОЈ ЛУЦИ
ФАКУЛТЕТ: ПРИРОДНО-МАТЕМАТИЧКИ

ИЗВЈЕШТАЈ
о оцјени урађене докторске дисертације

I ПОДАЦИ О КОМИСИЈИ

На основу члана 149. Закона о високом образовању, члана 54. Статута Универзитета у Бањој Луци, Наставно-научно вијеће Природно-математичког Факултета Универзитета у Бањој Луци, на сједници одржаној дана 16.11.2020. г., донијело је под бројем 19/3.2570/20 Одлуку о именовању Комисије за писање извјештаја, преглед, и оцјену урађене докторске дисертације кандидата мр Димитрија Д. Чвокића под називом: „*ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ (r|p) ХАБ-ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ*“, у сљедећем саставу:

1. Др Александар В. Пљасунов, доктор математичких наука са хабилитацијом, доцент на ужој научној области Дискретна математика и теоријска кибернетика (ВАК РФ 01.01.09.), Катедра за теоријску кибернетику Механичко-математичког факултета Националног истраживачког државног универзитета у Новосибирску (Новосибирск, Русија), предсједник;

2. Проф. др Милан В. Јовановић, редовни професор на ужој научној области Математичка анализа и примјене, Природно-математички факултет Универзитета у Бањој Луци, члан;

3. Др Драган Кораћ, доцент на ужој научној области Рачунарске науке, Природно-математички факултет Универзитета у Бањој Луци, члан;

4. Проф. др Александар Љ. Савић, ванредни професор на ужој научној области Нумеричка математика и оптимизација, Катедра за нумеричку математику и оптимизацију Математичког факултета Универзитета у Београду (Београд, Србија), коментор, члан;

5. Проф. др Јуриј А. Кочетов, редовни професор на ужој научној области Дискретна математика и теоријска кибернетика (ВАК РФ 01.01.09.) Катедра за теоријску кибернетику Механичко-математичког факултета Националног истраживачког државног универзитета у Новосибирску (Новосибирск, Русија), ментор, члан;

Одлуку Наставно-научног вијећа Природно-математичког факултета је потврдио Сенат Универзитета у Бањој Луци на сједници одржаној 26.11.2020. г., одлуком под бројем 02/02-3.2708/44/20. Комисија је у предвиђеном року прегледала докторску дисертацију кандидата мр Димитрија Д. Чвокића и о томе Наставно-научном вијећу Природно-математичког факултета и Сенату Универзитета у Бањој Луци подноси сљедећи извјештај.

1) Навести датум и орган који је именовао комисију;
2) Навести састав комисије са назнаком имена и презимена сваког члана, научно-наставног звања назива уже научне области за коју је изабран у звање и назива универзитета/факултета/института на којем је члан комисије запослен.

II ПОДАЦИ О КАНДИДАТУ

Име, име једног родитеља, презиме: *Димитрије (Душан) Чвокић*

Датум рођења, општина, држава: *08.11.1984. г., Ливно, Босна и Херцеговина*

Назив универзитета и факултета и назив студијског програма академских студија II циклуса, односно послиједипломских магистарских студија и стечено стручно/научно звање:

- Назив универзитета и факултета: *Математички факултет Универзитета у Београду*
- Студијски програм II циклуса: *Примењена математика*
- Стечено стручно/научно звање: *Мастер математичар*

Факултет, назив магистарске тезе, научна област и датум одбране магистарског рада:

- Факултет: *Математички факултет Универзитета у Београду*
- Назив магистарске/мастер тезе: *Гоморијеве одсијецајуће равни – развој и примјене*
- Научна област: *Математика*
- Датум одбране магистарског/мастер рада: 24.09.2012. г.

Година уписа на докторске студије и назив студијског програма: *2017, Математика.*

1) Име, име једног родитеља, презиме;
2) Датум рођења, општина, држава;
3) Назив универзитета и факултета и назив студијског програма академских студија II циклуса, односно послиједипломских магистарских студија и стечено стручно/научно звање;
4) Факултет, назив магистарске тезе, научна област и датум одбране магистарског рада;
5) Научна област из које је стечено научно звање магистра наука/академско звање мастера;
6) Година уписа на докторске студије и назив студијског програма.

III УВОДНИ ДИО ОЦЈЕНЕ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

Наслов докторске дисертације и орган који је прихватио тему

Тему докторске дисертације под називом „*ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ (r|p) ХАБ-ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ*“ (енг. „EXISTENCE AND SOLUTION METHODS FOR STACKELBERG EQUILIBRIA IN THE *(r|p)* HUB-CENTROID PROBLEM UNDER THE PRICE WAR”) кандидата мр Димитрија Д. Чвокића, прихватило је Наставно-научно вијеће Природно-математичког факултета Универзитета у Бањој Луци одлуком под бројем 19/3.1525/20, дана 15.07.2020. Сенат Универзитета у Бањој Луци одлуком под бројем 02/04-3.1604-27/20, дана 23.07.2020. г., дао је сагласност на поменути Извјештај о оцјени услова и подобности теме и кандидата за израду докторске дисертације студента III циклуса мр Димитрија Д. Чвокића, на Природно-математичком факултету Универзитета у Бањој Луци.

Садржај дисертације са страничењем

Contents

Основни подаци о докторској дисертацији

Језик дисертације је енглески, у складу са одлуком Сената под бројем 02/04-3.2708-43/20, дана 26.11.2020. г., при чему *насловна страна, захвалница, и резиме имају своје преводе на српски језик како је то уређено Правилником о садржају, изгледу, и дигиталном репозиторијуму докторских дисертација на Универзитету у Бањој Луци*. Дисертација има 154 стране А4 формата и издијељена је на седам глава: Introduction, The *(r|p)* Hub-Centroid Problem under the Price War, Solution Approach, Computational Experiments, Conclusion, Bibliography (127 библиографских јединица), и Appendices. Поред поменутих глава, присутне су прелиминарије, индекси, и биографија: списак скраћеница (List of Abbreviations), списак слика (List of Figures), списак табела (List of Tables), списак кодова (List of Codes), списак алгоритама (List of Algorithms), индекс појмова (Alphabetical Index), индекс имена (Name Index), и биографија (Curriculum Vitae). У тексту се налазе четири слике, 27 графикона, 11 табела, три пајтон-кода и пет алгоритама представљених псеудокодом.

1) Наслов докторске дисертације;
2) Вријеме и орган који је прихватио тему докторске дисертације
3) Садржај докторске дисертације са страничењем;
4) Истаћи основне податке о докторској дисертацији: обим, број табела, слика, шема, графикона, број цитиране литературе и навести поглавља.

IV УВОД И ПРЕГЛЕД ЛИТЕРАТУРЕ

Разлози за предузимање истраживања

Проблем размјештања разводних тачака (енгл. Hub Location Problem (HLP)) представља савременију област истраживања у теорији размјештања (енг. Location Theory). Уопштено, ријеч је о тражењу оптималног размјештаја разводних тачака и повезивања преосталих неразводних са разводним имајући у виду задат и јасно формулисан циљ образовања транспортне мреже. Разводне тачке служе за консолидацију и дисеминацију током рутирања робе или путника између полазишта и одредишта (енг. Origin-Destiantion pairs (O-D парови)). Овај концепт је уведен ради смањења укупног броја непосредних веза међу тачкама (чворовима диграфа) и искоришћавања економског ефекта који има повећана количина протока између разводних чворова (попуст на количину, енг. economies of scale), а који се природно јавља у оваквим поставкама. Због своје важности како из теоријског, тако и из практичног угла, HLP се у посљедње вријеме интензивно проучава.

Истраживање проблема размјештања се на Математичком институту „Собољев“ (Академ-град, Новосибирск, Русија) врши од раних 60-их година прошлог вијека. Међу њима се прије свега могу истаћи радови Бересњева В.Л., Гимадија Е.Х., Дементијева В.Т., Шамардина Ј.В., Колоколова А.А., Антипина А.С., Хамисова О.В., Васиљева И.Л., Забудског Г.Г., Левановке Т.В., и многих других. На Универзитету у Београду, тиме су се бавили Дугошија Ђ., Младеновић Н., Дражић З., Кратица Ј., Филиповић В., Станимировић З., Марић М., и други. Међутим, проблеми размјештања *разводних тачака* у конкурентном окружењу, и то у спрези са цијенама, тек одскора се разматрају у научној литератури.

У БиХ, као и земљама окружења, ово је прва дисертација у којој се проучава ова проблематика, а уједно и прва у којој се разматрају проблеми оптимизације у више нивоа.

Проблем истраживања

Двије конкурентне транспортне компаније намеравају да уђу на тржиште, што се сматра општепознатим (имајући у виду дефиницију тог термина у теорији игара). Обје теже ка максимизацији профита проналажењем најбољих разводних мрежа уз образовање одговарајуће цјеновне структуре. Управа једне компаније жели да размјести p разводних тачака, а друга планира да размјести њих r. Након инсталације одговарајућих мрежа, очекује се да ће компаније кренути са цјеновним надметањем. Рјешење овог надметања, ако постоји, је Бертранд-Нешов цјеновни еквилибријум у којем ниједној од компанија-такмаца није исплативо да једнострано промијени своју цјеновну одлуку.

Обично се у литератури разматрају два сценарија: симултани и секвенцијални улазак на тржиште. У првом сценарију цјеновно надметање је природна претпоставка. Проблем је у томе што бисмо могли очекивати да постоји више Нешових еквилибријума када је ријеч о разводним мрежама. Проналажење доминантног Нешовог еквилибријума за једну од компанија-такмаца би могао бити поприлично исцрпљујући задатак, гледано из рачунског угла. Штавише, доминантни Нешов еквилибријум не мора бити састављен од чистих стратегија, тј. може се пројавити у облику циклуса повезаних најбољих рекација. Стандардно тумачење Нешовог еквилибријума састављеног од мјешовитих стратегија није прихватљиво. Компанија неће „бацати новчић“ како би одабрала своју мрежну топологију. У другом сценарију, надметање ценама се не претпоставља, тј. прва компанија која уђе на тржиште посвећена је својој одлуци о размјештању разводних тачака и цијена. Ипак, коначно решење Штаклберга подразумијева постојање допустивог Бертранд-Нешовог цјеновног еквилибријума, који отвара врата кооперативној игри цијена са раздјељивим добитком. Стога, узимајући све ово у обзир, има смисла бити заинтересован за разматрање међуваријанте, тј. Штаклберговог надметања са цјеновним надметањем. Једна компанија, leader, улази на тржиште као први конкурент (први такмац), и свјесна је (зна) да ће за њом ући још једна – follower (други конкурент/такмац). С друге стране, цијене се формирају према рјешењу симултане цјеновне игре. Ова поставка имплицира да leader не намеће цијене follower-у, нити га игнорише, већ су оне у суштини резултат улаза follower-а на тржиште (надметања које тад креће). Могли бисмо рећи да је сценарио еквивалентан потрази за Штаклберговом стратегијом у симултаној игри (представљеној у нормалном/матричном облику), а могућа је и веза са потрагом за *status quo* позицијом када би се разматрале кооперативне цијене.

Предмет истраживања

Формално, *поставка проблема* почива на комплетном диграфу G = (N,A), гдје је N непразан скуп и A је скуп ивица. Разводном тачком може постати искључиво неки чвор $k \in N$, при чему чвор може бити дијељен међу такмацима и нема ограничења на капацитет. Одабир чвора који ће бити разводна тачка се сматра стратешком одлуком. Први и други такмац размјештају *p* и *q* разводних тачака, респективно. Неразводне тачке могу бити, такође, само чворови диграфа G, али различити од оних који су одабрани за разводне. За сваку ивицу $(i,j) \in A,$ постоји транспортни трошак по јединици робе или по путнику $c_{ij} \geq 0.$ За сваки O-D пар $(i,j) \in N^2,$ такмац може да понуди само једну транспортну руту. Вишеструка повезивања неразводних тачака са разводним су дозвољена. Једна неразводна тачка не може бити непосредно повезана са другом неразводном. Транспортни коефицијенти χ, α, и δ су познати за дато тржиште и респективно одговарају консолидацији робе/путника од полазишта до одговарајуће сусједне разводне тачке са којом је повезана, транспорту између двије разводне тачке, и дистрибуцији на крајње одредиште. Надовезивање ивица формира руту. Највише двије разводне тачке могу да буду на једној рути. Трошкови транспорта на рути i→k→l→j су дати према $c_{ij,kl} = \chi c_{ik} + \alpha c_{kl} + \delta c_{lj},$ за све $i,j,k,l \in N.$ Клијенти бирају руте превасходно према понуђеним цијенама. За оба учесника је карактеристично образовање mill-цијена. Логит-модел се користи за процјену конзумерског/дискретног избора (енг. Qualitative Choice Models). Карактерише га параметар осјетљивости Θ, са већ познатом ненегативном вриједношћу, добијеном економетријским испитивањем. За веће Θ клијенти су осјетљивији на разлике у цијенама, док за мање Θ, природно, важи обрнуто. Мрежа разводних тачака мора да покрива све чворове диграфа G, тј. услуге су понуђене свима на тржишту.

Циљ истраживања

У ширем смислу, циљ истраживања је проучавање спрега цијена и мреже разводних тачака (хабова) у конкурентном окружењу. Уже гледано, циљ истраживања је доказ постојања Штаклбергових еквилибријума у проблему *(r|p)* хаб-центроида са цјеновним надметањем (*(r|p)*HCPuPW), карактеризација оптималних рјешења, испитивање рачунске сложености, те конструкција матхеуристичких алгоритама за рјешавање проблема leader-а.

Хипотезе

Главна хипотеза: постоји Штаклбергов еквилибријум за проблем *(r|p)* хаб-центроида са цјеновним надметањем.

Помоћне хипотезе:

1. Постојање Штаклберговог еквилибријума је омогућено постојањем Бертранд-Нешовог цјеновног еквилибријума при логит-моделу конзумерског избора;
2. Руте са најмањим транспортним трошковима су оптималне;
3. Проблем follower-a се може реформулисати као мјешовито-цјелобројни линеарни математички програм;
4. *(r|p)*HCPuPW је *NP*-тежак;
5. Проблем follower-a (*(r|p)*HMPuPW и auxiliary) је *NP*-тежак;
6. Изведени алокацијски проблеми follower-a припадају класи *P*;
7. Надметање цијенама може значајно да утиче на топологију разводних тачака оба такмаца у разматраном Штаклберговом надметању.

Резултати претходних истраживања

Одлуке о размјештању и формирању мреже су према својој природи стратешке, док се образовање цијена налази на нижем, тактичком, или пак оперативном нивоу. Због овога се разматрање везе између размјештања и цијена може учинити необичним на први поглед. Ипак, показано је да одлука о топологији разводне мреже коју ће да користи компанија за транспорт увелико зависи од прихода. У том погледу, приходи су у исто вријеме зависни од цјеновне структуре. Стога, компанија која планира своју транспортну мрежу тешко да може донијети добру одлуку кад је ријеч о изгледу транспортне мреже, уколико неће узимати у обзир утицај цијена. С друге стране, на формирање цијена утиче сама транспортна мрежа, тј. имамо спрег стратешких и тактичких/оперативних одлука (Luer-Villagra and Marianov (2013), Čvokić et al. (2016), Čvokić at al. (2019), Čvokić and Stanimirović (2020)). Поред цијена, тржишта карактеришу и олигополи. Профит компаније не зависи само од одлуке њеног руководства, него и од потеза конкруенције (Sasaki and Fukushima (2001), Čvokić et al. (2019), Kononov et al. (2019)).

Улазак на тржиште компанија које логистику заснивају на топологији разводних мрежа се углавном у литератури разматрао из угла Штаклбергових надметања. Први рад на ову тему је објављен 2001. г (Sasaki and Fukushima (2001)). Разматран је непрекидни Штаклбергов HLP у којем се инкумбент бори за максимизацију профита узимајући у обзир више нападача. За сваку руту је само једна разводна тачка била дозвољена. Студија Махмутогуларија и Каре из 2016. г. (Mahmutogulari and Kara (2016)) ставља у фокус *(r|p)* хуб-центроид проблем, тј. HLP је комбинаторан по својој природи, а не непрекидан, а на рути могу бити једна или двије разводне тачке. У поменутом раду потражња се међу компанијама дијели према упроштеном моделу „побједник узима све“ – ко нуди нижу цијену за конкретан O-D пар „преузима“ све клијенте. Циљ проблема је максимизација удјела у тржишту. У раду је понуђен и егзактан приступ рјешавања.

Упроштен модел подјеле потражње „побједник узима све“ не одражава вјерно стварно стање. Поред надметања, у олигополима је потребно узети у обзир понашање клијената, а претпоставља се да су цијене важан критеријум у одлучивању корисника услуга. Логит-модел, заснован на логистичкој регресији, је изразито кориштен у транспортној литератури и тренутно је међу најчешће коришћеним моделима за моделовање конзумерског/дискретног избора (Garrow (2010), Luer-Villagra and Marianov (2013), Čvokić et al. (2019)). Омогућује добру процјену расподјеле клијената у случају варијабилних услуга са различитим цијенама, а додатно га карактерише затворен математички израз.

Упоређивање метахеуристичких алгоритама за конкруентне проблеме размјештања и образовања цијена (али не и размјештања разводних тачака) је дат у раду Кочетва и сарадника (Kochetov et al. (2015)). Може се рећи да је дати рад полазни за разматрање алгоритамског приступа у оваквим и сличним проблемима.

Поменуте референце

- Čvokić, D.D., Kochtov, Y.A., Plyasunov, A.V.: A leader-follower hub location and pricing problem under fixed markups. In; Y.A. Kochetov et al. (eds.) DOOR 2016 LNCS 9869, pp. 350-363, Srpinger, Cham (2016)
- Čvokić, D.D., Kochetov, Y.A., Playsunov, A.V., Savić, A.: The competitive hub location under the price war. In: Khacahy M., Kochetov Y., Pardalos P. (eds.) Mathematical Optimization Theory and Operations Research. MOTOR 2019. Lecture Notes in Computer Science, vol 11548, pp. 133-146. Springer, Cham (2019)
- Čvokić, D.D, Stanimirovic, Z.: A single allocation hub location and pricing problem. Comp. Appl. Math. 39, 40 (2020)
- Garow, L.A.: Discrete choice modelling and air travel demand: theory and applications. Routledge, London, UK (2010)
- Kochetov, Y., Panin, A., Plyasunov, A.: Comparison of metaheuristics for the bilevel facility location and mill pricing prolbem. Journal of Applied and Industrial Mathematics 9(3), 392-401 (2015)
- Kononov, A.V., Panin, A.A., Plyasunov, A.V.: A bilevel competitive location and pricing model with nonuniform split of demand. J. Appl. Ind. Math. 13, 500-510 (2019)
- Luer-Villagra, A., Marianov, V.: A competitive hub location and pricing problem. European Journal of Operational Research 231(3), 734-744 (2013)
- Mahmutogullari, I., Kara, B.: Hub location under competition. European Journal of Operational Research 250(1), 214-225 (2016)
- Panin, A., Paschchenko, M., Plyasunov, A.: Bilevel competitive facility location and pricing problems. Automation and Remote Control, 75(4), 715-727 (2014)
- Sasaki, M., Fukushima, M.: Stackelberg hub location problem. Journal of the Operations Research Society of Japan 44(4), 390-405 (2001)

Допринос тезе

Ријеч је о новом најреалистичнијем до сада моделу у два нивоа што се тиче размјештања разводних тачака (хабова), јер укључује конкурентно окружење са цјеновним надметањем. Математички је доказано постојање Бертранд-Нешовог цјеновног еквилибријума при логит-моделу конзумерског/дискретног избора. Из доказа су изведене нове трансцедентне једначине најбоље реакције. Дати резултати су послужили за доказивање постојања Штаклберговог еквилибријума у датом проблему. Поред тога, разматрана је карактеризација рута. Доказано је да оптималне руте одговарају оним које одликују најмањи транспортни трошкови (што не мора да буде случај у неким другим поставкама). Показано је како се може преформулисати иницијално нелинеаран модел follower-а (другог такмаца) у мјешовито-цјелобројни линеарни програм. Такође је показано како се auxiliary (допунски) модел може реформулисати као мјештовито-цјелобројни линеарни програм. Доказано је да су проблеми follower-а, auxiliary модела, и проблем $(r|p)$ хаб-центроида са цјеновним надметањем NP-тешки. С друге стране, доказано је да изведени алокацијски проблеми припадају класи P. Конструисани су алгоритми за проналажење Штаклберговог еквилибријума засновани на алтернацијској хеуристици и методу промјенљивих околина (МПО). За модификацију МПОа је доказано оптимизацијско својство. Стручни допринос се односи на имплементацију одговарајућих алгоритама за рјешавање овог проблема, њихово тестирање, и документовање испитивања на САВ-инстанцама. Другим ријечима, дати су одговори на све постављене хипотезе.

Очекивани научни и прагматични доприноси дисертације

Резултати овог истраживања наћи ће своје мјесто у теорији размијештања која је једна од тренутно најразрађенијих теорија операционих истраживања (научни допринос). Поред тога, ово истраживање може да послужи као основа за даљи развој у области размјештања разводних тачака у конкурентном окружењу, а представљени модели и теореме би могли бити укључени и у наставу математике и информатике на постдипломским студијама. Поред тога, могуће конкретне примјене (прагматични допринос) добијених резултата су:

1) образовање путне мреже за авиотранспорт путника и робе;
2) образовање путне мреже у поморском транспорту;
3) образовање мреже поштанских услуга;
4) образовање телекомуникационе мреже;
5) размјештај пунктова служби хитне помоћи, спасилачких служби, санитетског превоза, и патронажних служби;
6) *Supply Chain Management* логистика;
7) у теорији концентрације криминала и криминологији мјеста.

1) Укратко истаћи разлог због којих су истраживања предузета и представити проблем, предмет, циљеве и хипотезе;
2) На основу прегледа литературе сажето приказати резултате претходних истраживања у вези проблема који је истраживан (водити рачуна да обухвата најновија и најзначајнија сазнања из те области код нас и у свијету);
3) Навести допринос тезе у рјешавању изучаваног предмета истраживања;
4) Навести очекиване научне и прагматичне доприносе дисертације.

V МАТЕРИЈАЛ И МЕТОД РАДА

За теоријски дио су коришћени савремени методи операционих истраживања који укључују математичко моделовање оптимизацијских проблема, моделе конзумерског/дискретног избора, резултате конвексне анализе, теорију игара, теорију рачунске сложености, метахеуристике, и теорију локалног претраживања. Кад је ријеч о практичном (експерименталном) дијелу истраживања, конструисани алгоритми су тестирани на CAB-инстанцама, а за анализу нумеричких резултата је коришћена дескриптивна статистика. Треба имати у виду да сама операциона истраживања спадају у тзв. прескриптивну аналитику.

Теоријски правац истраживања је базиран на сљедећем:

- преглед, осврт, и проучавање постојеће научне и стручне литературе која се тиче проблема разводних тачака, Штаклберговог надметања, и образовања цијена;
- методе математичког моделовања оптимизацијских проблема који се односе на логистичке проблеме уз коришћење резултата теорије игара (енг. Game Theory) и конзумерског/дискретног избора;
- постављање хипотеза;
- коришћење познатих математичких резултата (теорема) и техника диференцијалног рачуна, конвексне анализе, теорије игара, и математичке оптимизације с циљем доказивања или пак оповргавања постављених хипотеза;
- технике реформулације нелинеарних математичких програма ради проналажења реформулације проблема follower-а и auxiliary модела које би омогућиле њихово рјешавање state-of-the-art рјешавачима;

- конструкцији алгоритама заснованих на мат- и метахеуристикама.

Практични дио истраживања је базиран на сљедећем:

- имплементацији алгоритама на рачунарском програмском језику Python, уз коришћење одговарајућих помоћних програмских пакета;
- извођењу рачунских огледа на CAB инстанцама;
- дескриптивној статистици за анализу добијених нумеричких резултата.

За обраду резултата рачунских огледа је коришћена дескриптивна статистика: минимум, максимум, средња вриједност, медијана, средње апсолутно одступање, и опсег. Размотрено је сљедеће:

1. стабилност приступа у рјешавању (алгоритма) процјењена на основу средњег апсолутног одступања и опсега;
2. квалитет рјешења добијених различитим алгоритмима упоређивањем и приказом расподјела поређења;
3. утицај параметара на профит, удио у тржишту и NPM;
4. сличност скупова разводних тачака коришћењем Жакардовог индекса сличности и Szymkiewicz-Simpson-овог коефицијента преклапања;
5. утицај auxiliary модела простим поређењем.

За обраду података су се користили програмски пакети NumPy, SciPy, и Pandas преинсталирани на Anaconda платформи.

Коришћене методе, како у теоријском, тако и у практичном дијелу у потпуности одговарају проблему и предмету истраживања. Како је ријеч о математичким резултатима у теоријском дијелу, коришћена методологија је сама по себи тачна и савремена. Што се тиче практичног дијела, Python спада у модерне и веома популарне динамички типизоване програмске језике са добро развијеним програмским библиотекама (пакетима) за научна израчунавања: NumPy, SciPy, Pandas, и Matplotlib. У самој дисертацији приказ поменутог је омогућен одговарајућим библотекама LaTeX-a (који је и коришћен за писање дисертације).

CAB скуп података је у научној литератури стандардан за формирање инстанци када је ријеч о проблемима размјештања разводних тачака.

Није дошло до промјене плана истраживања од пријаве теме дисертације.

Како је акценат на теоријском истраживању, јер је докторска дисертација из математике, сви разматрани параметри су довољни. Што се тиче практичног дијела, исто се може рећи и за развој алгоритама. Такође, статистичка обрада података се сматра адекватном.

1) Објаснити материјал који је обрађиван, критеријуме који су узети у обзир за избор материјала;

2) Дати кратак увид у примијењени метод истраживања при чему је важно оцијенити сљедеће:

1. Да ли су примијењене методе истраживања адекватне, довољно тачне и савремене, имајући у виду достигнућа на том пољу у свјетским нивоима;
2. Да ли је дошло до промјене у односу на план истраживања који је дат приликом пријаве докторске тезе, ако јесте зашто;
3. Да ли испитивани параметри дају довољно елемената или је требало испитивати још неке, за поуздано истраживање;
4. Да ли је статистичка обрада података адекватна.

VI РЕЗУЛТАТИ И НАУЧНИ ДОПРИНОС ИСТРАЖИВАЊА

Оригинални и најзначајни научни резултати истраживања

- уведен је нови, обогаћени проблем размјештања разводних тачака у Штаклберговом надметању при чему се разматра и надметање цијенама;
- формулисан је нелинеарни мјешовито-цјелобројни математички програм у два нивоа као математички модел за поменути проблем;

- доказано је постојање Бертранд-Нешовог цјеновног еквилибријума при логит-моделу конузмерског избора (теорема);
- доказано је постојање (сигурног) Штаклберговог еквилибријума (теорема);
- предложене су нове еквилибријумске једначине најбоље цјеновне реакције (став);
- испитано је гранично понашање оптималних вриједности функције циља leader-а и доказано је да је проблем follower-а *NP*-тежак (теореме);
- доказано је да изведени проблеми повезивања неразводних тачака са разводним припадају класи *P* (теореме и посљедице);
- доказано је да је проблем leader-а, тј. $(r|p)$ HCPuPW, *NP*-тежак (теорема);
- доказано је да су оптималне руте, за фиксиран размјештај разводних тачака, оне са најнижим транспортним трошковима (теореме);
- показано је како се проблем $(r|H_p)$ хаб-медианоида са цјеновним надметањем ($(r|H_p)$HMPuPW) може формулисати као мјешовито-цјелобројни линеарни програм;
- показано је како се допунски модел (АМ) може реформулисати као мјешовито-цјелобројни линеарни програм;
- показано је како се може имплементирати алтернацијска хеуристика (АХ) као метод за рјешавање проблема leader-а (алгоритам);
- показано је како се може имплементирати метод промјенљивих околина (МПО) за рјешавање проблема leader-а коришћењем АХ за добијање иницијалног рјешења (алгоритам);
- показано је како се може процијенити функција циља на основу линеарне релаксације модела follower-а;
- дата је теоријска основа за коришћење *k*-swap околина у модификацији МПО-а (ставови) за $(r|H_p)$HCPuPW;
- уведен је нови тип околина coreswap за претраживање МПО-ом.

Приказ добијених резултата

За све теоријске резултате су формулисане и *доказане* одговарајуће теореме, ставови (Propositions), и посљедице (Corollaries), како је то уобичајено у математичким наукама, чиме је и задовољен критеријум критичности.

Алгоритми су описани псеудокодом, што је уобичајено у математичким и информатичким међународним научним часописима, а за егзактно рјешавање доњег модела су приказани битни фрагменти рачунарског програмског кода на Python-у.

Приказ рачунских резултата је на адекватан и илустративан начин пропраћен одговарајућим графиконима, цртежима, и табелама. Графиконима су приказане неке функције ради појашњења доказа теорема, илустровани су сигурни Штаклбергов и Бертранд-Нешов цјеновни еквилибријум, те приказане CDF за Гаусову и Гамбелову расподјелу. Поред тога, у при предочењу резултата рачунских огледа коришћени су хистограм, barplot, violinplot, и boxplot. Табелама су приказане у литератури уобичајене карактеристике рачунских огледа на основу којих се процјењују пеформансе хеуристичких алгоритама. Резултати рачунских огледа су окачени на Харвардску DataVerse платформу за јавно дијељење података коришћених и добијених током научно-истраживачког рада. Кандидат је показао да је способан да прикупи, обради, и предочи резултате, те их упореди на одговарајући начин.

Нова сазнања

Нова сазнања се односе на добијене теоријске резултате исказане теоремама, ставовима, и њиховим посљедицама. Један нов проблем је уведен и за њега формулисан одговарајући нелинеарни мјештовито-цјелобројни математички програм у два нивоа. Формулисан и доказано је: 12 теорема, пет ставова, једна лема, и девет посљедица.

Такође, резултати рачунских експеримената говоре да је АХ добра за добијање иницијалног рјешења, али и да врло брзо упада у циклус (видно кратак). Поред тога, испоставило се да је *p*-HMLP боље полазно рјешење од *p*-HCLP за АХ. Разлика између МПО и АХ је у већини случајева мала, али зна бити значајна. Утицај параметра Θ је најуочљивији. Угрубо, leader може очекивати нижи профит за веће Θ, и обрнуто. Слично опажање вриједи и за параметар α. С друге стране, изгледа да Θ и α имају супротан ефекат на односе профита такмаца. Поред тога, није очекивано да leader-у буде профитабилно да заузима више од 50% тржишта, иако рачунски огледи сугеришу да leader има предност због „првог потеза“, што је, можда, и контраинтуитивно. Поред поменутог, резултати рачунских огледа сугеришу да у Штаклберговом надметању повећана осјетљивост на цијене имплицира неконзервативно понашање корисника услуга – више бивају везани за другог такмаца, тј. оног који улази на тржиште. Интересантно је да је ово опажање по својој природи више математичко-рачунска импликација, него што је социолошка, економска, или пак психолошка.

Нови истраживачки правци

Један смјер истраживања може да буде испитивање робусне варијанте овог проблема. Оно би захтјевало постављање новог модела, а тиме и другачији приступ у анализи и реформулацији. Поред тога, јединствени Бертранд-Нешов цјеновни еквилибријум указује да има смисла разматрати кооперативне цјеновне игре са раздјељивом добити јер је IATA својеврстан изузетак у међународним законима о заштити конкуренције. Такође, и за овај нови модел са кооперативним цијенама има смисла разматрати робусну варијанту.

Други од могућих смјерова будућих истраживања, а који се назире, би могао бити конструисање егзактног алгоритма за рјешавање *(r|p)*HCPuPW. Конструкција егзактног метода за нелинеарне оптимизационе проблеме у два нивоа у којима се мора дефинисати понашање follower-а се у литератури показала као изузетно тежак задатак. Штавише, из дисертације видимо да је и конструкција матхеуристичког алгоритма неуобичајена.

Поред развијања сложенијих модела и конструисања бољих приступа (оптималном) рјешењу, трећи правац истраживања који се назире јесте допунско испитивање сложености овог и сличних проблема, тј. тражење њиховог мјеста у полиномским и апроксимацијским хијерархијама.

1) Укратко навести резултате до којих је кандидат дошао;
2) Оцијенити да ли су добијени резултати јасно приказани, правилно, логично и јасно тумачени, упоређујући са резултатима других аутора и да ли је кандидат при томе испољавао довољно критичности;
3) Посебно је важно истаћи до којих нових сазнања се дошло у истраживању, који је њихов теоријски и практични допринос, као и који нови истраживачки задаци се на основу њих могу утврдити или назирати.

VII ЗАКЉУЧАК И ПРИЈЕДЛОГ

Докторска дисертација коју је написао студент III циклуса Димитрије Д. Чвокић, уводи и разматра нов проблем у области операционих истраживања, врло вјероватно најреалистичнији до сад кад је ријеч о размјештању разводних тачака у конкурентном окружењу. Кандидат је формулисао одговарајући математички оптимизациони модел у два нивоа, доказао постојање сигурног Штаклберговог еквилибријума, и конструисао алгоритме за њихово проналажење засноване на мета- и матхеуристикама. Научно истраживање, приказано у дисертацији, сасвим сигурно доприноси развоју теорије локације и матхеуристичких алгоритама. **Такође, оригиналност рукописа је провјерена и потврђена софтвером за откривање плагијаторства URKUND.** Докторска дисертација Димитрија Д. Чвокића задовољава све критеријуме оригиналног научног рада, самостално урађеног према општеприхваћеним важећим научно-истраживачким методима. На основу анализе рачунских огледа дошло се до одређених смјерница, тј. организационо-управљачких увида за формирање транспортних мрежа са разводним тачкама у конкурентном окружењу. Такође, дисертација се може сматрати довољно добрим полазиштем за нова истраживања.

С обзиром на претходно казано, Комисија је мишљења да је дисертација мр Димитрија Д. Чвокића квалитетнно урађена и да задовољава све критеријуме оригиналне тезе. Стога, Комисија предлаже Наставно-научном вијећу Природно-математичког Факултета и Сенату Универзитета у Бањој Луци да прихвате дисертацију и одобре њену јавну одбрану.

1) Навести најзначајније чињенице што тези даје научну вриједност, ако исте постоје дати позитивну вриједност самој тези;

2) На основу укупне оцјене дисертације комисија предлаже:
- да се докторска дисертација прихвати, а кандидату одобри одбрана,
- да се докторска дисертација враћа кандидату на дораду (да се допуни или измијени) или
- да се докторска дисертација одбија.

ПОТПИС ЧЛАНОВА КОМИСИЈЕ

Датум: 04.01.2021.

1. ....

Др Александар В. Пљасунов,
доцент на ужој научној области
Дискретна математика и теоријска кибернетика (ВАК РФ 01.01.09.),
Механичко-математички факултет
Националног истраживачког државног универзитета у Новосибирску,
Русија,
предсједник

2. Милан В. Јовановић

Др Милан В. Јовановић,
редовни професор на ужој научној области
Математичка анализа и примјене,
Природно-математички факултет Универзитета у Бањој Луци,

3. Драган Кораћ

Др Драган Кораћ,
доцент на ужој научној области Рачунарске науке,
Природно-математички факултет Универзитета у Бањој Луци,
члан

4. Сав Александар

Др Александар Љ. Савић,
ванредни професор на ужој научној области
Нумеричка математика и оптимизација,
Математички факултет Универзитета у Београду, Србија,
коментор, члан

5. Kochetov

Др Јуриј А. Кочетов,
редовни професор на ужој научној области
Дискретна математика и теоријска кибернетика (ВАК РФ 01.01.09.),
Механичко-математички факултет
Националног истраживачког државног универзитета у Новосибирску,
Русија,
ментор, члан

ИЗДВОЈЕНО МИШЉЕЊЕ: Члан комисије који не жели да потпише извјештај јер се не слаже са мишљењем већине чланова комисије, дужан је да унесе у извјештај образложење, односно разлог због којих не жели да потпише извјештај.

Изјава 1

Прилог 3.

ИЗЈАВА О АУТОРСТВУ

Изјављујем
да је докторска дисертација

Наслов рада ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ *(r|p)* ХАБ-ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ

Наслов рада на енглеском језику EXISTENCE AND SOLUTION METHODS FOR STACKELBERG EQUILIBRIA IN THE *(r|p)* HUB-CENTROID PROBLEM UNDER THE PRICE WAR

- ☒ резултат сопственог истраживачког рада,
- ☒ да докторска дисертација, у цјелини или у дијеловима, није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- ☒ да су резултати коректно наведени и
- ☒ да нисам кршио/ла ауторска права и користио интелектуалну својину других лица.

У Бањој Луци, дана 5. 2. 2021. године

Потпис докторанта
Димитрије Д. Чвокић

Изјава 2

**Изјава којом се овлашћује Универзитет у Бањој Луци
да докторску дисертацију учини јавно доступном**

Овлашћујем Универзитет у Бањој Луци да моју докторску дисертацију под насловом
ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ $(r|p)$ ХАБ-ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ

која је моје ауторско дјело, учини јавно доступном.

Докторску дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у дигитални репозиторијум Универзитета у Бањој Луци могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (*Creative Commons*) за коју сам се одлучио/ла.

- ◉ Ауторство
- ○ Ауторство – некомерцијално
- ○ Ауторство – некомерцијално – без прераде
- ◉ Ауторство – некомерцијално – дијелити под истим условима
- ○ Ауторство – без прераде
- ○ Ауторство – дијелити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на полеђини листа).

Потпис докторанта
Димитрије Д. Чвокић

У Бањој Луци, дана 5 . 2 .2021. године

Димитрије Д. Чвокић

Изјава 3

Изјава о идентичности штампане и електронске верзије докторске дисертације

Име и презиме аутора Димитрије Д. Чвокић

Наслов рада ПОСТОЈАЊЕ ШТАКЛБЕРГОВИХ ЕКВИЛИБРИЈУМА У ПРОБЛЕМУ *(r|p)* ХАБ-ЦЕНТРОИДА СА ЦЈЕНОВНИМ НАДМЕТАЊЕМ И АЛГОРИТМИ ЗА ЊИХОВО ПРОНАЛАЖЕЊЕ

Ментор проф. др Јуриј А. Кочетов

Изјављујем да је штампана верзија моје докторске дисертације идентична електронској верзији коју сам предао/ла за дигитални репозиторијум Универзитета у Бањој Луци.

У Бањој Луци, дана 5. 2.2021. године

Потпис докторанта
Димитрије Д. Чвокић