Public Transport Network Design for Competitive Service Providers

CHAN, Yon Sim Eddie

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Master of Philosophy
in
Systems Engineering & Engineering Management

November 2008

<u>Thesis/Assessment Committee</u>

Professor CHENG, Chun Hung (Chair)

Professor LEUNG, Janny M.Y. (Thesis Supervisor)

Professor LAM, Wai (Committee Member)

Professor LAM, Kokin (External Examiner)

Abstract of thesis entitled:

Public Transport Network Design for Competitive Service Providers

Submitted by CHAN, Yon Sim Eddie

for the degree of Master of Philosophy

at The Chinese University of Hong Kong in November 2008

We examine the game-theoretic equilibria within a setting of several profit maximising public-transit operators servicing multiple routes in various forms of road networks. Service providers maximise their individual profit by determining the set of routes to service while satisfying the capacity constraints (maximum number of routes it can offer service). Through numerical simulation using a game-theoretic model, we investigate the competition among the service providers under different infrastructural frameworks.

摘录

我們探討在各種不同形式的道路網絡及設定下，幾個服務多條路線並尋求最大利潤的公共交通服務營辦商之間的（納什）博弈均衡。在不違反容量限制（服務路線數量的頂點）之下，各服務商將個別選擇不同路線的集合，以優化其盈利。通過博弈論的方法，我們做了一些數值模擬，以便調查服務商間在不同基礎設施框架下的競爭。

Acknowledgement

I am indebted to my thesis committee members who provided me with sagacious advice throughout the revision of this thesis. They are Professor Cheng Chun Hung, Professor Janny Leung, Professor Lam Kokin and Professor Lam Wai.

I would like to take this opportunity to express my heartfelt gratitude to my thesis supervisor and mentor, Professor Janny Leung, for the kindness, encouragement and tireless guidance all the way through. A special thank you to Professor Ada Ng and Professor Leong Choon Heng for trusting me since the beginning of my study.

I am also very grateful to my fellow friends for their relentless support. Thank you, Alan Au, Alex Kam, Anthony Ng, Bao

Zhenhua, Cathy Wong, Chen Zhiyuan, Cui Xiangyu, Dickson Wong, Dorothy Wong, Eliphas Wong, Forest Kwan, Gao Jianjun, He Simai, Henry Hui, Ho Sin Cheong, Chen Jian, Jacky Wong, Katherine Chen, Kelvin Chan, Lin Yizhong, Lu Meng, Ma Lijun, Ong Hean Kuan, Pan Qi, Paul Fung, Qiu Jianchuang, Samson Tang, Sampson Chiu, Sharon Wong, Solomon Ding, Tao Zhijie, Wan Xiangwei, Wang Haifeng, Wang Chongyu, Xue Weili, Yang Di, Yang Yi, Yu Yikun, Zhang Feng, Zhang Nianbing, Zhou Ying, Zhu Hao and many others, which I cannot list all in here.

Last but not least, I would like to thanks Miss Ada So, Miss Iris Leung, Miss Mandy Tsoi, Miss Monica Ho, Miss Lolita Chan, Miss Phyllis Leung, Miss Tammy Yam, Miss Tracy Leung, Miss Yau Choi Oi, Mr. Edmund Cheng, Mr. Lam Tak Ming, Mr. Raymond Pang and Mr. Sam Wong for their kind assistance throughout my stay in the Chinese University of Hong Kong.

To the memory of my grandparents

and

To my beloved parents for their endless support

Contents

Abstract **i**

Acknowledgement **iii**

1 Introduction **1**

1.1 Motivation 1

1.2 Background 2

1.3 Literature Review 5

2 Game Theoretic Models For Competing Operators **10**

2.1 Competitive Equilibrium Model 12

2.1.1 Base Model Formulation 12

2.1.2 Capacitated Model Formulation 16

2.1.3 Solution Methods 19

2.2 Net Profit Maximizing 26

2.2.1 Equitable Route Assignment 27
2.3 Congestion Game Model with Player- and Route-dependent Operating Cost 31
2.3.1 Best-Response Algorithm 34
2.3.2 Integer Programming Formulation 41
2.3.3 Net Profit Maximizing 43

3 Network Design **45**
3.1 Network Structure 47
3.2 Comparison Between Two Network Structures . . 49
3.2.1 Routes with Same Ridership 49
3.2.2 Routes with Different Ridership 65
3.2.3 Network with Player- and Route-specific Profit Function 69

4 Elastic Demand **71**
4.1 Congestion Game Model with Service-Quality-Based Elastic Demand 71
4.1.1 Network with Service-Quality-Based Elastic Demand 82

5 Conclusion **84**

5.1 Future Work 84

5.1.1 Impact of Network Design and structure . 84

5.1.2 Non-cooperative and Cooperative Games . 86

5.1.3 Joint game-theoretic model of both passenger and providers 87

Bibliography **89**

List of Figures

2.1 Typical Solution from $(G2)$, $b_i \geq m$ 25

3.1 Complete point-to-point direct services vs. Hub-and-spoke with central interchange 47

3.2 Four Nodes Networks 49

3.3 $\frac{a}{\delta}$ versus $\pi_H - \pi_C$ 51

3.4 Plots with several other value of t 52

3.5 $log(\frac{a}{\delta})$ versus $\pi_H - \pi_C$ 53

3.6 Close up look for $\pi_H - \pi_C$ versus $\frac{a}{\delta}$ 62

3.7 Cases with different ridership on each route . . . 66

3.8 Approximate Range When Hub-and-spoke Network is Superior 66

3.9 Plots of $\frac{\sum a_j}{t \sum \delta_i}$ vs. $\pi_H - \pi_C$ — Player Specific Payoff Function 70

4.1 Case Example 74

4.2 Case Example's Solution 75

4.3 Plots of $log(\frac{\sum a_j}{\delta})$ vs. $\pi_H - \pi_C$ — Service-Quality-Based Elastic Demand 83

5.1 Other Prototypical Network Structures 85

List of Tables

3.1 $\frac{a}{\delta}$ Range for a Network where $n = t = 4$ 63

3.2 $\frac{a}{\delta}$ Range for a Network where $n = 4, t = 5$ 64

Chapter 1

Introduction

1.1 Motivation

Public Transport Issues in Malaysia

In Malaysia, the authority in transport policy and regulation, infrastructure, and operator licenses are distributed among three ministries: Ministry of Transport, Ministry of Works, and Ministry of Entrepreneurial and Cooperative Development. Lack of coordination among ministries result in poor transportation planning.

A long-distance bus service, 'Express Bus' is a popular means of inter-city travel in West Malaysia. Other than some safety issues — speeding — the system runs under a sustainable and profitable manner. On the other hand, town buses are underutilised and require government subsidy. Specialized services, such as school and factory buses, further affect the profitability. In Penang, poor public-bus operations (in terms of infrequent service, comfort level, slowness and accessability) make public transport a last-resort choice for the travelling public. For almost two-thirds of travellers, the preferred mode choice is 'private car'. This is primarily due to poor public transport; cheap petrol price, and an affordable national car (with about 60 percent market share) also exacerbate the decline in the use of public transport. Owning a motorcycle not only ease the user from all the aforementioned concerns, the corresponding long-term average-cost is also not much higher compared to taking public transport.

1.2 Background

In recent decades, a big portion of the population and economic growth was markedly due to metropolitan areas. In the United States, the 40 metropolitan areas (each with at least one million population in 1990) accounted for 53.4 percent of the nation's population[10]. The Bureau of Economic Analysis of the U.S. Department of Commerce stated that, "metropolitan areas produced 90 percent of U.S. current-dollar GDP; the five largest metropolitan areas accounted for 23 percent of the U.S. total" in 2005. As a result of the increasing population and economic activities, commuting time as well as travel distance for the metropolitan populace have also significantly increased. For instance, the net commuting into the 35 major U.S. metropolitan areas increased from less than 300,000 in 1980 to nearly 800,000 in 1990[10]. The average work trip distance in 2001 in the London boroughs was approximately 10 kilometres[2] and 13 kilometres for the new towns outside London's Greenbelt[3]. This indicates that a heavier burden is now being placed on the transport systems in metropolitan areas, leading to increased traffic congestion and attendant safety and environmental concerns.

In most places, development of transport infrastructure and public transit services failed to keep apace with the swell and sprawl of metropolitan areas. A typical public transit run by government authorities has the following difficulties. With limited budgets, services and connectivity are severely restricted. It is very common to have serious congestion in central business districts in contrast to insufficient coverage in peripheral areas. For example in Seattle, the bus service from the city centre to the major university suburb (that is only 20 minutes' drive away) runs only every half an hour and necessitates an intermediate bus transfer. In metropolises where public-transit services are provided by private firms in a relatively free market, operators tend to focus on high-profit routes and outlying smaller communities are usually under-served or ignored. Similar scenario can be seen in Hong Kong. The already congested Central business district is often jammed with half-empty double-decker buses from all the bus operators, while bus services to satellite communities in the New Territories are very infrequent and expensive.

This thesis investigates the competitive situation when several service providers offer public transit services, and the impact on the total set of services offered to the public and the resultant level of ridership of the system. We will investigate the competition among the service operators under different infrastructural frameworks by using a game-theoretic approach. The interplay between the basket of services offered and the overall ridership of the system will also be investigated.

With the modelling and analysis done here, government authorities may find it useful when tendering for public transit services. While providing insights and guidance on the number and types of routes (and their possible bundling) being offered for bidding, the result also helps in the decision regarding the provision of facilities or locations for transportation interchanges and hubs.

1.3 Literature Review

Although not explicitly acknowledged, concepts of game theory have been pervasively used in traffic studies. As pointed out by Fisk (1984) , the famous Wardrop's (1952) user-equilibrium principle is essentially the condition for a Nash (1950) game-theoretic equilibrium among road-users, since no driver can reduce his/her travel time by switching to a different route choice. Wardrop's principle has been a cornerstone in road traffic research for decades. Hollander and Prashker (2006) give an excellent review of recent literature on non-cooperative games in transport research.

Somewhat surprisingly, there has not been studies on the competitive situation amongst public transit operators in the literature. Castelli et al. (2004) modelled a game between two authorities (one determining flow, and the other capacities) in a freight transport network. This is obviously different to a passenger transit network as the transportee (freight) do not choose its route, a passenger does. Martin and Roman (2003) studied a game among airlines related to hub locations for each airline. According to Hollander and Prashker (2006), the "small number of such games is surprising, considering that NCGT (non-cooperative game theory) seems a natural tool for analysing relations between authorities"[11].

In this thesis, our models revolve around a type of game called *potential* games, but first we need to know what congestion games are. Congestion games were introduced by Rosenthal [16], and later generalised by Moderer and Shapley [14] to the class of *potential* game. Congestion games are non-cooperative games where the players' utilities (payoff functions that are *iden-*

tical for every player) depends on the choices of all players (to be exact, the total number of player choosing the same resource). Every congestion game admits a *potential* function and has at least one pure Nash equilibrium. A Nash equilibrium is a selection of strategies for all players such that no one player can unilaterally improve its payoff by switching to a different strategy. Potential games are games where the equilibrium can be computed by solving an auxiliary mathematical programme with the potential function as the objective function.

Milchtaich (1996) studied a class of congestion games with *player-specific* payoff functions (as contrast to the universally identical payoff function in traditional standard congestion games), with two assumptions. First, each player's strategy involves choosing only one resource (which in our case is the route selection) and second, the payoff received actually decreases (not necessarily so) with the number of other players selecting the same resource (monotonic payoff functions). Congestion games where each player chooses only one resource are sometimes called singleton/simple congestion games. As pointed out by Milchtaich, "these congestion games, while not generally admitting a *potential*, nevertheless always possess a Nash equilibrium in pure strategies" [13]. Similar to Milchtaich's model, we consider a model with player- and route-specific payoff function. The major different is the non-singleton congestion games in our setting. Players can choose more than one resource, and may be restricted by capacity constraint, if any.

On the other hand, Ieong et al. (2005) study a more generalized singleton congestion games. No simplifying assumption on the payoff/cost functions was made but they share the same restriction as of Milchtaich's first assumption. It was observed that Nash equilibrium solution in a singleton congestion games can serve as the Nash equilibrium solution of the subgame simply by excluding routes that are no longer in the complete-route-set of the subgame from the strategy sets. However, to construct a Nash equilibrium from two smaller equilibria (subgames), it must satisfy some conditions. With this, a better-response dynamic algorithm [12] was proposed. They showed that the algorithm finds the optimal Nash equilibrium for singleton games in polynomial time. Although we cannot construct algorithm for non-singleton congestion games which can be solved by breaking it into subgame, we proposed a best-response algorithm for finding the equilibrium.

Recently, there are many other works which study congestion games. They add different interpretations and perspectives to congestion games. Fabrikant et al.(2004) and Voecking (2006) proves that finding a Nash equilibrium in general congestion games is PLS-complete [1] [8], [17]. Chakrabarty et al. (2005) were the first to study centralized solutions for congestion games where the objective is to minimize the total cost[7]. Blumrosen and Dobzinski (2007), on the other hand aim to maximize the social welfare under a "centrally controlled" game. They also relate congestion games to combinatorial auctions and proposed an algorithm based on the useful connection between congestion games and combinatorial auctions [5].

□ **End of chapter.**

[1]Polynomial Local Search (PLS) problem was defined by Johnson et al. (1988) as an abstract class of local optimization problem.

Chapter 2

Game Theoretic Models For Competing Operators

Some preliminary investigation about the strategic gaming situation among competing public transit service providers have been carried out. In the first-cut model, we assume that all the operators have the same cost and price structure, and that the total ridership between each origin-destination pair is equally divided among all the operators that service that particular route. In this setting, a player of the game is the service provider, and its strategy is the set of routes that it chooses to offer service. Each player tries to maximize its total profit, and a Nash equilibrium is achieved when no player can improve its profit by unilaterally changing the set of routes it services.

We can show that this game can be modelled as what is known as a *congestion game* (first introduced by Rosenthal, 1973) where the equilibrium can be computed by solving an auxiliary mathematical programme. The original objective function is a non-linear function, we proposed an equivalent mixed-integer programming formulation to the problem.

A model that maximize the net profit was considered next, follow by a model that gives the most equitable route assignment given the maximized net profit. We still assume that the overall ridership will not be affected by the total number of route being offered services nor the total number of operator offering service on a particular route; and the assumption that the payoff functions are universal (identical for each player). Under these assumptions, the net profit maximizing model (centrally controlled) finds solution where at most one operator offer service per route while the competitive equilibrium model finds a solution that assigns as many operators as the market can bear to each route.

We further extend the model to incorporate player- and route-dependent operating cost. A best-response algorithm and greedy approach was introduced and compared with the equivalent potential game. Whilst the optimisation model (potential game) finds a Nash equilibrium, it may finds one that has low overall revenue, whereas the greedy approach finds the Nash equilibrium solution with highest total profit (optimal welfare).

2.1 Competitive Equilibrium Model

2.1.1 Base Model Formulation

We consider a game with n players (the service providers or operators) and m possible routes (origin-destination pairs). Throughout the discussion in this thesis, we will use route and origin-destination pair interchangeably). Let 2^M be the set of all subsets of the routes. Player i's strategy consist of a subset of routes, $S_i \subseteq M$. For each route j, let k_j denote the number of players choosing to offer service on the route, that is, $k_j = |\{i : j \in S_i\}|$. Note that k_j depends on the collection of strategies $\{S_1, S_2, \ldots, S_n\}$ but with an abuse of notation, we will omit the denotation for simplicity. Each player offering service on route j earns an identical payoff of $p_j(k_j)$ which depends on the number of players serving that route, namely,

$$p_j(k_j) = \begin{cases} \frac{a_j}{k_j} - \delta & , \text{ if } k_j > 0 \\ 0 & , \text{ if } k_j = 0 \end{cases},$$

where a_j is the total ridership of origin-destination pair j, and δ is the fixed operating cost of route j. Each player tries to maximize its total profit $\pi_i(S_1, S_2, \ldots, S_n) = \sum_{j \in S_i} p_j(k_j)$. A pure-strategy Nash equilibrium is a set of strategies $\{S_1^*, S_2^*, \ldots, S_n^*\}$ such that for each player i,

$$\pi_i(S_1^*, \ldots, S_{i-1}^*, S_i^*, S_{i+1}^*, \ldots, S_n^*) \geq \pi_i(S_1^*, \ldots, S_{i-1}^*, S_i, S_{i+1}^*, \ldots, S_n^*),$$

$\forall S_i \in 2^M$.

Following the framework of Rosenthal (1973), we can show that a pure-strategy Nash equilibrium can be obtained by solving the following auxiliary mathematical programme:

$$(G0): \quad Maximize \sum_{j=1}^{m} \sum_{y=1}^{k_j} p_j(y) \qquad (0.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = k_j, \ \forall j = 1, \ldots, m; \tag{0.1}$$

$$x_i^j \in \{0, 1\}, \forall i = 1, \ldots, n; \forall j = 1, \ldots, m. \tag{0.2}$$

where x_i^j indicates whether route j is in the set S_i of routes offered by player i according to its strategy S_i; that is, $x_i^j = 1$ if $j \in S_i$, zero otherwise. We note that this is not a straightforward binary linear programme, since the k_j values are also variables, so the objective is *not* a linear function. From the potential games point of view, the objective function of the auxiliary mathematical programme (0.0) is the potential function.

Pure-Strategy Nash Equilibria

Since solution to $(G0)$ exist, it suffices to show that any solution to $(G0)$ gives rise to a pure-strategy equilibrium. Let $\{x_i^{*j}, k_j^*\}, x_i^{*j} \in S_i^*$, solve $(G0)$, and suppose the associated strategy combination is not an equilibrium. Then for some l (player), there is a strategy $\hat{S}_l$ such that

$$\sum_{\substack{j \in \hat{S}_l \\ j \notin S_l^*}} p_j\left(k_j^* + 1\right) > \sum_{\substack{j \in S_l^* \\ j \notin \hat{S}_l}} p_j\left(k_j^*\right)$$

where S_l^* is the strategy used by l as indicated by the values of $\{x_i^{*j}\}$. Consider the new values $\{\hat{x}_i^j, \hat{k_j}\}$ associated with player l changing to the pure-strategy $\hat{S}_l$ (with the rest of the players not changing their strategies). (0.0) evaluated at $\hat{k_j}$ is:

$$\begin{aligned}
\sum_{j=1}^{m} \sum_{y=1}^{\hat{k_j}} p_j(y) &= \sum_{j=1}^{m} \sum_{y=1}^{k_j^*} p_j(y) + \underbrace{\sum_{\substack{j \in \hat{S}_l \\ j \notin S_l^*}} p_j(k_j^* + 1) - \sum_{\substack{j \in S_l^* \\ j \notin \hat{S}_l}} p_j(k_j^*)}_{>0} \\
&> \sum_{j=1}^{m} \sum_{y=1}^{k_j^*} p_j(y)
\end{aligned}$$

which contradict with the optimality of $\{x_i^{*j}, k_j^*\}$. ||

Intuitively, we can also argue why the solution of $(G0)$ give rise to a pure-strategy Nash equilibrium. At any point, if there exist an origin-destination pair j which provide an opportunity to a player (not yet offering service on this origin-destination pair) to increase its profit, origin-destination pair j will certainly be included in its strategy set. With each player trying to maximize their profit function defined as $\sum_{j=1}^{m} p_j(k_j)$, Nash equilibrium is guaranteed when each $p_j(k_j^*)$ took the minimum value that is greater than zero, leading to the maximum value of expression (0.0). No additional player will offer service on the origin-destination pair when $p_j(k_j^* + 1) < 0$.

2.1.2 Capacitated Model Formulation

In our base model, we assume that operators are not limited by the number of routes that they can serve (thus, an operator will serve a route as long as it is profitable). Due to capital limitation in equipment purchase, infrastructural investments, skilled labour availability, etc., it would be more realistic to assume that there is a capacity limit on the number of routes a service provider can offer. Let b_i define the maximum number of origin-destination pairs that player i can afford to provide service. The capacity restriction can then be modelled by adding the capacity constraints, $\sum_{j=1}^{m} x_i^j \leq b_i$, to $(G0)$, as follows:

$$(G1): \quad Maximize \quad \sum_{j=1}^{m} \sum_{y=1}^{k_j} p_j(y) \qquad (1.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = k_j, \ \forall j \in M; \tag{1.1}$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0, 1\}, \ i \in N; \ j \in M. \tag{1.3}$$

where $N = \{1, 2, ..., n\}$ and $M = \{1, 2, ..., m\}$ are sets of players and routes respectively.

For cases where operators have ample capacity, constraint (1.2) is automatically satisfied. Once the constraint (1.2) cannot be omitted (i.e. $b_i < m$), the problem become capacitated. With this,

1. Would the game amongst the operators still be a potential game? That is, would the optimal solution to the extended model ($G1$) still yield the Nash equilibrium?

2. How would the capacity limit alter the equilibrium solution?

3. How would the selection of services offered to the public be affected?

Introducing the capacity constraints may change the k_j value and hence the optimal solution. It is natural to expect a change in the potential function (objective function of the auxiliary mathematical program) or in worst case scenario, the game may not be a potential game when the capacity constraint is introduced. To answer these questions, we now proceed to argue that the previous proof still remain valid. Consider the case where there are still some profitable route, say route c, not selected by any player at the optimal point (due to the capacity constraints). The reason route c is not included in the optimal solution is that the least profitable route (among the chosen routes) provide greater gain than route c. In other word, it will only reduce the value of the objective function if we substitute route c with any selected route of the optimal solution. This is in accordance to the players' aim to maximize their own profit, it is not beneficial to include route c into their service. No player can do any better by unilaterally changing his strategy (set of routes to provide service) at the optimal point, which means equilibrium achieved. Note that this particular equilibrium obtained by solving the auxiliary mathematical programme does not guarantee maximal gain for each player and the total sum of all players' profit. It is only the point that no player will deviate from their strategies.

It may not be profitable for an operator to serve some routes, even without the capacity constraint. Due to social responsibility, the authority may require certain origin-destination pairs to be served although it is not profitable. However, for simplicity, we assume operators will only service origin-destination pairs which give positive payoff throughout our work. The model can be easily extended to represent a much practical case simply by setting a cut off point, say f_j, to each origin-destination pair. That is, operators will offer services only if $p_j(k) > f_j$. Alternatively, we can also consider to shift the payoff function so that the cut off point become negative, which may be regarded as a fixed amount of government subsidy.

2.1.3 Solution Methods

Solution Approach for $(G0)$

Assuming there is no restriction on how many origin-destination pair each operator can serve, $b_i \geq m$ for all i; which means operators can choose to serve zero to m origin-destination pairs depending on their own strategy. Without lost of generality, assume each ridership contribute to one dollar in revenue (the travel fare). As we have mentioned earlier, operators will offer services only on profitable routes, and they will continue to "enter the market" for a route as long as it is profitable. With this, we can set the value of k_j^* before solving the problem,

$$k_j^* = \max\left\{0, max\left\{h \in \mathbb{Z}_+ : \frac{a_j}{h} - \delta > 0, h > 0\right\}\right\}$$

Let x_i^j denotes the decision variables that indicates whether route j is in the set S_i of routes offered by player i according to its strategy S_i; $x_i^j = 1$ if $j \in S_i$, zero otherwise. Since we already know the value of k_j^*, thus as long as the number of player choosing each origin-destination pair j equals to k_j^*, the solution to $(G0)$is automatically an equilibrium. We will discuss later in Section 3.1 two different structure of the network and compare them according to the total profit, $\sum_{j=1}^{m}\left[p_j(k_j^*)\sum_{i=1}^{n} x_i^{*j}\right]$. Note that the total profit is equivalent to $\sum_{j=1}^{m} k_j^* p_j(k_j^*)$, which in fact can be calculated once $\{k_j^*\}$ is set. The steps of solving the problem are:

step 1: determine the value of k_j^*

step 2: evaluate all $p_j(k_j^*)$

step 3: assign the routes to the operators according to the determined k value and calculate the total profit.

Thus, $(G0)$ can be solved by following these few simple steps.

Solution Approach for $(G1)$

To solve $(G2)$, an equivalent mixed-integer programming formulation for $(G1)$ is presented below. This formulation will allow us to analytically (or computationally) compare the Nash-equilibrium solution to one that is "centrally" controlled and maximizes the total operator net profit. where function. A further variation of the model would give a solution that is "fair" in the sense of minimizing the profit difference among operators, after we obtained an optimal total net-revenue (an upper bound of the total net profit of the network).

An equivalent formulation for $(G1)$ is as follows:

Let variable

$$y_{jk} = \begin{cases} 1 & \text{if } k_j = k \\ 0 & \text{otherwise} \end{cases}$$

Also we define:

$$P_j(k) = \begin{cases} \sum_{z=1}^{k} p_j(z) \text{ for } k > 0 \\ 0 \quad \text{if } k = 0 \end{cases}$$

where

$$p_j(z) = \frac{a_j}{z} - \delta$$

$(G1)$ is equivalent to:

$$(G2): \quad Maximize \sum_{j=1}^{m} \sum_{k=0}^{n} P_j(k)\, y_{jk} \tag{2.0}$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \ \forall j \in M; \tag{2.1}$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0,1\}\,, \ i \in N; \ j \in M; \tag{1.3}$$

$$\sum_{k=0}^{n} y_{jk} = 1, \forall j \in M; \tag{2.4}$$

$$y_{jk} \in \{0,1\}\,, \ k \in N \cup \{0\}; \ j \in M. \tag{2.5}$$

Constraint (2.4) would specify a unique "k_j" for each j. This ensure the correct RHS for (2.1) and for the objective function.

If $b_i \geq m$, for all i, then we can see that:

Observation 1:

Since $P_j(k)$ is maximized at k_j^* where

$$k_j^* = \max\{0, max\{h \in \mathbb{Z}_+ : p_j(h) > 0\}\}$$

$(G2)$ is optimized when operators continue to "enter the market" for route j until there are k_j^* operators offering service. The intuition behind it is as follows. With each player trying to maximize their profit function defined as $\sum_{j=1}^{m} p_j(k_j)$, a player will include any profitable route in its strategy whenever possible. That is, when the player has not reached its capacity limit and there are still origin-destination pair/s (not yet included in its strategy) which provide an opportunity to increase its profit. No additional player will serve the origin-destination pair where $p_j(k_j^* + 1) \leq 0$. Hence the equilibrium is reached when each $p_j(k_j^*)$ attains the minimum positive value, which will also lead to the maximum value of objective function (2.0).

Given a set of values of a_j's and δ, we can first calculate all the possible $p_j(k)$ values, where $k = 1, 2, ..., n$. Using $(G2)$, we can then get the solution as depicted in Figure 2.1. Note that k_j^*'s are at their maximum value where $p_j(k_j^*)$ greater than zero. Also, $\sum_{j=1}^{m} k_j^* p_j(k_j^*)$ can be calculated once we obtained k_j^*'s, as mentioned earlier in the first approach.

a_1	a_2	a_3	a_4	a_5	a_6
466	650	693	250	59	507

δ
219

k_j	$p_1(k_1)$	$p_2(k_2)$	$p_3(k_3)$	$p_4(k_4)$	$p_5(k_5)$	$p_6(k_6)$
1	247.00	431.00	474.00	31.00	-160.00	288.00
2	14.00	106.00	127.50	-94.00	-189.50	34.50
3	-63.67	-2.33	12.00	-135.67	-199.33	-50.00
4	-102.50	-56.50	-45.75	-156.50	-204.25	-92.25

k_1^*	k_2^*	k_3^*	k_4^*	k_5^*	k_6^*
2	2	3	1	0	2

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	
Player 1	1	1	1	1	0	1	
Player 2	1	1	1	0	0	1	
Player 3	0	0	1	0	0	0	
Player 4	0	0	0	0	0	0	$\sum_{j=1}^{m} \sum_{y=1}^{k_j^*} p_j(y)$
$\sum_{y=1}^{k_j^*} p_j(y)$	261.00	537.00	613.50	31.00	0.00	322.50	1765
							$\sum_{j=1}^{m} k_j^* p_j(k_j^*)$
$k_j^* p_j(k_j^*)$	28.00	212.00	36.00	31.00	0.00	69.00	376

Figure 2.1: Typical Solution from $(G2)$, $b_i \geq m$

2.2 Net Profit Maximizing

We might want to study the potential maximum profit that can be sought from a particular network. Now consider the problem of maximizing the sum of net profit for all the operators:

$$(G3): \quad Maximize \sum_{j=1}^{m}\sum_{k=0}^{n} k\, p_j(k) y_{jk} \qquad (3.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \ \forall j \in M; \qquad (2.1)$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \qquad (1.2)$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \qquad (1.3)$$

$$\sum_{k=0}^{n} y_{jk} = 1, \forall j \in M; \qquad (2.4)$$

$$y_{jk} \in \{0,1\}, \ k \in N \cup \{0\}; \ j \in M. \qquad (2.5)$$

Observation 2:

Since $k\, p_j(k) = k\left(\frac{a_j}{k} - \delta\right) = a_j - \delta\, k$, for $k \neq 0$, the objective is optimized when $k_j = 1$ if $a_j > \delta$. Hence if $b_i \geq m$, for all i, the optimal solution is $y_{j1} = 1$ *or* $y_{j0} = 1$, for all j. That is, there is at most one operator for each route.

Observation 3:

For the uncapacitated case, the "competitive" equilibrium solution has all operators offering service on all routes (until $p_j(k) < 0$), whereas the "centrally controlled" solution has at most one operator per route!

2.2.1 Equitable Route Assignment

After maximizing the net profit, another problem arises as to how to assign the route to operators. The problem of finding a most "equitable" solution can be modelled as:

$$(G4): \quad Minimize \;\; Z_{max} - Z_{min} \tag{4.0}$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \;\; \forall j \in M; \tag{2.1}$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \;\; \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0,1\}\,, \;\; i \in N; \;\; j \in M; \tag{1.3}$$

$$\sum_{k=0}^{n} y_{jk} = 1, \;\; \forall j \in M; \tag{2.4}$$

$$y_{jk} \in \{0,1\}\,, \;\; k \in N \cup \{0\}; \;\; j \in M; \tag{2.5}$$

$$Z_{max} \geq \sum_{j=1}^{m} \left(\sum_{k=0}^{n} p_j(k)\, y_{jk} \right) x_i^j, \;\; \forall i \in N; \tag{4.6}$$

$$Z_{min} \leq \sum_{j=1}^{m} \left(\sum_{k=0}^{n} p_j(k)\, y_{jk} \right) x_i^j, \;\; \forall i \in N; \tag{4.7}$$

$$\sum_{j=1}^{m} \sum_{k=0}^{n} k\, p_j(k)\, y_{jk} \geq \sum_{j=1}^{m} \hat{k_j}\, p_j(\hat{k_j}). \tag{4.8}$$

where $\hat{k_j}$ is the derived optimal values from the optimal solution for $(G3)$; Z_{max} and Z_{min} are the profit of operator with the most and least total profit respectively.

Constraint (4.8) ensure that we do not get a trivial solution such as $x_i^j = 0$, for all i and j. Note that (4.6) and (4.7) are non-linear constraints!

These can be linearized by introducing binary variables w_{ijk}, intended to be:

$$w_{ijk} = \begin{cases} 1 & \text{if } x_i^j = 1 \text{ and } y_{jk} = 1 \\ 0 & \text{otherwise} \end{cases}$$

This can be enforced by the following constraints:

$$\begin{aligned} \sum_{k=0}^{n} w_{ijk} &\leq x_i^j, \ \forall i \in N; j \in M; \\ \sum_{i=1}^{n} w_{ijk} &\leq k\, y_{jk}, \ \forall k \in N \cup \{0\};\ j \in M; \\ x_i^j + y_{jk} &\leq w_{ijk} + 1, \ \forall i \in N;\ j \in M;\ k \in N \cup \{0\}. \end{aligned}$$

So a mixed-integer programming formulation of $(G4)$ is given by:

$$(G4') : \qquad Minimize\ \ Z_{max} - Z_{min} \tag{4.0}$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \ \forall j \in M; \tag{2.1}$$

$$\sum_{j=1}^{m} x_i^j \le b_i, \ \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \tag{1.3}$$

$$\sum_{k=0}^{n} y_{jk} = 1, \ \forall j \in M; \tag{2.4}$$

$$y_{jk} \in \{0,1\}, \ k \in N \cup \{0\}; \ j \in M; \tag{2.5}$$

$$Z_{max} \ge \sum_{j=1}^{m} \sum_{k=0}^{n} p_j(k)\, w_{ijk}, \ \forall i \in N; \tag{4.6'}$$

$$Z_{min} \le \sum_{j=1}^{m} \sum_{k=0}^{n} p_j(k)\, w_{ijk}, \ \forall i \in N; \tag{4.7'}$$

$$\sum_{j=1}^{m} \sum_{k=0}^{n} k\, p_j(k)\, y_{jk} \ge \sum_{j=1}^{m} \hat{k_j}\, p_j(\hat{k_j}); \tag{4.8}$$

$$\sum_{k=0}^{n} w_{ijk} \le x_i^j, \ i \in N; \ j \in M; \tag{4.9}$$

$$\sum_{i=1}^{n} w_{ijk} \le k\, y_{jk}, \ k \in N \cup \{0\}; \ j \in M; \tag{4.10}$$

$$x_i^j + y_{jk} \le w_{ijk} + 1, \ i \in N; \ j \in M; \ k \in N \cup \{0\}; \tag{4.11}$$

$$w_{ijk} \in \{0,1\}, \; i \in N; \; j \in M; \; k \in N \cup \{0\}. \qquad (4.12)$$

Note that when $b_i \geq m$, for all i, constraint (1.2) can be omitted as it will automatically satisfied.

Observation 4

Solving $(G4)$ require hours or even days when there is no fixed cost, $\delta = 0$.

2.3 Congestion Game Model with Player- and Route-dependent Operating Cost

In reality, each operator might have different operating costs on different routes. Hence we try to model a congestion game with player- and route-specific payoff functions. With a simple monotonic revenue function, we manage to extend our basic congestion model to a congestion game with player- and route-specific payoff functions which admit a *potential*. In contrast to Milchtaich's work [13], our model is not a singleton game; players can choose more than one route in their strategy sets.

We define $p(i,j,k_j) = \frac{a_j}{k_j} - \delta_{ij}$, a payoff function which depends on player (i), route (j) and the total number of player choosing that particular route (k_j). We extend the basic congestion model to the following:

There are n operators (players) competing for m routes, where the *revenue* to each player i serving route j is

$$r_j(k_j) = \frac{a_j}{k_j} \text{ for } k_j > 0$$

and k_j is the number of players offering services on route j.

We claim that a Nash equilibrium of this problem can be found by solving the following auxiliary problem:

$$(GD1): \quad Maximize \sum_{j=1}^{m} \sum_{y=1}^{k_j} r_j(y) - \sum_{j=1}^{m} \sum_{i=1}^{n} \delta_{ij}\, x_i^j \quad (1d.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = k_j, \ \forall j \in M; \quad (1.1)$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \quad (1.2)$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \quad (1.3)$$

The proof is a simple extension of the approach of Rosenthal (1973). Essentially, we show that (1d.0) is a potential function of the potential game.

Proof:

Let x^* be an optimal solution of $(GD1)$, and let $(S_1^*, S_2^*, ..., S_n^*)$ be the corresponding strategies of the players.
If this is not an equilibrium solution, then for some player l, there exist another strategy $\hat{S}_l$ such that

$$\text{payoff}_l(S_1^*, ..., S_{l-1}^*, \hat{S}_l, S_{l+1}^*, ..., S_n^*) > \text{payoff}_l(S_1^*, ..., S_{l-1}^*, S_l^*, S_{l+1}^*, ..., S_n^*),$$

that is:

$$\sum_{j \in \hat{S}_l \setminus S_l^*} \frac{a_j}{k_j^* + 1} - \sum_{j \in S_l^* \setminus \hat{S}_l} \frac{a_j}{k_j^*} - \sum_{j \in \hat{S}_l \setminus S_l^*} \delta_{lj} + \sum_{j \in S_l^* \setminus \hat{S}_l} \delta_{lj} > 0. \qquad (**)$$

Let $(\hat{x}, \hat{k})$ be the solution of $(GD1)$ corresponding to $(S_1^*, ..., S_{l-1}^*, \hat{S}_l, S_{l+1}^*, ..., S_n^*)$. Then

$$\sum_{j=1}^{m} \sum_{y=1}^{\hat{k}_j} r_j(y) - \sum_{j=1}^{m} \sum_{i=1}^{n} \delta_{ij}\, \hat{x}_i^j$$

$$
\begin{aligned}
&= \sum_{j=1}^{m}\sum_{y=1}^{k_j^*} r_j(y) - \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}\, x_i^{*j} + \\
&\left(\sum_{j\in \hat{S}_l \setminus S_l^*} \frac{a_j}{k_j^*+1} - \sum_{j\in S_l^* \setminus \hat{S}_l} \frac{a_j}{k_j^*} - \sum_{j=1}^{m} \delta_{lj}\, \hat{x}_i^j + \sum_{j=1}^{m} \delta_{lj}\, x_i^{*j} \right) \\
&= \sum_{j=1}^{m}\sum_{y=1}^{k_j^*} r_j(y) - \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}\, x_i^{*j} + \\
&\underbrace{\left(\sum_{j\in \hat{S}_l \setminus S_l^*} \left(\frac{a_j}{k_j^*+1} - \delta_{lj} \right) - \sum_{j\in S_l^* \setminus \hat{S}_l} \left(\frac{a_j}{k_j^*} - \delta_{lj} \right) \right)}_{>0 \text{ by } (**)} \\
&> \sum_{j=1}^{m}\sum_{y=1}^{k_j^*} r_j(y) - \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}\, x_i^{*j}
\end{aligned}
$$

So we have another solution $\hat{x}$ with a better objective value, contradicting the optimality of x^*.||

2.3.1 Best-Response Algorithm

Before proceeding to the correspond mixed-integer programming formulation, let us consider a best-response algorithm on solving this problem. For simple illustration, we will consider the case where the operating cost is operator(player) dependent only. A player- and route-dependent case can be extended easily. Priority will be given to operator with lower operating cost. The argument is that operator with better operating efficiency will be able to offer a more competitive charge (not necessary so, or perhaps provides better service that is beneficial to the passenger if the authority controls the fare charges and keeps it at certain prices) and can stay longer in competition. Hence, even operator may cheat by declaring a lower than actual cost, they may not be able to sustain its operation for long and will be replaced by other operator sooner or later. In contrast, we can do it the other way round — give opportunity for operator with higher operating cost to enter the competition first. They will be eliminated as time passes. But in that case, it is possible for operators with lower operating cost to end up servicing little or no route at all. Should we penalize them because of their high efficacy? For the welfare of the passenger and to keep a healthy competition, we therefore stick to the former way. The best-response steps are as follows:

1. First, rank the operator based on their operating efficiency

using their respective operating cost, δ_i — lower cost higher rank.

2. We can start at any point and let the best-respond dynamics lead the solution to an equilibrium. Here, we randomly assign the route to operators, then determine the correspond $\tilde{k}_j$'s as the initial point for the next step's iteration.
3. Then start the iteration with the following simple algorithm:

```
while (non-equilibrium)/
          (exist operator that can improve his profit) do
    for (last  to first ranked operator) do
         exclude non-profitable route from current set of service
    end for
    for (first  to last ranked operator) do
         include profitable route into current set of service
    end for
end while
Return Strategy set
```

Note that best-response dynamics simulate local search on the potential function; improving moves for players increase the value of the potential function. Even if operators exclude service to a route, it only exclude routes which are unprofitable (i.e. when $\delta_{ij} > r_j(k)$, which will result in the increment of the potential function). Since the algorithm will terminate whenever no player can unilaterally improves their payoff (Nash equilibrium achieved), it may not always maximizes the potential function. As it is possible to reach an equilibrium with higher overall cost, $\sum_{j=1}^{m}\sum_{i=1}^{n}\delta_{ij}x_i^j$. We will discuss in further detail later in the proof for **Theorem 1**. Best-response dynamics always converge to a Nash quilibrium, however.

Observation 5:

From 1000 cases each for network with 4, 5, 6, 7 and 8 nodes studied, it took at most four iterations in solving the problem. The best-response dynamics converge fairly quickly.

To look for a Nash equilibrium with higher overall total profit from all operators, consider the following greedy approach. Similar to the previous algorithm, the greedy algorithm is as follows:

1. First, rank the operator based on their operating efficiency using their respective operating cost, δ_i — lower cost higher rank.

2. Instead of start at a random point, we initiate with all $\tilde{k}_j$'s and x_i^j's equal zeroes.

3. Then start the iteration with the following simple algorithm:

```
while (non-equilibrium)/
         (exist operator that can improve his profit) do
   for (first  to last ranked operator) do
       include profitable route into current set of service
   end for
end while
Return Strategy set
```

Observation 6:

As before, operators will continue to "enter the market" until it is unprofitable to do so, assuming that the operators are not capacity constrained (i.e. $b_i \geq m$).
In this case the optimal value k_j^* is defined by

$$k_j^* = max\{k : r_j(k) - \delta_{[k]j} > 0\}$$

where $[\cdot]$ is a permutation of $\{1, 2, ..., n\}$ such that

$$\delta_{[1]j} \leq \delta_{[2]j} \leq ... \leq \delta_{[n]j}$$

Theorem 1:

The solutions obtained from the greedy algorithm are Nash equilibrium with the highest overall total profit from all operators.

Proof:

Before proving Theorem 1, we will first show that while the best-response algorithm always converge to a Nash equilibrium, it does not promises a high overall total profit from all operators. With that, it will reinforce the rationale behind Theorem 1.

Lets have a closer look at the best-response algorithm. By setting $k_j^* = max\{k : r_j(k) - \delta_{[k]j} > 0\}$, we know that route j can be serviced by at most k_j^* operator/s simultaneously without bearing any losses. Note that $\delta_{[k]j}$'s are sorted according to their respective operating efficiency — lowest to highest cost. Hence operator/s with cost greater than $\delta_{[k_j^*]j}$ who happened to servicing route j will leave route j once there are k_j^* operator/s in service, or when it is no longer profitable to service route j (whichever come first). Follow by the entering of operator with lower cost. It is also possible for operator/s with cost higher than $\delta_{[k_j^*]j}$ to stay in service when there are k_j^* operator/s servicing route j, i.e. when there exist an operator c where $\frac{a_j}{k_j^*} - \delta_{[c]j} > 0$, for $n \geq [c] > k_j^*$. In other words, that particular operator with cost higher than $\delta_{[k_j^*]j}$ (which already in service) would not leave as it is still profitable. However, it will be replaced by other operator/s if there exist another operator d where $\frac{a_j}{k_j^*+1} - \delta_{[d]j} > 0$, for $k_j^* \geq [d] \geq 1$. Note that the aforementioned scenarios would not occur in the greedy algorithm since operators are assigned according to its respective operating efficiency. No matter how, the solution will eventually converge to the equilibrium where $k_j^* = max\{k : r_j(k) - \delta_{[k]j} > 0\}$, with the overall cost, $\sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}x_i^j$, vary depending on the initial point of the best-response algorithm. This is true because as long as the total number of operators in the network is less than k_j^*, there will be at least one operator can unilaterally improves it profit. Now we have describe the process in the best-response algorithm, the proof for Theorem 1 is then very obvious. Also with $k_j^* = max\{k : r_j(k) - \delta_{[k]j} > 0\}$, the greedy algorithm provides Nash equilibrium with the highest overall total profit from all operators. This is due to the priority rule, which ensure it captures the least possible overall cost from all operators when there are k_j^* operator/s servicing route j, i.e. $\sum_{i=1}^{k_j^*} \delta_{[i]j}$. ||

2.3.2 Integer Programming Formulation

Similar to the basic model, $(GD1)$ can be formulated as an equivalent integer programming:

$$(GD2): \quad Maximize \; \sum_{j=1}^{m}\sum_{k=0}^{n} R_j(k)\, y_{jk} - \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}\, x_i^j \quad (2d.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \; \forall j \in M; \quad (2.1)$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \; \forall i \in N; \quad (1.2)$$

$$x_i^j \in \{0,1\}, \; i \in N; \; j \in M; \quad (1.3)$$

$$\sum_{k=0}^{n} y_{jk} = 1, \forall j \in M; \quad (2.4)$$

$$y_{jk} \in \{0,1\}, \; k \in N \cup \{0\}; \; j \in M. \quad (2.5)$$

where

$$R_j(k) = \sum_{y=1}^{k} r_j(k) = \frac{a_j}{1} + \frac{a_j}{2} + ... + \frac{a_j}{k}$$

Observation 7:

Whilst the optimisation model finds a Nash equilibrium, it may finds one that has low overall profit, whereas the greedy algorithm finds the highest-net-profit Nash equilibrium solution. The major different between the solutions is that $k_j^* = max\{k : r_j(k) - \delta_{[k]j} > 0\}$ for the greedy approach while $k_j^* = max\{k : r_j(k) - \delta_{[k]j} \geq 0\}$ for the game theoretical approach. And it happened only when there exist a route where the last operator joining the service is making zero profit on that particular route, consequently drag down all the involving operators' profits. Since by incorporating operator/s with zero profit into the solution will not violate the equilibrium and the objective function of the auxiliary mathematics programme for the game theoretical approach.

2.3.3 Net Profit Maximizing

Now consider the "centralized" model where the objective is to maximize "total" operator profit. This can be found by solving

$$(GD3): \quad Maximize \sum_{j=1}^{m}\sum_{k=0}^{n} k\, r_j(k)\, y_{jk} - \sum_{j=1}^{m}\sum_{i=1}^{n} \delta_{ij}\, x_i^j \quad (3d.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \ \forall j \in M; \quad (2.1)$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \quad (1.2)$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \quad (1.3)$$

$$\sum_{k=0}^{n} y_{jk} = 1, \forall j \in M; \quad (2.4)$$

$$y_{jk} \in \{0,1\}, \ k \in N \cup \{0\}; \ j \in M. \quad (2.5)$$

Observation 8:

Note that $k\, r_j(k) = \left(\frac{a_j}{k}\right) k = a_j$ for $k \neq 0$. So assuming all $\delta_{ij} > 0$, then clearly the optimal solution is to have each route j operated by the operator i with minimum δ_{ij} if $a_j > \delta_{ij}$ for that operator; and with no operator for that route if $a_j < \delta_{ij}$ for all i, assuming that the operators are not capacity constrained (i.e. $b_i \geq m$). To get the solution with maximal total net profit ("centrally" controlled), it is therefore equivalent to solve $(G3)$ using the top ranked operator's profit function.

So again we get the dichotomy of "at most one operator per route" for the "centrally controlled" case versus "as many operators as the market can bear" for the "competitive" case.

When $b_i < m$ for some operators, then the assignment of routes to operators is not so clear-cut and $(GD3)$ is a "real" optimization problem to be solved!

□ **End of chapter.**

Chapter 3

Network Design

Using the framework developed in the previous chapter, we will explore the impact of the network structure on the profit for the service providers. We consider a service area with t townships (nodes) and compared the equilibrium solution for a network structure where direct services are offered between every pair of townships to the equilibrium for a hub-and-spoke network, where every route between any two townships involves an interchange via a central hub. (See Figure 3.1) In the second network structure, the routes offered are between a township and the central hub, and the total ridership from each origin (to all destinations) is consolidated into the ridership from the origin to the central hub. For each service provider, the profit from a route depends on the operating cost of offering the service, the total profit due to the ridership and the number of competitors also servicing that route.

We examined numerically one thousand cases assuming the ridership and operating cost is the same for all origin-destination pairs, for various combinations of values for the ridership (a) and route operating cost (δ). Interestingly, the numerical results seems to suggest that the ratio $\frac{a}{\delta}$ is pivotal. For low and high values of the ratio (when operating costs are very low or very high relative to the ridership), the service providers (at a Nash equilibrium situation) make a *higher* profit with a hub-and-spoke network than with a complete direct-service network, and the cut-off values seem to be quite sharp. Analytical results are obtained that can clearly delineate when one network structure is favoured by the service providers over the other. We then investigate networks with different ridership among origin-destination pairs. The switch-over value of $\frac{\sum a_j}{\delta}$ between the two network structure are not as clear-cut. Computational experiments for cases where the payoff functions are player- and route-specific were also studied. Again, we observed a not very clear-cut switch-over value of $\frac{\sum a_j}{\sum \delta_i}$ between the two network structures.

3.1 Network Structure

As mentioned earlier, we examine two different structure of a network. The first one is a complete network with t nodes and $\sum_{u=1}^{t-1} u = \frac{t(t-1)}{2}$ links, one for each origin-destination pair. The second is a hub-and-spoke network with $t+1$ nodes and t links. Figure 3.1 illustrate these two networks with $t = 6$. For the second structure, we assume no loss of ridership due to the change in network structure, as explained below.

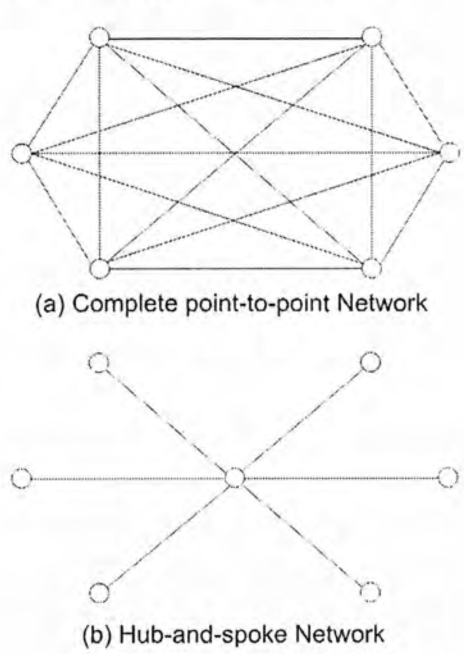


Figure 3.1: Complete point-to-point direct services vs. Hub-and-spoke with central interchange

From the original $\frac{t(t-1)}{2}$ origin-destination pairs in the first structure to t origin-destination pairs in the second structure, the ridership will be redistributed as illustrated in the following manner. Consider a four nodes network with respective ridership as depicted in Figure 3.2. When links AB, BC, CD, DA, AC, and BD is replaced by links AO, BO, CO, and DO, passenger who wish to travel from or to A must pass through link AO, hence we assign a ridership of $a_1 + a_4 + a_6$ for link AO. The ridership for BO, CO, and DO are obtained similarly. Since the hub-and-spoke network require each passenger to travel on two links to reach the destination, we assume that the fare for each link is cut to half a dollar.[1] As mentioned earlier, to compare the two structure, we assume that the system operate according to a competitive equilibrium but we measure the total profit generated from the network, which is, $\sum_{j=1}^{m} k_j\, p_j(k_j)$, to compare the two networks. When $b_i \geq m$, it is sufficient to solve $(G1)$ to obtain a competitive equilibrium solution.

[1]Other fare structure can be studied with this same model.

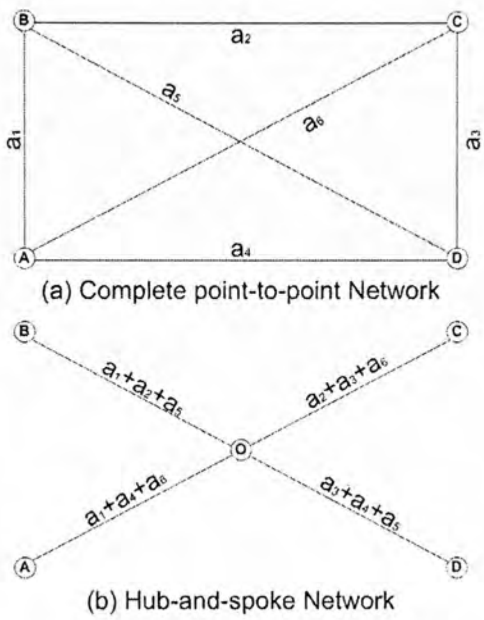


Figure 3.2: Four Nodes Networks

3.2 Comparison Between Two Network Structures

3.2.1 Routes with Same Ridership

We try to study a simple scenario first. Assuming a network with four players, and four nodes, where all the origin-destination pairs have the same ridership, $a_i = a$, $\forall i$, and operators are not capacity constrained, $b_i \geq m$, $\forall i$. One thousand cases with random ridership, a, and operating cost, δ, were simulated. Both a and δ are pseudo-random numbers uniformly distributed between zero and 1000 (generated using Excel's random function). We then compare the total profit of the two network structure under a competitive environment. Refer to Figure 3.3 for the plot $\frac{a}{\delta}$ versus $\pi_H - \pi_C$, where the y-axis refer to the revenue/cost ratio while the x-axis refer to the difference in total profit between the two network structure. Total profit for complete structure and hub-and-spoke structure are denoted by π_C and π_H respectively. Out of the 1,000 cases, 167 cases are irrelevant as the fixed operating cost is far greater than the total profit. From the remaining 833 cases, 68% of the cases perform better when hub-and-spoke structure is imposed. The improvement ranging from 0.22% to 10,933%. While about 11% of the cases show that the two structure perform equally well. The rest are cases where hub-and-spoke structure is inferior. One interesting observation is that hub-and-spoke structure is inferior when the $\frac{a}{\delta}$ ratio fall between clear cut ranges — $\left[\frac{4}{3}, 2\right)$ and $\left[\frac{8}{3}, 3\right)$.

Several other combination of different nodes (4,5,6,7,8) and players number (4,5,6) were investigated. Similar observation was found. Figure 3.4 shows the results for different number of nodes, t. While Figure 3.5 gives the plots when we take the logarithm value of the revenue-cost ratio.

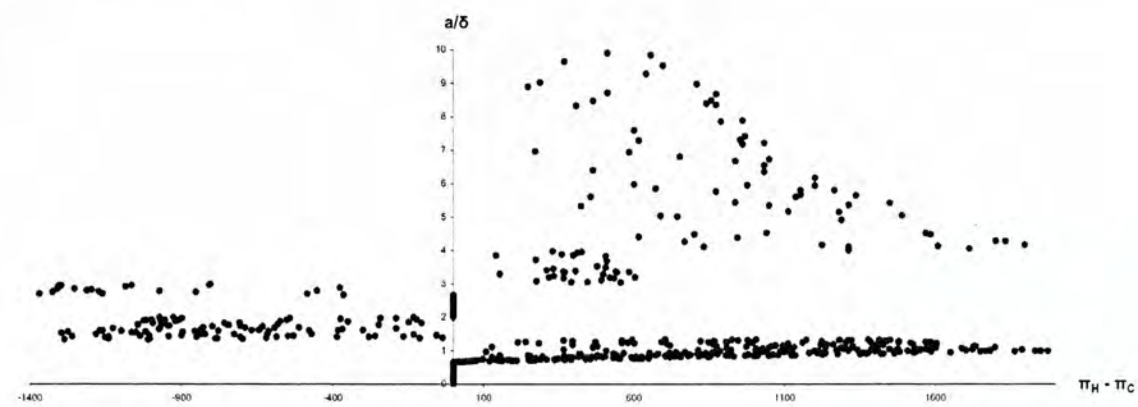


Figure 3.3: $\frac{a}{\delta}$ versus $\pi_H - \pi_C$

Analytical Result

Let π_C and π_H denote the total profit for complete network and hub-and-spoke network respectively. The total profit is obtained by summing up profit of each route from all operators. Since all routes have the same ridership and there will be only k operators offering services on each route, thus the total profit for the complete network is $\sum_{i=1}^{n}\sum_{j=1}^{m} x_i^j\, p_j(k) = \frac{t(t-1)}{2} k\, p_j(k)$, where

$$p_j(k) = \begin{cases} \frac{a}{k} - \delta & , \text{ for } k > 0 \\ 0 & , \text{ for } k = 0 \end{cases}.$$

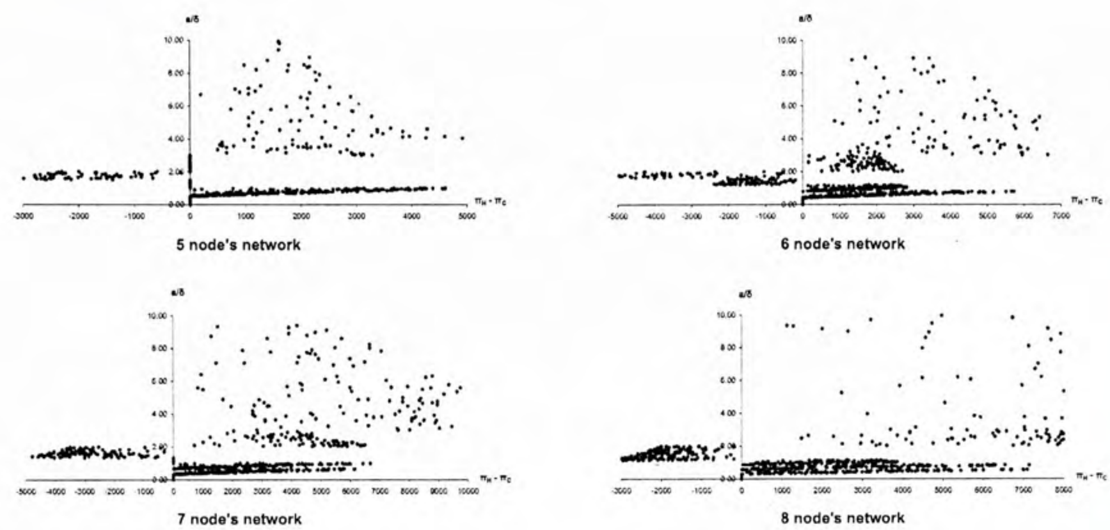


Figure 3.4: Plots with several other value of t

Therefore

$$\pi_C = \begin{cases} \frac{t(t-1)}{2}\,(a - \delta\, k) & , \text{ for } k > 0 \\ 0 & , \text{ for } k = 0 \end{cases}$$

where $k = max\left\{h : \frac{a}{h} > \delta, 0 \leq h \leq n\right\}$

or

$$k = \begin{cases} 0 & , \text{ for } a < \delta \\ n & , \text{ for } \frac{a}{n} > \delta \\ \left\lfloor \frac{a}{\delta} \right\rfloor & , \text{ otherwise} \end{cases}$$

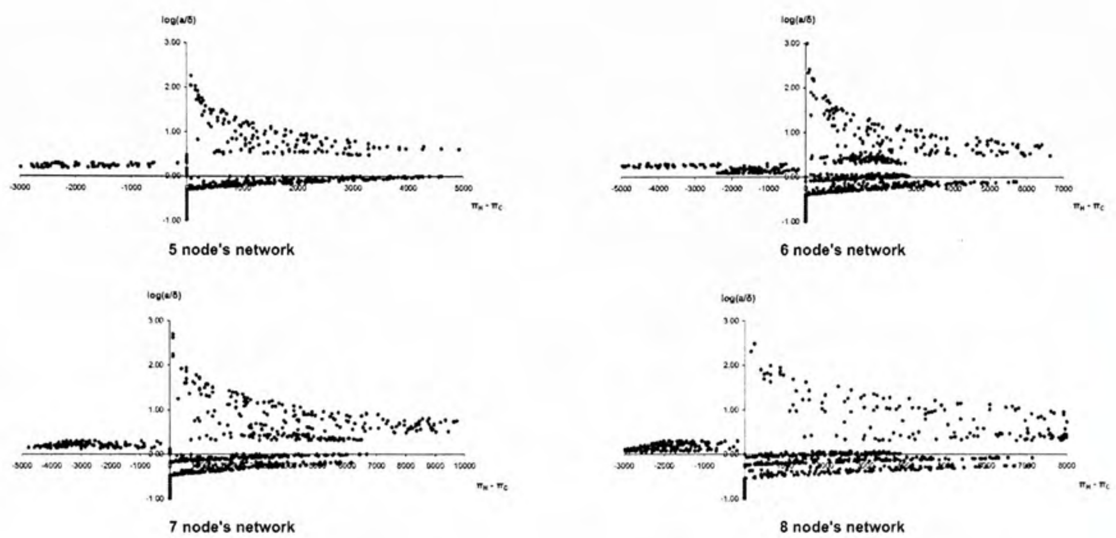


Figure 3.5: $log(\frac{a}{\delta})$ versus $\pi_H - \pi_C$

For hub-and-spoke structure, we have t routes and

$$p_j(k') = \begin{cases} \frac{a(t-1)}{2k'} - \delta & , \text{ for } k' > 0 \\ 0 & , \text{ for } k' = 0 \end{cases}.$$

Therefore

$$\pi_H = \begin{cases} t\left(\frac{a(t-1)}{2} - \delta\, k'\right) & , \text{ for } k' > 0 \\ 0 & , \text{ for } k' = 0 \end{cases}$$

where $k' = max\left\{h : \frac{a(t-1)}{2h} - \delta > 0, 0 \leq h \leq n\right\}$

or

$$k' = \begin{cases} 0 & , \text{ for } \frac{a(t-1)}{2} < \delta \\ n & , \text{ for } \frac{a(t-1)}{2n} > \delta \\ \left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor & , \quad \text{otherwise} \end{cases}$$

We wish to know which network structure perform better, and under what circumstances. For case $t \leq 2$, the two networks are identical, therefore to compare the two network structures, we assume that t, the original number of nodes in the network, is not less than three. Both k and k' cannot take value greater than n due to the fact that the total number of operators offering service is at most equal to the total number of operators in the network. Since $\left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor \geq \left\lfloor \frac{a}{\delta} \right\rfloor$ for $t \geq 3$, that is $k' \geq k$, there are five scenarios to be considered: -

Case 1 — $\left(k' = \left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor\right) = \left(k = \lfloor \frac{a}{\delta} \rfloor\right) = 0$

This implies

$$\begin{aligned} \tfrac{a(t-1)}{2\delta} &< 1 \\ \tfrac{a}{\delta} &< \tfrac{2}{t-1} \end{aligned}$$

In this case, $\pi_C = \pi_H = 0$.

Case 2 — $n \geq \left(k' = \left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor\right) > \left(k = \lfloor \frac{a}{\delta} \rfloor\right) = 0$

Note that in this case,

$$\begin{aligned} && \pi_C &\leq \pi_H \\ \iff && 0 &\leq t\left(\tfrac{a(t-1)}{2} - \delta \left\lfloor \tfrac{a(t-1)}{2\delta} \right\rfloor\right) \\ \iff && \left\lfloor \tfrac{a(t-1)}{2\delta} \right\rfloor &\leq \tfrac{a(t-1)}{2\delta} \end{aligned}$$

Also,

$$\begin{aligned} \because && n \geq \left\lfloor \tfrac{a(t-1)}{2\delta} \right\rfloor > 0 \quad \text{and} \quad \lfloor \tfrac{a}{\delta} \rfloor = 0 \\ \iff && n+1 > \tfrac{a(t-1)}{2\delta} \geq 1 \quad \text{and} \quad \tfrac{a}{\delta} < 1 \\ \therefore && 1 \leq \tfrac{a(t-1)}{2\delta} < \tfrac{t-1}{2} \quad \text{or} \quad 1 \leq \tfrac{a(t-1)}{2\delta} < n+1 \end{aligned}$$

$$\iff \frac{a}{\delta} \in \left[\frac{2}{t-1}, min\left(1, \frac{2(n+1)}{t-1}\right)\right) \text{ for this case.}$$

Therefore, when $\frac{a}{\delta} \in \left[\frac{2}{t-1}, min\left(1, \frac{2(n+1)}{t-1}\right)\right)$

$$\pi_C < \pi_H \quad \text{if} \quad \frac{a(t-1)}{2\delta} \notin \mathbb{Z}^+,$$

$$\text{else} \quad \pi_C = \pi_H.$$

Case 3 — $n \geq \left(k' = \left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor\right) \geq \left(k = \left\lfloor \frac{a}{\delta} \right\rfloor\right) > 0$

In this case

$$\frac{a(t-1)}{2\delta} \;<\; n+1,$$

and

$$\frac{a}{\delta} \;>\; 1$$

so $\frac{a}{\delta} \in \left(1, \frac{2(n+1)}{t-1}\right)$.

Hence,

$$\begin{aligned}
& \pi_C \lesseqgtr \pi_H \\
\iff & \frac{t(t-1)}{2}\left(a - \delta\left\lfloor \frac{a}{\delta}\right\rfloor\right) \lesseqgtr t\left(\frac{a(t-1)}{2} - \delta\left\lfloor \frac{a(t-1)}{2\delta}\right\rfloor\right) \\
\iff & \left\lfloor \frac{a(t-1)}{2\delta}\right\rfloor \lesseqgtr \frac{t-1}{2}\left\lfloor \frac{a}{\delta}\right\rfloor
\end{aligned}$$

That is, the turning points will occur each time $\frac{a}{\delta}$ or $\frac{a(t-1)}{2\delta}$

reach an integer value. Where,

$$\pi_C > \pi_H \quad \text{if} \quad \tfrac{a}{\delta} \notin \mathbb{Z}^+, \tfrac{a(t-1)}{2\delta} \in \mathbb{Z}^+$$

$$\pi_C = \pi_H \quad \text{if} \quad \tfrac{a}{\delta} \in \mathbb{Z}^+, \tfrac{a(t-1)}{2\delta} \in \mathbb{Z}^+$$

$$\pi_C < \pi_H \quad \text{if} \quad \tfrac{a}{\delta} \in \mathbb{Z}^+, \tfrac{a(t-1)}{2\delta} \notin \mathbb{Z}^+$$

Figure 3.6 shows a closer look at the plot for **Case 3** (from a network with four players, four routes).

To examine the situation when $\frac{a}{\delta} \notin \mathbb{Z}^+$ and $\frac{a(t-1)}{2\delta} \notin \mathbb{Z}^+$, we consider the two cases when t is odd and t is even.
Let $\frac{a}{\delta} = k + \varepsilon$ where $k \in \mathbb{Z}^+$ and $0 < \varepsilon < 1$.

i) t is odd, say $t = 2q + 1$ with $q \in \mathbb{Z}^+$.
Then

$$\begin{aligned} \left\lfloor \tfrac{a(t-1)}{2\delta} \right\rfloor &= \lfloor (k+\varepsilon)q \rfloor \\ &= \lfloor kq + \varepsilon q \rfloor \\ &= kq + \lfloor \varepsilon q \rfloor \end{aligned}$$

and

$$\frac{t-1}{2}\left\lfloor\frac{a}{\delta}\right\rfloor = qk$$

so

$$\pi_C \geq \pi_H.$$

We also note that when t is odd, then

$$\frac{a}{\delta} \in \mathbb{Z}^+ \Rightarrow \frac{a}{\delta}\left(\frac{t-1}{2}\right) \in \mathbb{Z}^+$$

so we would never have $\frac{a}{\delta} \in \mathbb{Z}^+$ and $\frac{a}{\delta}\left(\frac{t-1}{2}\right) \notin \mathbb{Z}^+$.

Thus, when t is odd, we have

$$\pi_C \geq \pi_H$$

with equality holding when $\frac{a}{\delta} \in \mathbb{Z}^+$.

ii) t is even, say $t = 2q$ with $q \in \mathbb{Z}^+$.

Then

$$\left(\frac{t-1}{2}\right)\left\lfloor\frac{a}{\delta}\right\rfloor = \left(q - \frac{1}{2}\right)k,$$

and

$$\begin{aligned}\left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor &= \left\lfloor (k+\varepsilon)\left(q-\tfrac{1}{2}\right) \right\rfloor \\ &= \left\lfloor k\left(q-\tfrac{1}{2}\right) + \varepsilon\left(q-\tfrac{1}{2}\right) \right\rfloor\end{aligned}$$

Hence, if k is even, then

$$\pi_C \geq \pi_H,$$

else if k is odd, then

$$\begin{aligned}\pi_C \geq \pi_H, \ \text{when} \ \varepsilon\left(q-\tfrac{1}{2}\right) &\geq \tfrac{1}{2} \\ \text{or} \qquad \varepsilon &\geq \tfrac{1}{t-1},\end{aligned}$$

and

$$\pi_C < \pi_H, \ \text{when} \qquad \varepsilon < \tfrac{1}{t-1}.$$

Case 4 — $\left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor > (k' = n) \geq \left(k = \left\lfloor \frac{a}{\delta} \right\rfloor\right) > 0$

$$\begin{aligned} \pi_C &\lesseqgtr \pi_H \\ \Longleftrightarrow \quad \tfrac{t(t-1)}{2}\left(a - \delta\left\lfloor \tfrac{a}{\delta} \right\rfloor\right) &\lesseqgtr t\left(\tfrac{a(t-1)}{2} - \delta\, n\right) \\ \Longleftrightarrow \quad \tfrac{2n}{t-1} &\lesseqgtr \left\lfloor \tfrac{a}{\delta} \right\rfloor \end{aligned}$$

Also,

$$\because \qquad \left\lfloor \tfrac{a(t-1)}{2\delta} \right\rfloor > n \geq \left\lfloor \tfrac{a}{\delta} \right\rfloor$$

$$\Longleftrightarrow \qquad \tfrac{a(t-1)}{2\delta} \geq n+1 \quad \text{and} \quad n+1 > \tfrac{a}{\delta}$$

$$\therefore \qquad n+1 > \tfrac{a}{\delta} \geq \tfrac{2(n+1)}{t-1}, \quad \text{i.e.} \quad \tfrac{a}{\delta} \in \left[\tfrac{2(n+1)}{t-1}, n+1 \right).$$

Note that if $\frac{2n}{t-1}$ is integer valued,

then since,

$$\begin{aligned} \tfrac{a}{\delta} &\geq \tfrac{2n}{t-1} + \tfrac{2}{t-1} \\ \left\lfloor \tfrac{a}{\delta} \right\rfloor &\geq \tfrac{2n}{t-1} \end{aligned}$$

so

$$\pi_H \geq \pi_C.$$

If $\frac{2n}{t-1}$ is not integer valued, say,

$$\frac{2n}{t-1} = q + \frac{p}{t-1}$$

where $q \in \mathbb{Z}^+, p \in \mathbb{Z}^+, 0 < p < t-1,$ and let

$$\begin{aligned} \tfrac{a}{\delta} &= \tfrac{2(n+1)}{t-1} \\ &= q + \tfrac{p}{t-1} + \tfrac{2}{t-1} + r, r \geq 0 \end{aligned}$$

then

$$\begin{aligned} \lfloor \tfrac{a}{\delta} \rfloor \geq \tfrac{2n}{t-1} \quad \text{when} \quad \tfrac{p+2}{t-1} + r &\geq 1 \\ \text{or when} \quad r &\geq 1 - \tfrac{p+2}{t-1}. \end{aligned}$$

Hence $\pi_H \geq \pi_C$ when

$$\frac{a}{\delta} \geq \left\lceil \frac{2n}{t-1} \right\rceil .$$

Case 5 — $\left\lfloor \frac{a(t-1)}{2\delta} \right\rfloor \geq \lfloor \frac{a}{\delta} \rfloor > (k' = k = n)$

$$\begin{aligned} \pi_C &\leq \pi_H \\ \iff \tfrac{t(t-1)}{2}(a - \delta\, n) &\leq t(\tfrac{a(t-1)}{2} - \delta\, n) \\ \iff t &\geq 3 \end{aligned}$$

$$\begin{aligned} &\because \text{ we only consider networks with } t > 3 \\ &\therefore \quad \pi_H > \pi_C, \text{if } \lfloor \tfrac{a}{\delta} \rfloor > n \end{aligned}$$

With these inequalities we can determine when π_C is larger (less) than or equals to π_H, in other words, complete network (hub-and-spoke) is preferable or both network structures perform equally well in terms of profitability.

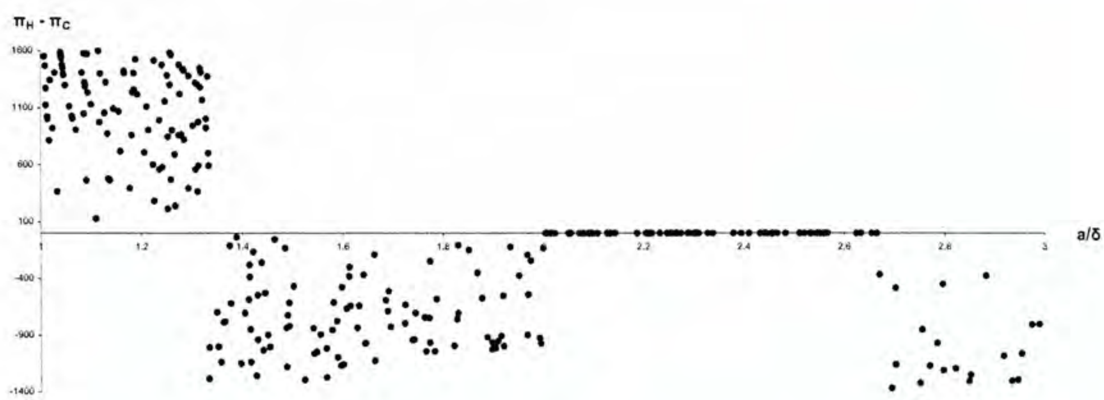


Figure 3.6: Close up look for $\pi_H - \pi_C$ versus $\frac{a}{\delta}$

Table 3.1 and table 3.2 are examples for network with $n = t = 4$ and $n = 4, t = 5$ respectively.

Under the assumption that all the operators share the same cost structure and that the ridership is the same for all origin-destination pair, we have just shown that there is an analytical relationship between the relative revenue/cost ratio and the choice of the network structure (complete network vs. hub-and-spoke) by the service providers. In the next part, we will try to extend the analytical result to a more realistic case, where each origin-destination pair may have different ridership, to see if similar analytical relationships between the revenue/cost ratio and the network structure exists.

Case	$\frac{a}{\delta}$	$\frac{a(t-1)}{2\delta}$	k	k′	$\lfloor \frac{a}{\delta} \rfloor$	$\frac{t-1}{2} \lfloor \frac{a}{\delta} \rfloor$		$\lceil \frac{a(t-1)}{2\delta} \rceil$	π_C		π_H
1	0.666	0.999	0	0	0	0	=	0	0	=	0
	$\frac{2}{3}$	1	0	1	0	0	<	1	0	=	0
2	0.667	1.0005	0	1	0	0	<	1	0	<	0.002
	0.999	1.4985	0	1	0	0	<	1	0	<	1.994
	1	1.5	1	1	1	1.5	>	1	0	<	2
	1.332	1.998	1	1	1	1.5	>	1	1.992	<	3.992
	$\frac{4}{3}$	2	1	2	1	1.5	<	2	2	>	0
3	1.999	2.9985	1	2	1	1.5	<	2	5.994	>	3.994
	2	3	2	3	2	3	=	3	0	=	0
	2.666	3.999	2	3	2	3	=	3	3.996	=	3.996
	$\frac{8}{3}$	4	2	4	2	3	<	4	4	>	0
	2.999	4.4985	2	4	2	3	<	4	5.994	>	1.994
	3	4.5	3	4	3	4.5	>	4	0	<	2
	3.333	4.9995	3	4	3	4.5	>	4	1.998	<	3.998
	$\frac{10}{3}$	5	3	4	3	4.5	<	5	2	<	4
	3.999	5.9985	3	4	3	4.5	<	5	5.994	<	7.994
4	4	6	4	4	4	6	=	6	0	<	8
	4.666	6.999	4	4	4	6	=	6	3.996	<	11.99
	$\frac{14}{3}$	7	4	4	4	6	<	7	4	<	12
	4.999	7.4985	4	4	4	6	<	7	5.994	<	13.99
5	5	7.5	4	4	5	7.5	>	7	6	<	14

Table 3.1: $\frac{a}{\delta}$ Range for a Network where $n = t = 4$

Case	$\frac{a}{\delta}$	$\frac{a(t-1)}{2\delta}$	k	k'	$\lfloor\frac{a}{\delta}\rfloor$	$\frac{t-1}{2}\lfloor\frac{a}{\delta}\rfloor$		$\lfloor\frac{a(t-1)}{2\delta}\rfloor$	π_C		π_H
1	0.499	0.998	0	0	0	0	$=$	0	0	$=$	0
2	0.5	1	0	1	0	0	$<$	1	0	$=$	0
	0.501	1.002	0	1	0	0	$<$	1	0	$<$	0.01
	0.999	1.998	0	1	0	0	$<$	1	0	$<$	4.99
3	1	2	1	2	1	2	$=$	2	0	$=$	0
	1.499	2.998	1	2	1	2	$=$	2	4.99	$=$	4.99
	1.5	3	1	3	1	2	$<$	3	5	$>$	0
	1.999	3.998	1	3	1	2	$<$	3	9.99	$>$	4.99
	2	4	2	4	2	4	$=$	4	0	$=$	0
	2.499	4.998	2	4	2	4	$=$	4	4.99	$=$	4.99
4	2.5	5	2	4	2	4	$<$	5	5	$=$	5
	2.999	5.998	2	4	2	4	$<$	5	9.99	$=$	9.99
	3	6	3	4	3	6	$=$	6	0	$<$	10
	3.499	6.998	3	4	3	6	$=$	6	4.99	$<$	14.99
	3.5	7	3	4	3	6	$<$	7	5	$<$	15
	3.999	7.998	3	4	3	6	$<$	7	9.99	$<$	19.99
	4	8	4	4	4	8	$=$	8	0	$<$	20
	4.499	8.999	4	4	4	8	$=$	8	4.99	$<$	24.99
	4.5	9	4	4	4	8	$<$	9	5	$<$	25
	4.999	9.998	4	4	4	8	$<$	9	9.99	$<$	29.99
5	5	10	4	4	5	10	$=$	10	10	$<$	30

Table 3.2: $\frac{a}{\delta}$ Range for a Network where $n = 4, t = 5$

3.2.2 Routes with Different Ridership

Another 1,000 cases were simulated. This time around, each origin-destination pair may have different ridership. a_j's and δ were generated similarly (pseudo-random numbers uniformly distributed between zero and 1000). 89 cases were excluded due to the extreme operating cost, $\delta \gg a_j$. About 68% of the remaining cases showed that hub-and-spoke structure is superior, with the improvement ranging from 0.19% to 49,950%. Whereas around 2% of the cases share the same profitability between two structures. In other words, less than 31% of the cases are actually in favour for the complete network structure. No clear cut range can be obtained as of the previous study but a pattern is observed, refer to Figure 3.7. However, we can still approximately draw lines to separate ranges where the hub-and-spoke structure outperform the complete structure. Figure 3.8 is the plot for a network with four players and eight nodes. The revenue-cost ratio between 18 and 41 (i.e. $18 < \frac{\sum\limits_{j \in M} a_j}{\delta} < 41$) is the indistinct range where no clear indication as of which network structure is better. As the number of nodes increases, the dividing line become clearer and a shrinking of the ambiguous range is observed.

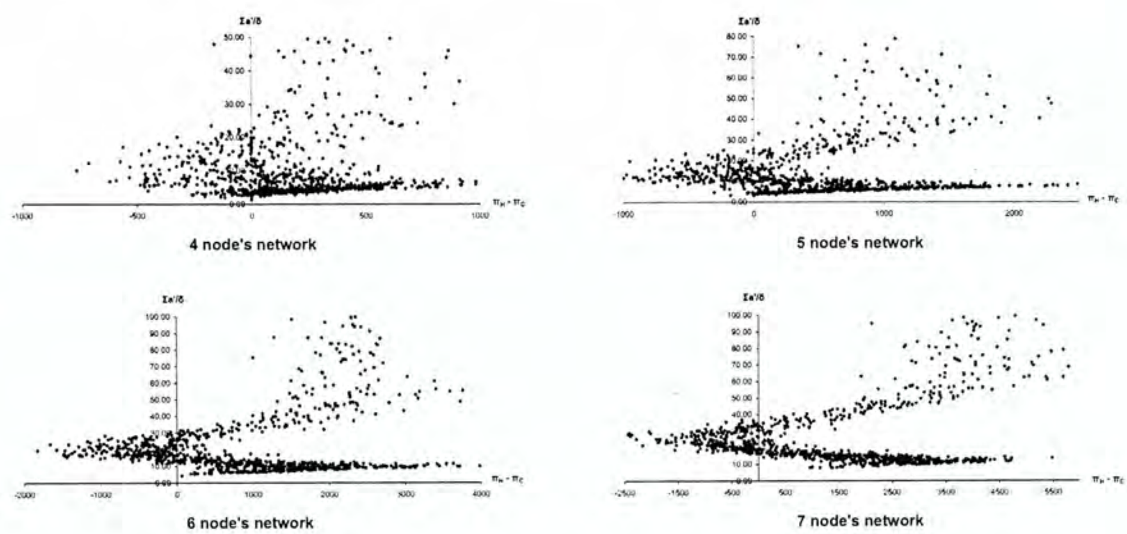


Figure 3.7: Cases with different ridership on each route

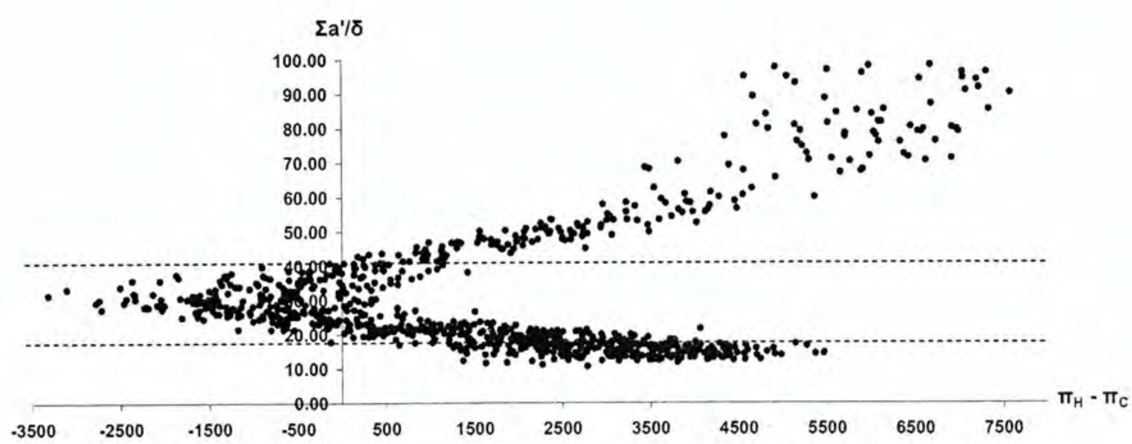


Figure 3.8: Approximate Range When Hub-and-spoke Network is Superior

Analytical Result

Now, the total profit for the complete network is $\sum_{i=1}^{n}\sum_{j=1}^{m} x_i^j\, p_j(k_j) = \sum_{j=1}^{\frac{t(t-1)}{2}} k_j\, p_j(k_j)$, where

$$p_j(k_j) = \begin{cases} \frac{a_j}{k_j} - \delta & , \text{ for } k_j > 0 \\ 0 & , \text{ for } k_j = 0 \end{cases} .$$

Therefore

$$\pi_C = \begin{cases} \sum_{j=1}^{\frac{t(t-1)}{2}} a_j - \delta \sum_{j=1}^{\frac{t(t-1)}{2}} k_j & , \text{ for } k_j > 0 \\ 0 & , \text{ for } k_j = 0 \end{cases}$$

where $k_j = min\left\{n, max\left\{0, max\left\{h : \frac{a_j}{h} - \delta > 0, h > 0\right\}\right\}\right\}$

or

$$k_j = \begin{cases} 0 & , \text{ for } \left\lfloor \frac{a_j}{\delta} \right\rfloor = 0 \\ \left\lfloor \frac{a_j}{\delta} \right\rfloor & , \text{ for } 0 < \left\lfloor \frac{a_j}{\delta} \right\rfloor \leq n \\ n & , \text{ for } \left\lfloor \frac{a_j}{\delta} \right\rfloor > n \end{cases}$$

For hub-and-spoke structure, we have t routes and the correspond ridership are redistributed as mentioned earlier in section

3.1. Let $a'_j = \sum_{a_i \in \mathcal{A}_j} a_i$, where $\mathcal{A}_j$ is the set of ridership (in the original complete network) which were partially served by current route j (in the correspond hub-and-spoke network). Refer back to Figure 3.2 for example, $\mathcal{A}_1 = \{a_1, a_4, a_6\}$, $a'_1 = a_1 + a_4 + a_6$ etc. Hence

$$p_j(k') = \begin{cases} \frac{a'_j}{2k'_j} - \delta & , \text{ for } k'_j > 0 \\ 0 & , \text{ for } k'_j = 0 \end{cases},$$

and

$$\pi_H = \begin{cases} \sum_{j=1}^{t} \frac{a'_j}{2} - \delta \sum_{j=1}^{t} k'_j & , \text{ for } k'_j > 0 \\ 0 & , \text{ for } k'_j = 0 \end{cases}$$

$$\text{where } k'_j = min\left\{n, max\left\{0, max\left\{h : \frac{a'_j}{2h} - \delta > 0, h > 0\right\}\right\}\right\}$$

or

$$k'_j = \begin{cases} 0 & , \text{ for } \left\lfloor \frac{a'_j}{2\delta} \right\rfloor = 0 \\ \left\lfloor \frac{a'_j}{2\delta} \right\rfloor & , \text{ for } 0 < \left\lfloor \frac{a'_j}{2\delta} \right\rfloor \leq n \\ n & , \text{ for } \left\lfloor \frac{a'_j}{2\delta} \right\rfloor > n \end{cases}$$

Again, we wish to know which network structure perform better, and under what circumstances. The original number of nodes in the network, t, must not less than three and all k_j and k'_j cannot possibly greater than n. Note that $\sum_{j=1}^{\frac{t(t-1)}{2}} a_j = \sum_{j=1}^{t} \frac{a'_j}{2}$. However, we cannot further simplify k'_j as a'_j, which depends on $\mathcal{A}_j$, vary for different t. Thus, no general conclusion can be drawn here. Further more, to compare the two network structure for routes with different ridership, there are more than two ranges of variables to consider; as oppose to only two — the range of k and k' — in networks where all routes have the same ridership.

3.2.3 Network with Player- and Route-specific Profit Function

Similar to the previous comparison, we observed a pattern of the plot $\frac{\sum a_j}{t \sum \delta_i}$ versus $\pi_H - \pi_C$. Refer to figure 3.9 for the plots of network with 4, 5, 6 and 7 nodes. With the same difficulty faced in the previous part, we cannot obtain any analytical result here either.

□ **End of chapter.**

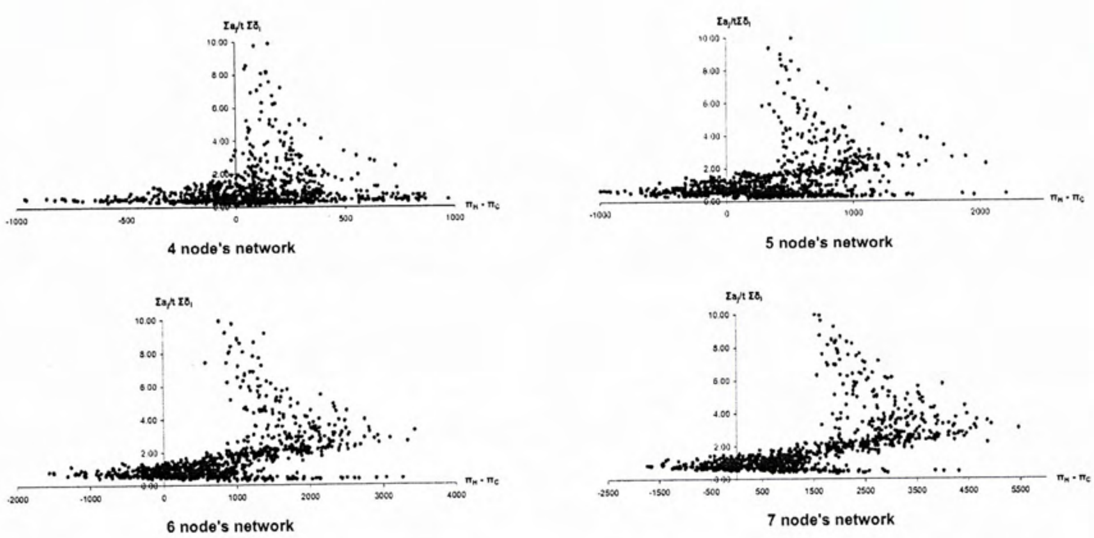


Figure 3.9: Plots of $\frac{\sum a_j}{t \sum \delta_i}$ vs. $\pi_H - \pi_C$ — Player Specific Payoff Function

Chapter 4

Elastic Demand

In this chapter, we investigate networks with service-quality-based elastic demand. We only manage to construct a model that maximizes the total net profit (which give "centrally controlled" solutions). The "competitive" equilibrium solution is harder to get as we cannot find the respective potential function for the game.

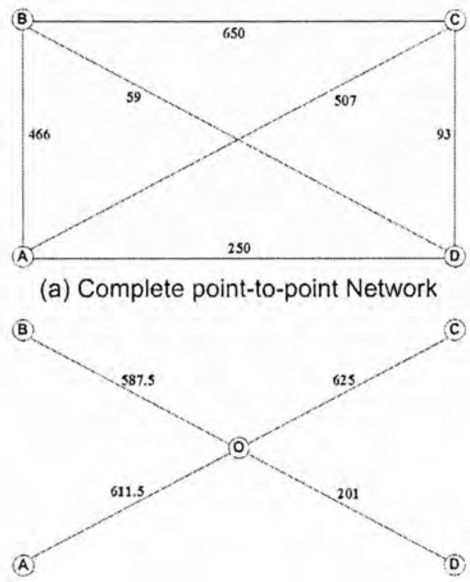
4.1 Congestion Game Model with Service-Quality-Based Elastic Demand

So far in all presented models, they share a common restrictive assumption — network structure and the total service bundle being offered will not affect the overall ridership. That is to say, all passenger who would travel from town X to town Y with direct service will still travel even if the trip involves an interchange via a central hub. Clearly , the convenience level, travel time and possibly travel cost will not be the same for the two trips. Also, apparently with wider service coverage and greater number of operators servicing a particular route (higher frequency), it will attract more ridership. Hence, a more realistic approach is to assume that both the network structure and the total service bundle will affect the overall ridership. Incorporating this assumption in the base model would mean that the profit function of a route may depend not only on the number of operators serving that route, but on the entire set of services offered by all the operators.

Recall that the ridership redistribution assumption made in section 3.1(Figure 3.2), from a complete network structure to a hub-and-spoke one. No loss of ridership is assumed, even if some of the routes may not be serviced. Consider the following example. Figure 4.1 shows the ridership on each network structure, and 4.2 is the solution. Note that for the hub-and-spoke network (right hand side of Figure 4.2), route AO, BO, CO, and

DO are denoted by $j = 1, 2, 3, 4$ respectively. The solution for the hub-and-spoke problem is to service route AO, BO, and CO while abandoning DO. Strictly speaking, operators who are servicing those three aforementioned routes will suffer some loss of ridership. As passenger who wish to commute from A, B, or C to D can only reach O under this service bundle. Assuming the same profit function described earlier, the fee charges are cut to half as compare to the charges on complete network, which is one dollar. The loss of ridership is therefore at most $a_1 + a_4 + a_6$ and the subsequent loss in revenue would be a sum of $\frac{(a_1+a_4+a_6)\theta}{2}$ less than the revenue acquired by solving $(G2)$, where θ denote the percentage lost of ridership. Then it make sense to anticipate that at least one operator will cover the losing route, DO, as long as the loss induced by the loss of ridership is greater than the loss from route DO alone.

With the more realistic model, is it still possible to model the problem as a potential game? We may need to develop a formulation of the game for this model and investigate algorithms for finding the Nash equilibrium for the different prototypical network structures. We can even look for models where additional



(a) Complete point-to-point Network

(b) Hub-and-spoke Network

Figure 4.1: Case Example

ridership may be induced, for example, in a backbone network where the main truck routes are bus-only lanes that enable the bus to traverse the congested city centre faster than private cars. Or the ridership between an origin-destination pair will be affected by the number of interchanges required in the trip and the expected travel time and cost. Below, we will show the proposed model, $(GD4)$, which maximize the total net profit. We assume that the ridership between an origin-destination pair is affected by the total number of routes being serviced, following a logistic distribution. Note that we can always try with other distribution according to specific situation and model it similarly as follows, as long as the profit is the function of h and k,

Complete Structure

a_1	a_2	a_3	a_4	a_5	a_6
466	650	93	250	59	507

δ
219

k_j	$p_1(k_1)$	$p_2(k_2)$	$p_3(k_3)$	$p_4(k_4)$	$p_5(k_5)$	$p_6(k_6)$
1	247.00	431.00	-126.00	31.00	-160.00	288.00
2	14.00	106.00	-172.50	-94.00	-189.50	34.50
3	-63.67	-2.33	-188.00	-135.67	-199.33	-50.00
4	-102.50	-56.50	-195.75	-156.50	-204.25	-92.25

k_1^*	k_2^*	k_3^*	k_4^*	k_5^*	k_6^*
2	2	0	1	0	2

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	
Player 1	1	1	0	1	0	1	
Player 2	1	1	0	0	0	1	
Player 3	0	0	0	0	0	0	
Player 4	0	0	0	0	0	0	$\sum_{j=1}^{m}\sum_{y=1}^{k_j^*} p_j(y)$
$\sum_{y=1}^{k_j^*} p_j(y)$	261.00	537.00	0.00	31.00	0.00	322.50	1152
							$\sum_{j=1}^{m} k_j^* p_j(k_j^*)$
$k_j^* p_j(k_j^*)$	28.00	212.00	0.00	31.00	0.00	69.00	340

Hub-and-spoke Structure

a_1	a_2	a_3	a_4
611.5	587.5	625	201

δ
219

k_j	$p_1(k_1)$	$p_2(k_2)$	$p_3(k_3)$	$p_4(k_4)$
1	392.50	368.50	406.00	-18.00
2	86.75	74.75	93.50	-118.50
3	-15.17	-23.17	-10.67	-152.00
4	-66.13	-72.13	-62.75	-168.75

k_1^*	k_2^*	k_3^*	k_4^*
2	2	2	0

	$j=1$	$j=2$	$j=3$	$j=4$	
Player 1	1	1	1	0	
Player 2	1	1	1	0	
Player 3	0	0	0	0	
Player 4	0	0	0	0	$\sum_{j=1}^{m}\sum_{y=1}^{k_j^*} p_j(y)$
$\sum_{y=1}^{k_j^*} p_j(y)$	261.00	537.00	613.50	31.00	943
					$\sum_{j=1}^{m} k_j^* p_j(k_j^*)$
$k_j^* p_j(k_j^*)$	28.00	212.00	36.00	31.00	337

Figure 4.2: Case Example's Solution

where h denotes total number of routes under service, and k is the number of operators serving route j.

We define the profit on route j for an operator as:

$$p_j(h,k) = \begin{cases} \frac{a_j}{k\left(1+e^{-(h-m/2)}\right)} - \delta & \text{for } k > 0 \\ 0 & \text{if } k = 0, \end{cases}$$

Let

$$u_h = \begin{cases} 1 & \text{if} \sum_{j=1}^{m}\sum_{k=1}^{n} y_{jk} = h \\ 0 & \text{otherwise} \end{cases}, \text{for } h = 0, 1, ..., m;$$

and let

$$z_h = \sum_{j=1}^{m}\sum_{k=0}^{n} k\, p_j(h,k)\, y_{jk},$$

where u_h's are binary variables whereas z_h denotes the total revenue when there are h routes under service. We need to ensure a unique u_h, which can be enforced by the following constraints:

$$\sum_{h=0}^{m} u_h = 1;$$
$$\sum_{h=0}^{m} h\, u_h = \sum_{j=1}^{m}\sum_{k=1}^{n} y_{jk}.$$

$(GD4)$ is given by:

$$(GD4): \quad Maximize \sum_{h=0}^{m} u_h\, z_h \qquad (4d.0)$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\, y_{jk}, \ \forall j \in M; \tag{2.1}$$

$$\sum_{j=1}^{m} x_i^j \le b_i, \ \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \tag{1.3}$$

$$\sum_{k=0}^{n} y_{jk} = 1, \ \forall j \in M; \tag{2.4}$$

$$y_{jk} \in \{0,1\}, \ k \in N \cup \{0\}; \ j \in M; \tag{2.5}$$

$$\sum_{h=0}^{m} u_h = 1; \tag{4d.6}$$

$$\sum_{h=0}^{m} h\, u_h = \sum_{j=1}^{m} \sum_{k=1}^{n} y_{jk}; \tag{4d.7}$$

$$u_h \in \{0,1\}, \ \forall h \in M \cup \{0\}. \tag{4d.8}$$

$$z_h = \sum_{j=1}^{m} \sum_{k=0}^{n} k\, p_j(h,k)\, y_{jk}; \tag{4d.9}$$

Constraint (4d.6) would specify a unique h. This ensure the correct RHS for (4d.7) and for the objective function.

However, the objective function of $(GD4)$ is a non-linear function. An equivalent mixed-integer programming, $(GD5)$, will be presented below. We introduce binary variables v_{hjk}, intended to be:

$$v_{hjk} = \begin{cases} 1 & \text{if } u_h = 1 \text{ and } y_{jk} = 1 \\ 0 & \text{otherwise} \end{cases}$$

which can be enforced by the following constraints:

$$\begin{aligned} \sum_{j=1}^{m}\sum_{k=1}^{n} v_{hjk} &= h\,u_h, \ \forall h \in M \cup \{0\}; \\ \sum_{h=0}^{m} v_{hjk} &= y_{jk}, \ \forall k \in N \cup \{0\}; \ j \in M; \\ u_h + y_{jk} &\leq v_{hjk} + 1, \ \forall h \in M \cup \{0\}; k \in N \cup \{0\}; \ j \in M; \end{aligned}$$

A mixed-integer programming of $(GD5)$ is given by:

$$(GD5): \quad Maximize \sum_{h=0}^{m}\sum_{j=1}^{m}\sum_{k=0}^{n} k\,p_j(h,k)v_{hjk} \tag{5d.0}$$

subject to:

$$\sum_{i=1}^{n} x_i^j = \sum_{k=0}^{n} k\,y_{jk}, \ \forall j \in M; \tag{2.1}$$

$$\sum_{j=1}^{m} x_i^j \leq b_i, \ \forall i \in N; \tag{1.2}$$

$$x_i^j \in \{0,1\}, \ i \in N; \ j \in M; \tag{1.3}$$

$$\sum_{k=0}^{n} y_{jk} = 1, \ \forall j \in M; \tag{2.4}$$

$$y_{jk} \in \{0,1\}, \ k \in N \cup \{0\}; \ j \in M; \tag{2.5}$$

$$\sum_{h=0}^{m} u_h = 1; \tag{4d.6}$$

$$\sum_{h=0}^{m} h\, u_h = \sum_{j=1}^{m} \sum_{k=1}^{n} y_{jk}; \tag{4d.7}$$

$$u_h \in \{0,1\}, \ \forall h \in M \cup \{0\}. \tag{4d.8}$$

$$\sum_{j=1}^{m} \sum_{k=1}^{n} v_{hjk} = h\, u_h, \ \forall h \in M \cup \{0\}; \tag{5d.9}$$

$$\sum_{h=0}^{m} v_{hjk} = y_{jk}, \ j \in M; \ k \in N \cup \{0\}; \tag{5d.10}$$

$$u_h + y_{jk} \leq v_{hjk} + 1, \ , \ h \in M \cup \{0\}; \ j \in M; \ k \in N \cup \{0\}; \tag{5d.11}$$

$$v_{hjk} \in \{0,1\}, \ h \in M \cup \{0\}; \ j \in M; \ k \in N \cup \{0\}; \tag{5d.12}$$

$$\sum_{h=0}^{m} \sum_{j=1}^{m} \sum_{k=0}^{n} k\, p_j(h,k) v_{hjk} \geq 0. \tag{5d.13}$$

Constraint (4d.6) would specify a unique h. This ensure the correct RHS for (4d.7) and for the objective function. While binary variables v_{hjk} are enforced by constraints (5d.9) to (5d.11). Since operators will not offer any service if the overall operation is unprofitable, we introduce the last constraint, (5d.13), to reduce the feasible region.

Observation 9:

Constraint (5d.13) reduce the computational time from days to only a few minutes for some extreme cases.

Similar to ($G2$), we might want to model a competitive-demand-elastic model as a potential game model. Unfortunately, we have no idea as of how to model an equivalent potential game yet. However, we can proceed from the previous net profit maximizing solution to find the Nash equilibrium solution with the algorithm below:

```
    Solve (GD4) to get the initial k̃_j set

for (each route j where k̃_j=1) do
    if (the market can bear more than n, i.e.  when a_j/((1+e^{m/2-h})n) > δ) then
        k_j = total number of operators, n
    else
        k_j = as many operators as the market can bear, a_j/((1+e^{m/2-h})δ)
    end if
end for

for (each route j where k̃_j=0) do
    for (first to last ranked operator) do
        if (overall profit increase by including this route) then
```

```
            k_j = 1
            h = h + 1
        end if
    end for
end for

Return Strategy set
```

The idea is to find the maximum possible profit from the network first. Only then we allow the operators to compete for the route. That is, with the optimal solution from $(GD4)$, we proceeds to let as many operators as the market can bear to enter profitable routes. While having the number of routes under service, h, unchanged. Finally, we add in the following consideration. During the competition, operators will consider servicing unprofitable routes as long as the lose can be covered by the extra induced profit from (ridership on) the other servicing routes.

To take into consideration that when more operator servicing a route, more ridership induced, we may replace all h, u_h with

g, u_g respectively and define the following:

$$u_g = \begin{cases} 1 & \text{if} \sum_{j=1}^{m} \sum_{k=1}^{n} k\, y_{jk} = g, \\ 0 & \text{otherwise}; \end{cases} \quad \text{for } g = 0, 1, ..., m+n,$$

and

$$p_j(g,k) = \begin{cases} \frac{a_j}{k\left(1+e^{-(g-mn/2)}\right)} - \delta & \text{for } k > 0, \\ 0 & \text{if } k = 0, \end{cases}$$

while changing constraint (4d.7) to

$$\sum_{g=0}^{m} g\, u_g = \sum_{j=1}^{m} \sum_{k=1}^{n} k\, y_{jk}; \tag{4d.7'}$$

Note that g refers to the total number of routes offered by all operators, where the same route served by two operators counts as two routes.

4.1.1 Network with Service-Quality-Based Elastic Demand

We also numerically investigate the comparison of a hub-and-spoke and a complete network. Again, we observed some pat-

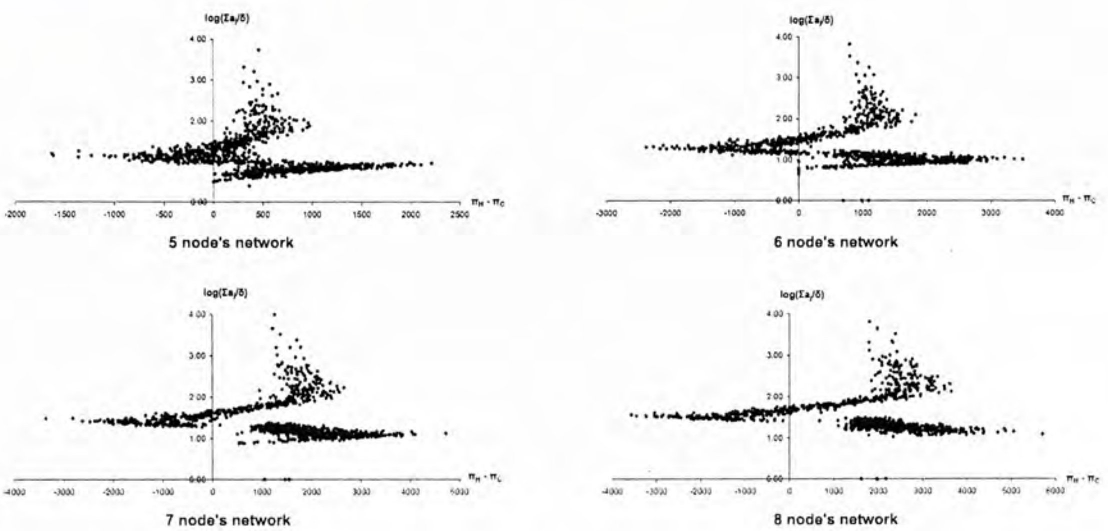


Figure 4.3: Plots of $log(\frac{\sum a_j}{\delta})$ vs. $\pi_H - \pi_C$ — Service-Quality-Based Elastic Demand

terns when we plot the graph for $log(\frac{\sum a_j}{\delta})$ versus $\pi_H - \pi_C$, similar to previous results. π_H and π_C refer to the overall total profit from all operators when a hub-and-spoke network and a complete network is implemented respectively. Figure 4.3 shows the plots of network with 5, 6, 7 and 8 nodes. We face the same difficulty in obtaining an analytical result.

□ **End of chapter.**

Chapter 5

Conclusion

5.1 Future Work

We will discuss various directions of research to be explored.

5.1.1 Impact of Network Design and structure

So far, we only consider two types of network — complete network and hub-and-spoke network. Other prototypical network structures may be worth investigating. Some such prototypical structures are spanning trees, Manhattan grids, and hub-and-spoke with ring-roads. (See Figure 5.1.) We would like to explore how the structure of the network and the revenue/cost ratio impact:

1. the equilibrium solution chosen by the operators?

2. the overall profit for the operators?

3. the set of routes being served?

For each network structure, would it be possible to identify, for each link in the network, a threshold value for the operating cost whereby operators would be induced to offer services on that link?

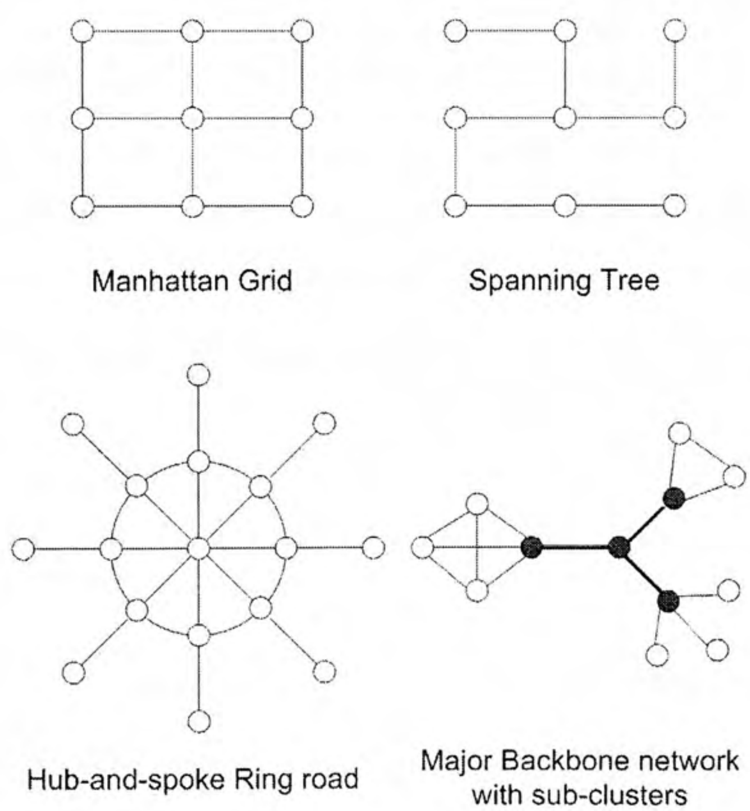


Figure 5.1: Other Prototypical Network Structures

5.1.2 Non-cooperative and Cooperative Games

The potential game framework is a non-cooperative game framework. Operators may consider to cooperate when their resources are limited (which is mostly true). One operator may offers services on a part of the network that serve as feeder links to the service provided by another operator, and vice versa. Developing a cooperative game-theoretic model may help us to compare and contrast the equilibrium solutions of both setting, the cooperative game and non-cooperative settings.

We are then left with questions like:

1. What is the impact on the resulting service level offered to the public transit ridership?
2. What is the impact on the overall profit for the service providers?
3. What is the appropriate profit-sharing scheme to induce higher profits or more comprehensive services for the public?

5.1.3 Joint game-theoretic model of both passenger and providers

Hopefully with this project we can develop a comprehensive game-theoretic model that enable us to seek for a balance among the service providers' strategic competition and the interplay between the basket of services provided, and the total ridership of the system as a whole. An appropriate representation of the model may be a bi-level one. Where the "upper-level" would represent the strategic game among the service providers as they select the services to be offered to maximise their individual profit. While the "lower-level" is the game between the public and the operators as a group, in that the public may be diverted to other forms of transport (e.g. taxis, private vehicles) if the availability and service quality (e.g. interchanges required, circuitous routes, travel time) of the basket of services offered by the operators are too low. The two levels of the bi-level problem are interlinked since the choice of the public — to utilise public transit or not — would affect the potential ridership of the system and thus impact the potential profit of the operators. We plan to develop this model and investigate algorithms for finding its equilibrium solution(s), and study the impact of the parameters (costs, network structure, demand function, etc.) on the resulting equilibrium.

By investigating this bi-level game, we may obtained insight into the relationship among the network infrastructure, competitive situation between operators and the impact on the type and level of services offered to the public. These relationship could further guide us in decision making on possible infrastructural investments and incentives to offer both operators and the riding public, which is very helpful for the government authorities and to ensure a public transit system that well-serves the public and benefits the community in terms of costs, convenience, quality, environmental impact and other concerns being designed.

□ **End of chapter.**

Bibliography

[1] Bureau of Economic Analysis, U.S. Department of Commerce 26^{th} September 2007. News Release: BEA Introduces New Measures of the Metropolitan Economy. http://www.bea.gov/newsreleases/regional/gdp_metro /gdp_metro_newsrelease.htm.

[2] Demographia, 7^{th} May, 2005. London Boroughs Commuting Distance and Geographical Size. http://www.demographia.com/db-lonborcommute.htm.

[3] Demographia, 9^{th} January, 2005. Southeast England New Towns Commuting Distance and Urban Area Geographical Size. http://www.demographia.com/db-seuknewtowns.htm.

[4] Malaysian Automotive Association (MAA) 2006/2007 Registration, Total Malaysia Analy-

sis - By Makes and Segment, 10^{th} April 2008. http://www.maa.org.my/info_summary.htm, Jan 2008. Zipped files under Market Outlook.

[5] L. Blumrosen and S. Dobzinski. Welfare maximization in congestion games. *IEEE Journal on Selected Areas in Communications*, 25(6), August 2007.

[6] L. Castelli, G. Longo, R. Pesenti, and W. Ukovich. Two-player noncooperative games over a freight transportation network. *Transportation science*, 38(2), 2004.

[7] D. Chakrabarty, A. Mehta, V. Nagarajan, and V. Vazirani. Fairness and optimality in congestion games. In *Proceedings of the 6th ACM conference on Electronic commerce, 2005.*

[8] A. Fabrikant, C. Papadimitriou, and K. Talwar. The complexity of pure nash equilibria. In *Proceedings of the ACM Symposium of Theory of Computing, 2004.*

[9] C. Fisk. Game theory and transportation systems modelling. *Transportation Research B*, 18B(4/5), 1984.

[10] R. Forstall and J. Fitzsimmons. Metropolitan growth and expansion in the 1980s. Working Paper No. 6, Population

Division, U.S. Bureau of Census, Washington D.C., 20233. U.S.A, 1993.

[11] Y. Hollander and J. Prashker. The applicability of non-cooperative game theory in transport system analysis. *Transportation*, 33.

[12] S. Ioeng, R. McGrew, E. Nudelman, Y. Shoham, and Q. Sun. Fast and compact: A simple class of congestion games. In *20th National Conference on Artificial Intelligence (AAAI)*.

[13] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13(1), March 1996.

[14] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14.

[15] J. F. Nash. Equilibrium points in n-person games. *Proceeding of the National Academy of Sciences*, 36(1), 1950.

[16] R. W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2, 1973.

[17] B. Voecking. Congestion games: Optimization in competition. In *Proc. of the 2nd Algorithms and Complexity in Durham Workshop, 2006.*

[18] J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers, Part II*, 1, 1952.

[19] H. Yang. and X. Zhang. Modelling competitive transit and road traffic information services with heterogeneous endogenous demand. *In Transportation Research Record: Journal of the Transportation Research Board, No. 1783, National Research Council, Washington, D.C.*, 2002.