A CONGESTION PRICING APPROACH TO REDUCE TRAFFIC JAMS IN URBAN ARTERIAL NETWORKS

AYŞE NİLAY ALPAY

AUGUST 2019

A CONGESTION PRICING APPROACH TO REDUCE TRAFFIC JAMS IN URBAN ARTERIAL NETWORKS

A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF NATURAL
AND APPLIED SCIENCES OF ÇANKAYA UNIVERSITY

BY

AYŞE NİLAY ALPAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER OF SCIENCE IN INDUSTRIAL ENGINEERING
DEPARTMENT

AUGUST 2019

Title of the Thesis: A Congestion Pricing Approach to Reduce Traffic Jams in Urban Arterial Networks

Submitted by Ayşe Nilay ALPAY

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Can ÇOĞUN
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Ferda Can ÇETİNKAYA
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Hakan ÖZAKTAŞ
Supervisor

Examination Date: 19.08.2019

Examining Committee Members

Assoc. Prof. Dr. Hediye TÜYDEŞ YAMAN	(METU)
Assist. Prof. Dr. Nureddin KIRKAVAK	(Çankaya Univ.)
Assist. Prof. Dr. Hakan ÖZAKTAŞ	(Çankaya Univ.)

STATEMENT OF NON-PLAGIARISM PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Ayse Nilay ALPAY

Signature: aysenilay

Date: 19.08.2019

ABSTRACT

A CONGESTION PRICING APPROACH TO REDUCE TRAFFIC JAMS IN URBAN ARTERIAL NETWORKS

ALPAY, Ayşe Nilay

M.Sc., Department of Industrial Engineering

Supervisor: Assist. Prof. Dr. Hakan ÖZAKTAŞ

AUGUST 2019

A novel approach is introduced to aid congestion toll management for a simple city network. Vickrey’s single bottleneck model is effective for a single commuter’s route; however even in simple road networks congestion in a certain zone will affect the flow values in the rest of the network. Using a step by step methodology to determine equilibrium flow values, it is possible to use the single bottleneck approach separately for all bottleneck queue formations with some simplifying assumptions. Average waiting time is used to evaluate the severity of a queue and is computed with simulation runs. An analytical derivation for the waiting time is also made and a paradoxical result is also highlighted.

Keywords: Traffic jams, congestion pricing, discrete event simulation

ÖZ

ŞEHİR İÇİ ANA ARTER AĞLARINDA TRAFİK TIKANIKLIĞINI GİDERMEK İÇİN SIKIŞIKLIK FİYATLANDIRMA YAKLAŞIMI

ALPAY, Ayşe Nilay

Yüksek Lisans, Endüstri Mühendisliği Anabilim Dalı

Tez Danışmanı: Assist. Prof. Dr. Hakan ÖZAKTAŞ

AĞUSTOS 2019

Basit bir şehir ağı için trafik sıkışıklığı yönetimine yardımcı olacak yeni bir yaklaşım getirilmektedir. Vickrey tek darboğaz modeli, tek bir işe gidiş rotası için etkilidir. Ancak basit yol ağlarında bile belli bir bölgedeki tıkanıklık, ağın geri kalanındaki akış değerlerini etkileyecektir. Denge akış değerlerini belirlemek için adım adım bir yöntem kullanarak, bazı basitleştirici varsayımlarla birlikte tüm darboğazdaki kuyruk noktaları için tekli darboğaz yaklaşımını ayrı ayrı kullanmak mümkündür. Kuyruğun derecesini ölçmek için kullanılan ortalama bekleme süresi simülasyon yaklaşımıyla hesaplanmıştır, bunun yanında bu değer için bir analitik formül de bulunmuştur. Ortalama bekleme süresi ile ilgili tespit edilen bir paradoks da vurgulanmaktadır.

Anahtar Kelimeler: Trafik tıkanıklığı, sıkışıklık fiyatlandırma, kesikli olay benzetimi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Increase of the number of private cars in city centers have resulted with traffic jams and congestion making city life unbearable. Public transportation and carpooling are obvious solutions however it is still necessary to discourage drivers from using their private cars to travel through city traffic. Naturally, congestion management has also been the subject of academic research.

We have started this study by reviewing worldwide toll implementations for traffic congestion. A summary of this review is given in Section 2.1. This will be followed by a short review of some important academic studies in Section 2.2. Since our work is primarily based on the single bottleneck model developed by Vickrey (1969) we have given a detailed description of the single bottleneck model in Sections 3.1 and 3.2. Vickrey’s elementary approach has been the basis of many successor studies in the later years.

Our main work has been explained through Section 3.3: We have started with a single bottleneck congestion model and developed a simulation tool to evaluate performance depending on the incoming traffic and the service capacity. The idea of the single bottleneck was extended so as to be used separately at different congested zones in a simple road junction network. The methodology is explained with two networks. An analytical derivation of the average waiting time which was originally obtained with a simulation tool is made. As a consequence of this derivation we have arrived at a paradox which is interesting for discrete event simulation models of queues.

CHAPTER 2

AN OVERVIEW OF CONGESTION PRICING

Tolling of motorways and bridges is common for many cities and is an effective approach to reduce congestion however the primary objective is usually to obtain revenue for municipalities. When congestion management is considered the idea is to discourage residents driving their cars in the busy city centers marked by a virtual cordon where entry/exit of vehicles are subject to fees such as the system implemented in central London.

Congestion management has five primary interrelated objectives (Yüksel, 2004):

- To reduce traffic jams
- To improve bus services
- To reduce uncertainty in travel duration estimations
- To have a fairer distribution of wealth
- To increase sustainability of resources

In this chapter the famous examples of congestion tolling are explained to be followed by an overview of the academic perspective to the congestion management problem.

2.1 Congestion Tolling Applications

Although traffic congestion is a very severe problem for most metropolitan areas of the world actual implementation examples are very limited due to reluctance of political decision makers (Giuliano, 1992; Banister, 2003). Reasonably successful examples have been implemented for London and Singapore. These two examples along with others will be briefly explained in this section.

2.1.1 Congestion Tolling Overview in London

London is one of the big cities which encountered traffic congestion, so a lot of proposals have been prepared since the early 1960s. The number of cars increased from 500.000 in 1958 to over one million in 1963 (Leape, 2006). Alternative congestion pricing mechanisms were proposed by the Smeed Report prepared in 1964. However, at that time, determining of optimal tolls for entering city center zones were considered to be complicated (unlike determining toll amounts through a bridge or a motorway) (Leape, 2006). Later, in 1974, tolls of £1.00 for private vehicles and £3.00 for commercial vehicles while entering the business districts were considered but unimplemented due to political reasons (Yüksel, 2004). A working team for Review of Charging Options for London (ROCOL) proposed a daily pass with unlimited number of entries instead of charging every entry into the zone.

After years of debate, congestion charging system (which had been promised by the newly elected mayor Ken Livingstone) started in central London on 17 February 2003 (Kearns, 2014). Tolling was implemented between 07:00 – 18:30 during weekdays only (excluding Saturdays, Sundays and other public holidays). Daily rate of £5.00 was charged regardless of the number of entries and exits within the day for all motor vehicles with the exception of taxis, buses and motorcycles (Anas & Lindsey, 2011), (Yüksel, et al., 2010). 90% deduction was applied for residents of the toll zone. Disabled drivers' licenses are exempt from tolling as well (Leape, 2006). Naturally, exemption of taxis caused an increase in their usage since the start of congestion charging (Anas & Lindsey, 2011).

Despite the worldwide fame and success of London congestion charging (LCC), obtained results have been below expectations simply because almost half of the vehicles were exempt (Ingles, 2009). While net income was £93 million in between 2004 and 2005, net income in 2006-2007 was £123 million (Yüksel, et al., 2010). The objective had been to dedicate all net income on public transportation, therefore 80% of £137 million net income in between 2007 and 2008 was spent to improve bus lanes, the rest of the income was spent to enhance planning, roads and bridges, road safety, freeways, walking trails and bicycle roads (Ingles, 2009).

It also turned out that the operating expenses of congestion charging were higher than expected. Instead of toll booths or booking offices admission charges could be paid in kiosks, gas stations. Payments through mail orders, credit cards (via phone, SMS or the Internet) were also accepted. A central database kept daily track of these payments and to check the payments automatic plate recognition technology was used with the help of cameras located in all entry and exit points of the toll zone (Yüksel, et al., 2010). Penalty fees were levied for vehicles entering the toll zone without any payment. Plate recognition technology did not work perfectly resulting with many customer complaints (Leape, 2006). Criticisms of unfairness have been made since flat rate tolling does not account the amount of distance travelled within the toll zone.

Therefore, Transport for London started to research new recognition technologies, and then identified that "Tag and beacon" systems have high accuracy rate. However, cost of tagging and detection of tags have been calculated more than the total revenue collected (Yüksel, et al., 2010), (Ingles, 2009), (Leape, 2006).

The number of all cars (including private cars, vans and trucks) entering the Central London toll zone decreased by %27 in between 2002 and 2003. Part of this decrease is due to people preferring public transportation, sharing private vehicles with colleagues (car-pooling), shifting travel hours to night-time, etc. all objectives being along the objective of LCC. However, it was also observed that some drivers circumnavigated central London zone to avoid paying tolls. Daily charge was increased to £8.00 after 4 July 2005 (Leape, 2006). At the end of 2005, the number of cars and traffic delays decreased respectively by %18 and %30. Also, %40-70 reduction was observed in accidental injuries during the year. 70% of London residents believe that the scheme reduces congestion.

The toll zone was extended towards west by 19 km^2 in 2007; however this extension was later withdrawn in 2010. The decision in 2010 was based on a public opinion survey among residents of London.

Before the implementation, there was concern about increase of accidents. However, the number of accidents in the zone and other London areas decreased as evenly after the implementation. Reduction of congestion had a positive effect on environmental conditions. Although environmental impact was difficult to evaluate an analysis of traffic volume to environmental index showed that emissions decreased in the city center (Kearns, 2014).

2.1.2 Congestion Management in Singapore

Singapore is a city state located at the end of the Malay Peninsula. Public transportation was not effective during the 70s in Singapore. People relied on their private cars in a densely populated area hence traffic congestion became a severe problem by the 70s, so the government decided to discourage private car usage. According to the results of two major researches which were done between 1967 and 1974 some restrictions on private vehicles turned out to be inevitable. Short-term aim was to decrease congestion in the central zone of Singapore and the long-term aim was to persuade drivers to change their attitudes towards owning and using private cars.

Singapore Government initially set a goal of reducing the peak-hour traffic by 25 – 30% without impeding the accessibility to the central business district. This would necessitate alternative transportation availability for drivers who chose not entering the central city traffic (Watson & Holland, 1976). Moreover, the charging scheme would require simplicity and ease of implementation for the government and its citizens (Watson & Holland 1976).

Singapore congestion charging system has evolved since its first implementation in 1975 with additional stages. The restricted zone is an area surrounded by a cordon (the central business district which has an area of 5 km^2) and vehicles entering the zone are charged a certain amount. In 1975 the initial stage implementation called the Area Licensing System (ALS) became operational (Yüksel, 1998). The charging zone was determined considering several issues: (i) There should be some diversion routes for drivers, (ii) Entry points should be minimized for simple control, (iii) Additional parking lots should be constructed close to the central business district (Watson & Holland, 1976). Drivers had to buy a paper license which was cost of S$3 (equivalent to 1.3 U.S. dollars) per day or S$60 (equivalent to 26 U.S. dollars) per month (Yüksel, 1998). The daily and monthly charges were determined by trial and error as there was no previous experience (Watson & Holland, 1976).

The ALS was applied between 7.30 AM and 9.30 AM, Mondays through Saturdays (May, 1992). But after that, the tolling interval was revised as 7.30—10.15 AM since the initial implementation caused a severe congestion from 9.30 AM to 10.00 AM. After this revision significant improvement over traffic congestion was achieved (Watson & Holland, 1976).

Manual toll collection was done by police officials (Yüksel, 2007). Alternatively, monthly paper licenses were sold from kiosks and post offices. An inspection was conducted in 28 control points. The penalty of entering any zone without paper license was 23$. The penalty was increased for repeated offence (Yüksel, 1998). There were exemptions for emergency vehicles, motorcycles, taxis and other commercial vehicles (May, 1992).

For drivers an encouragement scheme was started so that any car carrying more than three passengers could pass through a control point toll-free (Xian, 2014), (May, 1992). This so-called carpooling scheme was implemented with the idea of convincing private car owners to share their vehicles with their friends and neighbors and hence reduce the amount of vehicles entering the city. However carpooling was abandoned later, because it was realized that car owners took along public transportation users instead of their fellow car owners, thus the expected reduction in private car usage while entering the restricted zone was not realized. Exemptions for motorcycles, taxis and commercial transport vehicles were lifted in 1989 (Xian, 2014).

Congestion in the charging zone was decreased by 44% during the first seven months of its implementation. However, ALS also had some side-effects. Firstly ALS immediately caused a shift in congestion hours (between 7.00 AM and 7.30 AM, instead of 7.30 AM and 9.30 AM) and places (control points, instead of city centers) (Phang & Toh, 2004). Secondly, although speed of cars increased 20% while entering and exiting the zone, the speed decreased close to control points during the charging intervals. Thirdly, additional costs had to be incurred by drivers due to change in timing and routes of their journeys. Fourthly, some drivers started using their commercial vehicles (while making their private trips) to benefit from toll exemption (May, 1992). Another consequence was congestion shifting to routes circumnavigating the restricted zone caused by drivers trying to bypass the toll checkpoints. (Phang & Toh, 2004).

Parking fees in the restricted zone were increased by almost 100%, and The Park and Ride Scheme was encouraged for drivers going to the central business district (Phang & Toh, 2004). A parking area with a capacity of 15,000 was arranged. The charge for this parking area was S$30 ($13) monthly and regular shuttle services carried the drivers into the central business district (Watson & Holland, 1976).

The traffic in Singapore was continuously monitored for improving the charging scheme: Traffic counts, household interviews, speed/flow measurements, interview with businessmen were made. Observations of pedestrian and parking behavior and pollution data were collected for feedback. These data provided a basis for performance measurement of ALS. Additionally, the accumulated information could be useful for scholars who were interested in applying the congestion management scheme for other major cities of the world (Watson & Holland, 1976).

At the beginning of ALS tolls were implemented only during the morning rush-hours with the expectation of a mirror effect as a remedy for the evening rush-hours. However, congestion during the evening hours occurred (caused by vehicles entering the city center during the evening) so after 1989 tolls were implemented within 4.30-7.00 PM during the weekdays (Xian, 2014).

Weekend Car Scheme was started in 1991. The scheme encouraged drivers to use their own cars within the unrestricted hours. Drivers were able to enter the restricted zone toll-free during the holidays, or the night-time during weekdays (between 7.00 PM and 7.00 AM), and after 3.00 PM on Saturdays. If drivers registered their cars as weekend car, they had benefits such as discount of the annual road tax (a discount of 70%). Weekend Car Scheme was strongly criticized of favoring ownership of luxury cars since annual road tax was primarily based on engine size. When Off-Peak Car Scheme replaced Weekend Car Scheme a fixed reduction of S$800 for the road tax was applied instead of the 70% discount (Phang & Toh, 2004).

In 1994, Whole Day ALS was implemented at 7.30 AM to 6.30 PM from Monday to Friday and at 7.30 AM to 2 PM on Saturdays with a two-tier shoulder pricing system.

The aim of this system was to reduce charge of prior and later peak hours. According to results of Whole Day ALS, traffic decreased at first two-tier restricted hours while the traffic increased at last two-tier restricted hours. Furthermore, the traffic increased in the morning and evening while the traffic decreased in the midday. This means that peak hours could be smoothed by the help of right timing and shoulder pricing (Phang & Toh, 2004).

In 1995, as a part of congestion management plan the East Coast Parkway was made a toll-road (Road Pricing Scheme). The number of cars entering to East Coast Parkway decreased from 12,400 to 7,300 between May and August after this implementation (Phang & Toh, 2004).

Management of Area Licensing System had become complicated with 16 different types of licenses (whole day license, part-day license, daily license, monthly license, etc.) and an increase of population. Existing number of personnel at checkpoints was insufficient which caused slowing down of traffic as well as increased use of counterfeit licenses (Phang & Toh, 2004). So there was a necessity to replace the manual tolling system. Studies of Electronic Road Pricing (ERP) began in 1989. ERP implementation began in 1998 (Xian, 2014).

ERP included three components: (1) In-vehicle Unit (IU) attached to the windshield along with a smart card (cash card) of the car owner, (2) ERP Gantries, (3) Control Centre (Xian, 2014). These components are explained in detail as follows:

(1) In-vehicle Units (IUs) vary for different types of vehicles (attached to front windshields of automobiles or handlebars of motorcycles) (Xian, 2014; Phang & Toh, 2004). IUs are coded with different colors for different types of licenses to prevent confusion. Following an entry into toll-area the payment amount is deducted from the smart card balance. Similarly, when a car enters/exits a parking lot the parking fee is deducted from the car owner with the use of Electronic Parking System (EPS). Drivers can use their smart-cards to ensure that sufficient balance is present for their IUs (Xian, 2014).

(2) There are two gantries for every entering point. First gantry communicates with IU of vehicle, and the license plate is recorded with an optical device. If the vehicle

IU is valid and the balance in the smart card is sufficient toll deduction is made at the second gantry. Passing of vehicles was provided. Any violated entry results with photographing of the license plate of the vehicle to be followed with legal action. Records of violations are stored for 6 months (Xian, 2014). Shortly, while first gantry controlled smart-cards and determined charges, second gantry identified place of cars, type of cars and collected charges (Phang & Toh, 2004).

(3) Control Centre processes all ERP transactions, stores all smart-card deductions and reports violations (invalid IUs, insufficient balance, etc.) and direct maintenance teams whenever problems with the ERP arise (Xian, 2014).

The system was implemented on weekdays from 8.00 AM to 8.00 PM for three cordons; on Saturdays between 12.30 PM to 8.00 PM for two cordons; and between 7.00 AM to 9.30 AM for expressways and arterial roads. Ambulances, police cars and fire vehicles are exempt from the charge (Xian, 2014). ERP is more sophisticated than the ALS in following respects: (1) Charges can be time varying to make smooth transition between tolling intervals, (2) Charges can be based on speed of the vehicles to discourage drivers entering the restricted zone during busy hours, (3) Less expensive than ALS because of reduced need for manual labor (Xian, 2014).

The success of ERP has been outstanding. Despite an increase of 3% increase of the number of vehicles from 1998 to 2012 the traffic volume increased only by 0.8% (Xian, 2014).

2.1.3 Other Implementations

Stockholm

Congestion pricing was implemented in Stockholm for a trial duration of 6 months in 2006. Entry and exit for the city center between 6.30 AM and 6.30 PM was charged per vehicle and favorable results were obtained in terms of traffic volume reduction possibly aided by Stockholm’s very good public transportation system. Congestion charging became permanent in 2007 (Eliasson, 2014).

Hong Kong

In the mid-80s, electronic toll system had been started to be tested in Hong Kong. Although encouraging results were obtained, the trial program had to be stopped because of public opposition. Criticisms of the electronic toll system were also made because of privacy intrusion (Demirtaş, 2009).

United States

The congestion tolling applications in the United States are not cordon type, but tolling through motorways. As an alternative to tolling through motorways and bridges, the high occupancy toll (HOT) project has been implemented first in San Diego in 1996 as express lanes for vehicles paying a certain amount of toll (Buckeye, 2014).

There are other cities throughout the world with successful implementations such as Bergen, Oslo (May, 1992) and Milan (Moroni, 2014). It is anticipated that many metropolitan city councils will impose electronic congestion charging systems in the near future.

2.2 A Review of the Literature

In Section 2.1, a summary of congestion pricing applications in certain cities and metropolitan areas have been given. This section includes an overview of the related studies available in the academic literature. Congestion management problems have gained popularity among academic scholars as the increasing number of vehicles caused traffic problems in city centers. Academic articles can be distinguished primarily as analytical models versus practical studies. Analytical papers are based on those including mathematical models and numerical tools for solving the congestion problem. Practical papers are heavily based on a simulator program which can implement what-if scenarios to determine performance statistics.

Analytical models have been developed as early as the 60s and have become popular since the 70s and 80s. An initial work based on the supply-demand dynamics and costs of traffic congestion has been studied by Walters (1961). The single bottleneck model where the effects of time-varying tolls have been analyzed is due to Vickrey

(1969). The study of Vickrey is a basic approach for single bottleneck model for identical commuters on a route. Vickrey's model serves as a basis to many analytical models developed afterwards. Equilibrium of supply-demand dynamics when the system is not at steady state is studied by Hurdle (1981). The work of Hurdle aims that the basic supply-demand function is improved in consideration of the changing situations based on the time of day. An alternative approach considering driver interactions is developed by Else (1981) whereas in the conventional approach, economic analysis is between the traffic flow and the cost of using a route, but it is claimed that the relationship interests in between the costs and the number of vehicles on the road. The study of Smith (1984) focuses on single bottleneck on a route where it is assumed that every commuter has only one appropriate route and all the routes pass through a single bottleneck. The only decision made by the drivers is the departure time for work. The articles mentioned so far assume identical commuters which is a restrictive assumption. Cohen (1987) has analyzed the single bottleneck problem for non-identical commuters. As a result of this study, while commuters with higher income are favored by the tolling system, commuters with lower income incur losses. Analyses of the traffic flow during peak hours have been made by Hendrickson and Kocur (1981). It is concluded that the increase of road capacity or service rate do not remove the congestion during rush hour periods as well as the study of Arnott and Small (1994). The results of these works indicate that the increase of capacity or service rate cannot totally eliminate the queue waiting time in traffic. Although increased capacity reduces the peak values and delays, free flow traffic cannot be ensured. Ben-Akiva et al. (1986) and Braid (1996) analyze case of alternative parallel roads except only one road. Also, the study of Ben-Akiva et al. (1986) aims to improve a model which analyzes the effects of alternative precautions to eliminate the congestion during peak periods and predict the volume of traffic in the bottleneck and time distribution of delay.

Chen et al. (2015) have studied multi-step tolls for non-identical commuters in the bottleneck model. The model is constituted as a mathematical program. Step tolls are approximation strategies for the time-varying toll suggested initially in Vickrey's model (Lindsey et al. 2012). Arnott et al. (1994) have researched the effects of toll implementation for the non-identical commuters. Drivers may have different costs of travel, different work hours, and also differing preferences for early or late arrival to work. Unlike previous studies, Akamatsu et al. (2015) have analyzed the case of a single commuter’s freeway with multiple entry points and multiple bottlenecks along the route.

Zhang et al. (2005) have analyzed examples of traffic stabilization via elastic working hours. The single bottleneck model by Xin and Levinson (2015) is a stochastic queuing problem, among articles with deterministic models. Danielis & Marcucci (2002) and An & Zhang (2012) have considered models for which the commuters can use choices of alternative transportation instead of private cars. De Palma et al. (2017) has examined the departure time choice for the commuters who travel to work by train or metro which is not unlike the departure time choice of drivers in a congested traffic.

Some of the academic papers are based on the implementation of simulation programs. There are real city network applications as well as hypothetical urban network simulations to derive results. An artificially created network is used by de Palma et al. (2005) for which alternative tolling approaches is compared by using the simulation tool METROPOLIS. Comparison of results of alternative tolling approaches indicates that step tolls are better than flat tolls. The claim of the authors of this paper is that a simulation tool for road pricing is much more capable than analytical models and procedures. The claim makes sense given that analytical models are usually simple and limited whereas simulation programs allow the users to test with various scenarios of congestion management. Kristofferson and Engelson (2009) have studied the Stockholm city network in a project called SILVESTER (SimuLation of choice between Starting TimEs and Routes) and CONTRAM has been developed as the route choice model. Other better known simulation programs CORSIM and VISSIM have been compared by Bloomberg and Dale (2000). The SMARTEST-project which is developed as an open-source by Krajzewicz and Rössel is explained in Krajzewicz et al. (2002).

Having provided an overview of the academic studies we will outline our congestion pricing model for urban arterial networks which has been developed as an extension of the single bottleneck approach.

CHAPTER 3

AN EXTENSION OF THE SINGLE BOTTLENECK APPROACH FOR CONGESTION MANAGEMENT IN URBAN ARTERIAL NETWORKS

In this chapter an introductory analysis will be provided for the Vickrey's single bottleneck model to be followed by a revised interpretation of this approach. After the simulation runs for the single bottleneck approach we will outline how to make use of this approach for simple city networks to develop average waiting times during congested traffic hours. The chapter will be concluded by an analytical derivation of the average waiting time of the single bottleneck problem along with a paradox of its computed value. We start our discussion with a conventional understanding of the traffic congestion problem from the economist's perspective: supply and demand.

3.1 A Conventional Supply Demand Analysis of Congestion Tolls

The conventional supply demand structure for traffic congestion on a single route has been based on the analysis of Walters (1961) and illustrated in Figure 1. For simplicity it is assumed that all vehicles are identical so the cost of a driver passing through this route is constant as long as there is free flow traffic. However, as the number of vehicles increases at some point the route capacity is exceeded and after that the cost incurred by the driver will start to increase (curve outlined as C_1). Meanwhile, an increase in traffic congestion (hence time lost) will cause some of the drivers to seek alternative means of transportation which means a decrease of demand to use that route (curve outlined as D). Therefore, the supply-demand equilibrium realizes at point E which means drivers incur a larger cost for crossing through this route and enduring some amount of traffic congestion while doing so.

Taking into consideration the relevant marginal social costs induced by the traffic (as well as the private costs incurred by the drivers) and supposing that to reduce congestion through this route (most possibly only during certain hours of the day) a toll is implemented. The toll has a discouraging effect (having reduced the traffic flow on the route from F_E to F_S) so that the new equilibrium realizes at S with a toll amount equivalent to the magnitude of the line segment between points S and R (Else, 1981; Hurdle, 1981). An alternative interpretation is that some drivers are willing to pay tolls to reduce the valuable time lost due to traffic congestion (Hendrickson & Kocur 1981).

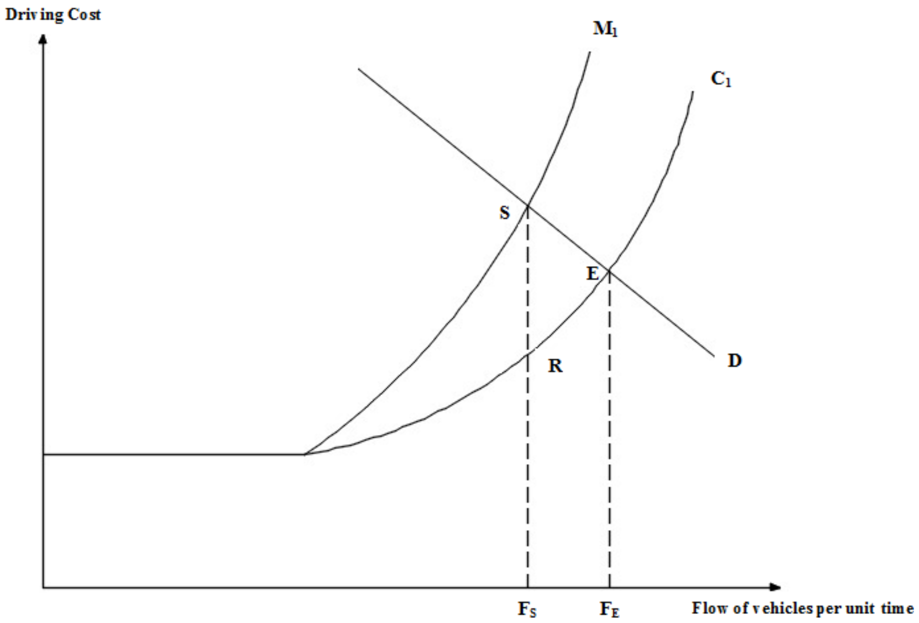

Figure 1 Driving cost of a single vehicle through the route versus traffic flow through the route as vehicles per unit time (graph adapted from Else (1981))

3.2 Vickrey’s Single Bottleneck Model

Vickrey’s single bottleneck model is a deterministic congestion model with uniform traffic flow through a road with one entry and one exit (Cohen, 1987). Despite the simplicity of the model, one can visualize that the single bottleneck along a route is a problem faced in many cities and towns, most typically the commuter traffic joining a residential area and the city center. Drivers go to their workplaces during peak period of the morning by using a specific route which includes a single bottleneck (see Figure 2). The primary assumption is that any driver will choose his/her departure time from home to minimize total travel cost. In the absence of any traffic congestion owing to its being a deterministic model the travel time to work is fixed for each driver so that he/she would leave home exactly at the time which can be computed as the desired arrival time minus fixed travel time assuming free flow traffic along the route. However, if the route capacity (for simplicity assuming at a certain bottleneck location) is exceeded during morning rush-hours, a queue builds up at this bottleneck location (Vickrey, 1969; Cohen, 1987). Therefore, in the absence of congested traffic the travel cost of a driver is the fixed cost of his/her journey (for this instance, the costs incurred to drive the vehicle from home to work for a fixed duration of time). On the other hand, if a commuter comes to a bottleneck in his/her morning route, the commuter causes a social cost-apart from the indicated private cost because of the undesirable interaction with the other drivers and the resulting traffic congestion. Such a social cost would not be the case had there been uninterrupted free flow traffic through the route (Cohen, 1987).

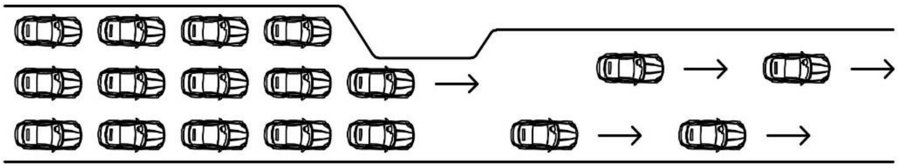

Figure 2 Sketch of a bottleneck on a commuters' route during rush hour traffic causing the formation of a queue

The formation of the bottleneck will induce the driver to make a choice: Either to leave home early to be at the workplace on time, or to leave home at a suitable time as if no bottleneck is present and pass through the congested traffic and arrive at the workplace late. Supposing that the morning demand of traffic along the route is 120 cars per minute and at the bottleneck point the capacity (the supply) allows only 100 cars per minute, then the 7200 cars are unable to cross the route in 60 minutes (actually a duration of 72 minutes is necessary). Drivers who arrive at their workplaces early face less congestion in start of the peak period. However, they arrive at their jobs early which are undesirable for them. Commuters travelling in peak times of the congestion pay either low cost or nothing, but they are faced with much longer travel times. Moreover, they very possibly arrive at their jobs late which are also very undesirable.

A deterministic model to represent the bottleneck on a single route and the preferences of individual drivers (assuming that each individual has the same preference) was developed by economist William Vickrey. This model tries to discourage early arrivals to workplace but at the same time late arrivals are obviously very undesirable as well. For this purpose payoffs are defined in the following manner:

$\boldsymbol{w_h}$: Payoff for time spent at home (received per minute)

$\boldsymbol{w_j}$: Payoff for time spent in office during work hours (received per minute), say between 9.30 AM and 5.30 PM (late arrival payoff)

$\boldsymbol{w_q}$: Payoff for time spent queuing (received per minute) at a traffic bottleneck

$\boldsymbol{w_p}$: Payoff for time spent in office outside work hours (received per minute), for example before 09.30 AM (early arrival payoff)

These payoffs are reasonable if the following condition is satisfied:

$$\boldsymbol{w_j > w_h > w_p > w_q = 0} \tag{1}$$

The equilibrium for this model results with a linearly increasing waiting time at the queue caused by the bottleneck on the route until a certain time point between 8.00 AM and 9.00 AM. The queue waiting time starts to linearly decrease until it diminishes meaning an end of the congestion and the start of the free flow traffic along the route. The maximum waiting time in the queue and the timing position of this peak is identified by (i) the bottleneck capacity; (ii) cost parameters defined above (see Figure 3). The calculation of the queue buildup and decline rates as well as toll values based on these payoffs are given by Vickrey (1969) and will not be explained in detail as we consider a single bottleneck problem disregarding the payoffs representing the individual preferences of spending time for early arrivals, late arrivals, and spending time in traffic in Section 3.3.

As stated earlier some drivers will leave home earlier than the case with no traffic congestion. For those, it means an additional cost because of loss of valuable time in traffic (waiting in the bottleneck queue) as well as early arrival at workplace (the payoff for early arrival being less than the payoff at home). Some drivers might prefer to travel after their desired times so that they try to avoid a long waiting time in the queue and an undesired early arrival at work. In the single bottleneck model, there are some people who arrive at their workplaces earlier or later than their co-workers.

The assumption of the model is that the additional costs incurred by the drivers who leave home early are willing to pay a toll amount equivalent to this cost. Since the additional cost amount depends on the crossing time, the toll should be variable by time (and no tolls should be implemented when there is no congestion at all). Tolling is supposed to force some car owners to use public transportation, or commuters to share their cars and some drivers to change their travel times and routes as well. Therefore, implementing a time-varying toll for the bottleneck on the route is a way to reduce congestion (Cohen, 1987).

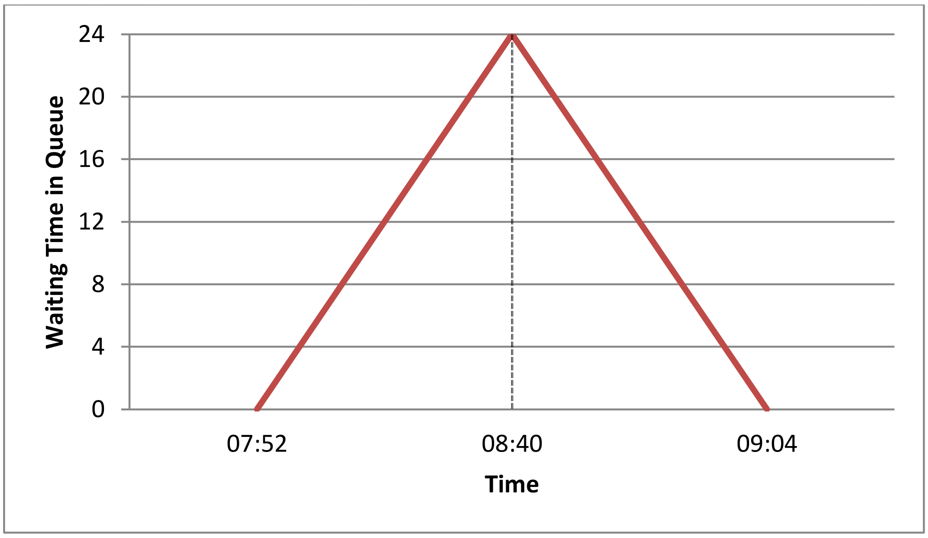

Figure 3 Waiting time observed by drivers arriving at the queue for the single bottleneck model of Vickrey (1969). The graph represents discrete arrivals of 7200 cars supposed to cross the bottleneck zone in 60 minutes (within 8.00 - 9.00 AM), however all of the cars can cross in a duration of 72 minutes.

As also seen from the graph in Figure 3, outside the rush-hours there is free flow of traffic through the bottleneck point (traffic flow is below bottleneck capacity). At a certain time in the morning (7.52 AM) congestion starts at the bottleneck and cars will be waiting for a while in queue until they pass the bottleneck and continue their journey in free flow. The wait in the queue time increases linearly as time passes. For this particular example the maximum waiting time in the queue is 24 minutes and is experienced by the driver who leaves the bottleneck at 8.40 AM. Right after this point the waiting time reduces at a linear rate and starting with 9.04 AM there is no longer any bottleneck and free flow prevails again.

If drivers pass from the bottleneck before peak time, they arrive at their offices early. Because of the bottleneck, these drivers have to leave their home early, spend time in queue and also incur an early arrival cost in their office. In addition, a driver who crosses bottleneck before 8.40 AM (two-thirds of drivers falling in this category) has to pay cost for waiting in the queue because payoff for time spent waiting to cross the bottleneck is zero whereas time spent in office or at home have positive payoffs. In case of congestion tolls this also has to be added to total travel cost (Vickrey, 1969).

If there was no congesting bottleneck through this route drivers would not be paying the additional cost items (queuing cost, early or late arrival in the office cost) because they would leave their home at the latest minute to arrive on time for their work in free flow traffic. It is not unreasonable to say that drivers will be willing to pay tolls to reduce the traffic congestion. The implementation of the toll is expected to reduce the arrival rate (120 cars between 8.00 and 9.00 AM) because some drivers will switch to public transportation and some others will start sharing their cars with their colleagues. Introduction of toll shifts the supply-demand equilibrium towards left when drivers pay tolls equivalent to the cost of queuing. In this way, the flow rate during the rush-hours will be reduced (see Figure 3).

3.3 A revision of the single bottleneck model

The single bottleneck model which is deterministic model has one direction route. As explained in the Section 3.2 Vickrey’s model assumes that some drivers prefer to leave home early which takes back the queue formation to 07:52 instead of 08:00. Since the intention in this work is not to evaluate an economical supply-demand analysis we will assume that the queue formation begins at the beginning of the 60 minutes interval and the maximum queue length is reached at the 60th minute. The queue starts to dissolve when the arrival rate falls below the departure rate at that instant.

Similar to Vickrey's model there is a single highway joining residential areas with the business center of a town and for a bottleneck zone on this route for a fixed time interval when the arrival rate exceeds the departure rate is considered. We will consider only morning time congestion for the bottleneck (evening time congestion will take place in the reverse direction on the same route). The free flow traffic continues until 7.30 AM and after that, the congestion will occur for a duration exceeding 60 minutes. The arrival rate is increased suddenly at 7.30 AM causing the start of queue buildup in the bottleneck zone. The queue length reaches a maximum value at 8.30 AM and after that point the arrival rate decreases suddenly to a value below the departure rate so that dissolving of the queue starts. This assumption although not very appropriate in real life will simplify the computation of the average waiting time at the end of each simulation run.

We have assumed a linear rate of increasing and decreasing queue lengths during the queue buildup and queue dissolving phases respectively similar to that in the single bottleneck model described in Section 3.2 (see Figure 4). The queue buildup phase (called phase-1, time interval between 7.30 AM and 8.30 AM) has an arrival rate of $\boldsymbol{a_1}$ and a departure rate of $\boldsymbol{d}$. The queue dissolving phase (called phase-2) starts at 8.30 AM with arrival rate $\boldsymbol{a_2}$ whereas the departure rate remains unchanged as $\boldsymbol{d}$. In order the guarantee queue formation in phase-1 and queue dissolving in phase-2 the following condition should be satisfied:

$$\boldsymbol{a_1 > d > a_2} \qquad (2)$$

The duration of phase-1 is fixed as 60 minutes, but the duration of phase-2 depends on the value of $\boldsymbol{a_2}$. In most of the simulation runs we have chosen $\boldsymbol{a_2}$ value equivalent to 80% or 20% of the departure rate. We investigate the effect of various choices of the $\boldsymbol{a_2}$ value in section 3.3.3.

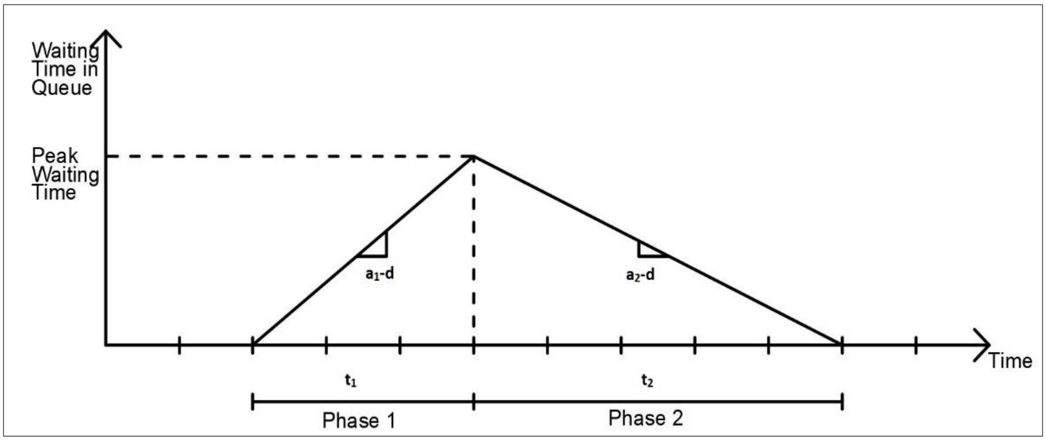

Figure 4 Waiting time of vehicles given the arrival time to bottleneck zone during the rush-hour traffic. The graph is the continuous version of what should actually display discrete arrivals. For $\boldsymbol{a_1}$ = 80 vehicles per minute, $\boldsymbol{d}$ = 60 vehicles per minute, $\boldsymbol{a_2}$ = 48 vehicles per minute arrivals occur every 45 tertias in phase-1 and every 75 tertias in phase-2.

Arrivals and departures happen at discrete time instants, so we need finer intervals than seconds. For this reason, we have used one-sixtieth of a second as tertia. Also, the arrival and departure rate is calculated in terms of drivers per minute to avoid computations with many unnecessary digits. During phase-1 the queue length faced by each arriving car as well as the queue waiting time increase linearly, and similarly during phase-2 these performance statistics decrease linearly (most possibly the increase and decrease rates are different).

Consider the following example; let $\boldsymbol{a_1 = 80}$ cars per minute (cpm) after 7.30 AM. New car arrives at the bottleneck zone for every 45 tertias assuming uniform arrivals. Also, (uniform departure rate) $\boldsymbol{d = 60}$ cpm exits from the bottleneck for every 60 tertias. First car arriving after 7.30 AM waits 15 tertias, second car waits 30 tertias, etc. The simulation program generates a list of arriving cars. For convenience we have printed these values on a minutely base (see Appendix-1). After one minute has passed in the first interval the 80th car arrives at the bottleneck zone, sees a queue length of 20 cars and waits in the queue for 1200 tertias (0.33 minutes). After two minutes have passed the 160th car arrives, sees a queue length of 40 cars and waits 2400 tertias (0.67 minutes). The 4800th car arrives at 8.30 AM and waits 72000 tertias (20 minutes). For this example based on the simulation run the average waiting time of cars having arrived at the bottleneck during phase-1 is 36000 tertias (10 minutes). It is assumed that the arrival rate is decreased suddenly in phase-2 to $\boldsymbol{a_2 = 48}$ cpm (equivalent to 80% of d). The queue length starts to decrease after 8.30 AM. After 8.30 AM arrivals take place every 75 tertias at the bottleneck zone and the departures take place every 60 tertias (similar to phase-1). The 9598th car waits 30 tertias in the queue while the 9599th car waits 15 tertias in the queue. Eventually, the 9600th car which arrives at the bottleneck (at 10.10 AM) sees no queue; hence the waiting time is zero after that. There will be free flow traffic until the next morning at 7.00 AM.

Simulation results for the single bottleneck problem are provided in Appendices 2, 3, and 4 for various arrival and departure parameters. The minute based list is enough for displaying the linear increase and decrease in queue length faced with each car.

3.3.1 A Simple Network

We will assume that the congested networks which we study fulfill flow conservation based on a pipe network analogy (Daganzo & Garcia, 2000) in the sense that whatever comes in a traffic junction should go out without loss of water. Such cases are typical for zones of transition where buildings, facilities, residence places are not available, but the congestion network serves as a junction of intersecting highways where congestion might occur.

We also assume deterministic flow values through unidirectional routes (for cases of bidirectional one might consider parallel routes with flow in the other direction). Similar to the single bottleneck problem described earlier we will consider fixed rush-hour duration so that congestion starts at the beginning of this duration at several bottleneck zones and the bottlenecks gradually disappear at the end of this duration.

For convenience we will provide a network with initial flow values (initial meaning the flow values which would be valid if the problem had no capacity restrictions, hence no bottleneck zones and no congestion). The initial flow values are represented as number of cars crossing through a zone per minute. Each such defined zone has a serving capacity and during the rush-hours some of these zones have flow values exceeding the capacities resulting with bottlenecks and queue formation. This is similar to the situation of the single bottleneck problem occurring simultaneously at multiple zones. However, unlike the single bottleneck case each bottleneck has effects on the rest of the network as will be outlined. A sample network is given in Figure 5.

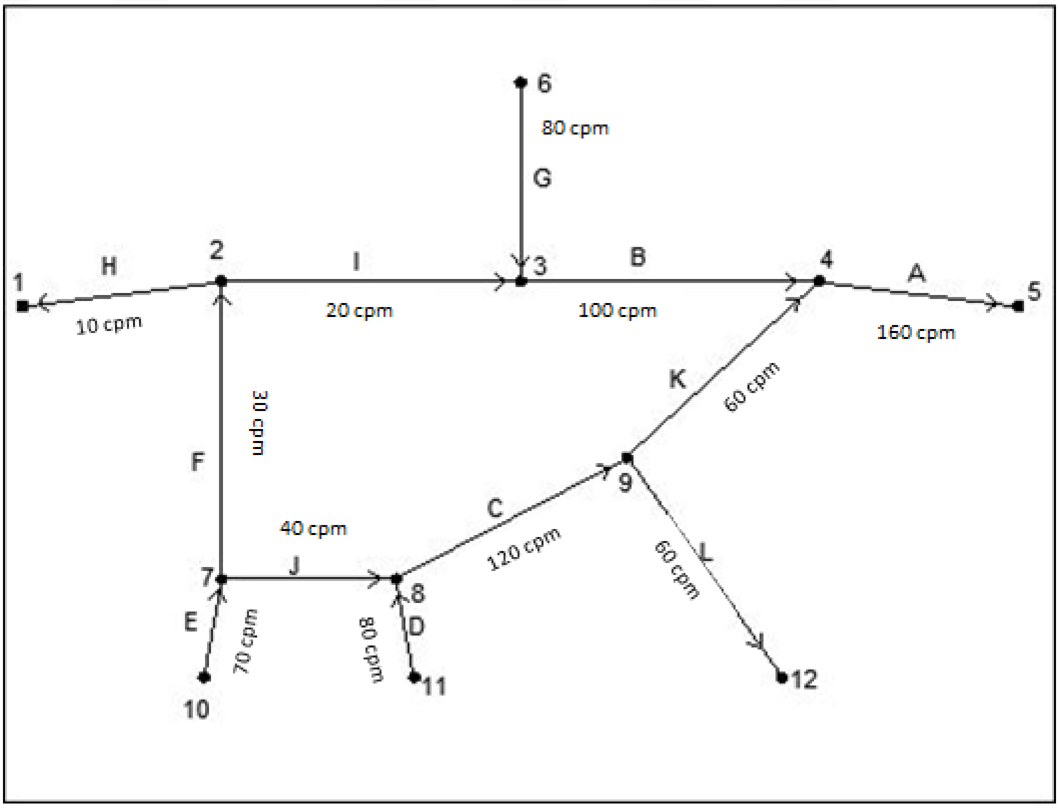

Figure 5 The initial flows diagram of a simple traffic network (possibly a hub where several freeway routes meet). Flow rates are expressed in cars per minute (cpm). Node-demand values are given in Table 1.

Table 1 Node-demand values

Node	**Direction(Link)**	**Demand (in cpm)**
1	-	-
2	1 (H)	10
2	2 (I)	20
3	1 (B)	100
4	1 (A)	160
5	-	-
6	1 (G)	80
7	1 (F)	30
7	2 (J)	40
8	1 (C)	120
9	1 (K)	60
9	2 (L)	60
10	1 (E)	70
11	1 (D)	80
12	-	-

For this network the flow conservation equations will be:

Flow (A) = Flow (B) + Flow (K) (3a)

Flow (C) = Flow (K) + Flow (L) (3b)

Flow (C) = Flow (J) + Flow (D) (3c)

Flow (B) = Flow (G) + Flow (I) (3d)

Flow (F) = Flow (H) + Flow (I) (3e)

Flow (E) = Flow (F) + Flow (J) (3f)

The approach in this section will be to make use of the simulation tool developed for the single bottleneck model. As stated the bottleneck zones in a network are not independent of each other. Therefore, we should compute the equilibrium flow values first, and then analyze each bottleneck separately. To determine the equilibrium flow values we will employ the processing algorithm summarized in Figure 6.

Initialization	• Assign all effective incoming and outgoing flow values to zero. Choose the links which do not have any successors and put them in the 'To be processed list' (TBP list).
Main iteration	• For the next link in the TBP list update the effective incoming flow and effective outgoing flow values based on initial flow values (try to assign the initial flow values if possible). If eventually, effective incoming flow value is equal to effective outgoing value then there will be free traffic flow through that link. Otherwise, there will be a bottleneck zone through that link. • For each processed link, choose the links which follow that link and place them in the TBP list to be considered for the next iteration. • When all links in the TBP link are processed, move on to the next iteration with the next TBP link.

Termination	• The algorithm is terminated either (i) When it is not possible to place any link in the TBP list for the next iteration (no-cycle case), OR (ii) When it is not possible to place any link in the TBP list for the next iteration and/or the updated effective incoming flow and effective outgoing flow values no longer change in successive iterations (cycled case).

Figure 6 The algorithm to compute the equilibrium flow values (to be used for no-toll and tolled scenarios)

The initial flows of the first example network are illustrated in Figure 5. Given this network with possible bottlenecking zones one should compute the effective flows and the resulting no-toll equilibrium as will be explained later. Due to congestion zones the initial flows are not necessarily the realized flow values during the rush hours.

The effect of tolling would shift the supply-demand intersection by discouraging more drivers to shift to alternative means of transportation (municipal buses, car-pooling with co-workers, etc.) causing a reduction of the flow values. A clever approach would be to test alternative tolling scenarios with the aim of removing the bottlenecks. About tolling certain zones of this network we can make some reasonable assumptions:

- Tolls are implemented only through routes (tolling is not cordon-type or zonal).
- Congestion at some zones should be local; any congestion zone through a route should not cause additional queuing at an adjacent route or congestion zone (in real life, this assumption may not be valid).
- If toll is implemented through some route then it will reduce the flow. Some simple reduction rules can be applied:

    - If toll is implemented through some route then it will affect all downward flows emanating from that route. We will call this a-type effect.
    - If toll is implemented through some route then it will affect all upward flows ending at that route. We will call this b-type effect.
    - If toll is implemented through some route additional effects will be observed due to secondary and third-order relations. We will call this c-type effect.
- Tolls are imposed at several points during only rush-hours. It is logical to assume that tolls will reduce the flow at these points and these reductions will affect the flow values in the downward direction. Hence, the flow values should be revised. Some bottlenecks can be eased while some bottlenecks might even disappear. This will be the new toll-equilibrium. There will be alternative toll equilibriums for different toll scenarios.

We will consider only a-type effects of tolling at present following the analogy of a network of river flows and reservoir areas (meaning bottleneck zones). b-type and c-type effects are considered too complicated for determining the equilibrium flows.

For this network example, there are 12 linking routes where flows were measured. If there are capacity values which are less than the given flow values (for at least one route), then traffic congestion is observed. The constituted network has three exit points and three entry points. The values given in Figure 5 are not the actual flow values during rush hours because of capacity limitations of some of the links. These flow values represent the demand-side. When the supply-side is considered, the realized values should be calculated because of the capacity limitations. If there were no capacity limitations at any link during the rush-hour period, the flow values in this figure would be the actual values. Since we only consider a-type effects, the bottleneck at zone C should not cause congestion at zones D or E.

When it is examined on the network, the number of cars passing through link E is 70 cpm. The road is divided into two routes: flow rate of link F is 30 cpm and flow rate of link J is 40 cpm to fulfill flow conservation. This means that if the flow through link E is reduced, the flow values through F and J will be reduced proportionately. All the initial flow values in Figure 5 are generated to fulfill the flow conservation equations in eq-3. From three entry links E, D, and G there is a total incoming flow of 230 cpm and from three exit links A, H and L there is a total outgoing flow of 230 cpm. Assume that due to capacity restrictions through links A, C, and G bottleneck queue formations will take place at these zones. Let link A has a capacity of 120 cpm, and similarly links C and G have respective capacity values of 100 and 60 cpm. The question should be asked: given the bottlenecking traffic at zones C and G will there be an incoming flow equivalent to 160 cpm at zone A, the intuitive answer would be no but let us proceed a step by step evaluation.

The No-Toll Equilibrium

The effective flow values should be computed step by step in the downstream direction to arrive at the no-toll equilibrium. For example, 120 cars arrive at link C per minute and its capacity is 100 cpm. Due to this capacity limitation, the incoming cpm to link C is 120, and the outgoing cpm from link C will be 100. As a result the effective incoming value for links K and L will be 50 cpm instead of 60 cpm (the revised incoming values preserve initial flow proportionality). Similarly, through link G incoming flow rate is 80 cpm and outgoing flow rate is 60 cpm due to capacity limitation. Therefore, the effective incoming flow rate through link B will be 80 cpm (instead of 100 cpm) to fulfill flow equation “Flow (B) = Flow (G) + Flow (I)”.

Due to revised outgoing flow values through links B and K, the effective incoming flow value through link A becomes 130 cpm to fulfill flow conservation (meaning that an incoming flow rate of 160 cpm does not reach to link A due to bottlenecks at zones C and G). Due to the capacity limitation of link A the outgoing flow rate will be 120 cpm. As a result there will be three bottleneck zones in this network during morning rush hours: zone C (incoming flow 120, outgoing flow 100 cpm), zone G (incoming flow 80, outgoing flow 60 cpm), zone A (incoming flow 130, outgoing flow 120 cpm). Note that the effective incoming flow rate for link A is different than the initial flow value given in Figure 5. For the no-toll equilibrium, the total incoming flow from three entry links E, D, and G is 230 cpm. Total outgoing flow through three exit links A, H and L is:

120+10+50 = 180 cpm

The difference of this quantity from 230 cpm is because we have three bottleneck zones C, G and A:

(120-100) + (80-60) + (130-120) = 50 cpm

Due to three separate queue formations there is this flow imbalance which will be relieved when the morning rush hours are over.

The no-toll equilibrium is given in Figure 7.

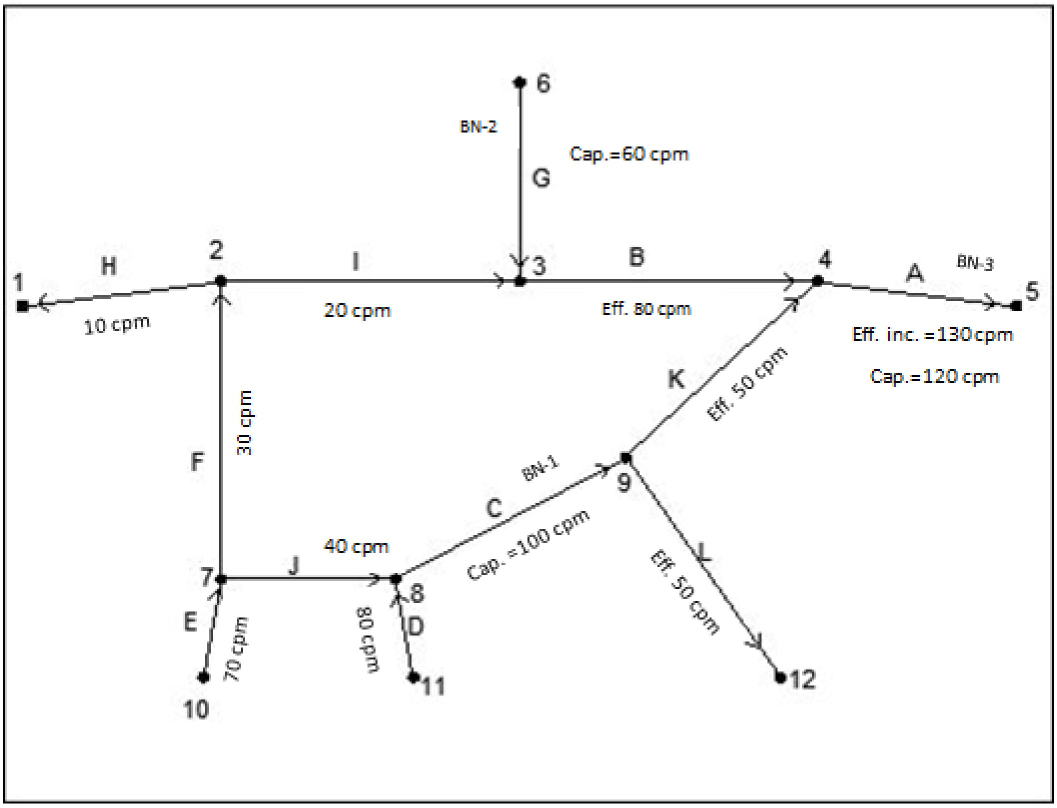

Figure 7 No-toll equilibrium for a network with three bottlenecks

Now that the equilibrium flow values are available we can run the single-bottleneck simulation tool for the bottleneck zones C, G, and A independently to derive the average waiting times for cars during morning rush hours. The summarized results are given in Table 1. Capacity values are given for links only with binding constraints, hence causing the formation of bottleneck queues.

Table 2 No-Toll Equilibirum

Zone (Link)	**Succ. of**	**Foll. by**	**Initial Flow**	**Capacity**	**Eff. Inc. Flow**	**Eff. Outg. Flow**	**Average Waiting Time in Queue (in minutes)**
A	B, K	-	160	120	130	120	2.50
B	G, I	A	100		80	80	0.00
C	J, D	K, L	120	100	120	100	6.00
D	-	C	80		80	80	0.00
E	-	F, J	70		70	70	0.00
F	E	H, I	30		30	30	0.00
G	-	B	80	60	80	60	10.00
H	F	-	10		10	10	0.00
I	F	B	20		20	20	0.00
J	E	C	40		40	40	0.00
K	C	A	60		50	50	0.00
L	C	-	60		50	50	0.00

It should be noted that once the equilibrium flows are computed for the rush-hours (the effective incoming and outgoing flows) we can make use of the single bottleneck simulation tool described in Section 3.3 for any link where queuing occurs during the rush-hours. In Table 1 links having equal effective incoming and outgoing flows have no bottleneck queues hence zero average waiting times. Wherever the effective incoming flow is larger than the effective outgoing flow, a bottleneck queue is observed through that link. For such a case the effective incoming flow is equivalent to $\boldsymbol{a_1}$ and the effective outgoing flow is equivalent to $\boldsymbol{d}$ discussed in

Section 3.3. In the summarized results we have not mentioned $\boldsymbol{a_2}$, the arrival rate for second phase (which can be any value less than $\boldsymbol{d}$), because in Section 3.3.3 it will be shown that the value of average waiting time is independent of the value of $\boldsymbol{a_2}$. Duration of phase-1 is 60 minutes ($\boldsymbol{t_1}$in Section 3.3), duration of phase-2 ($\boldsymbol{t_2}$ in Section 3.3.) depends on the value of $\boldsymbol{a_2}$.

Alternative Tolling Scenarios

As stated previously, we will simply assume that tolling will shift some drivers to alternative transportation methods (in the following section we will consider shifting to alternative routes by some drivers) and hence a certain reduction on the demand side will take place. Since we consider only a-type effects consider tolling of the entry links E, D, and G for the network given in Figure 5. Assume that these flow rates are reduced to 60, 70 and 70 cpm respectively instead of 70, 80 and 80 cpm.

Tolling at E reduces the flow rate to 60 cpm and consequently the flow rate through F becomes 26 cpm and the flow rate through J becomes 34 cpm. Here, we distribute the outgoing flow from E preserving the original proportion of flows among F and J. Similarly tolling the entry from D, results with 70 cpm. The 34 cpm outgoing flow from J and the 70 cpm passing through D arrive at link C as 104 cpm based on flow equation Flow (C) = Flow (D) + Flow (J). However, the outgoing flow at link C is 100 cpm because of the capacity limitation. In the same way, the flow of C is equivalent to the total flow of L and the flow of K (Flow (C) = Flow (L) + Flow (K)), so 100 cpm outgoing from C is divided into two routes as K and L. As a result of tolls in E and D links, the effective value is 50 cpm in link K and the effective value is 50 cpm in link L. On the other hand, the outgoing flow is 26 cpm at F and consequently the flow rate through H becomes 10 cpm and the flow rate through I becomes 17 cpm (because of flow equation Flow (F) = Flow (H) + Flow (I)). Tolling through G reduces the flow rate to 70 cpm, but 60 cpm passing through link G because of the capacity limitation of link G. 77 cpm pass through link B when 17 cpm in I and 60 cpm in G connected in link B (according to flow equation Flow (B) = Flow (G) + Flow (I)). Eventually, 127 cpm instead of 130 cpm entry to link A when the flow rate of B and the flow rate of K are connected at link A. However, 120 cpm exit from link A in the network because of the capacity limitation at link A. After updating of all links, the toll-equilibrium is reached in the network (see Figure 8).

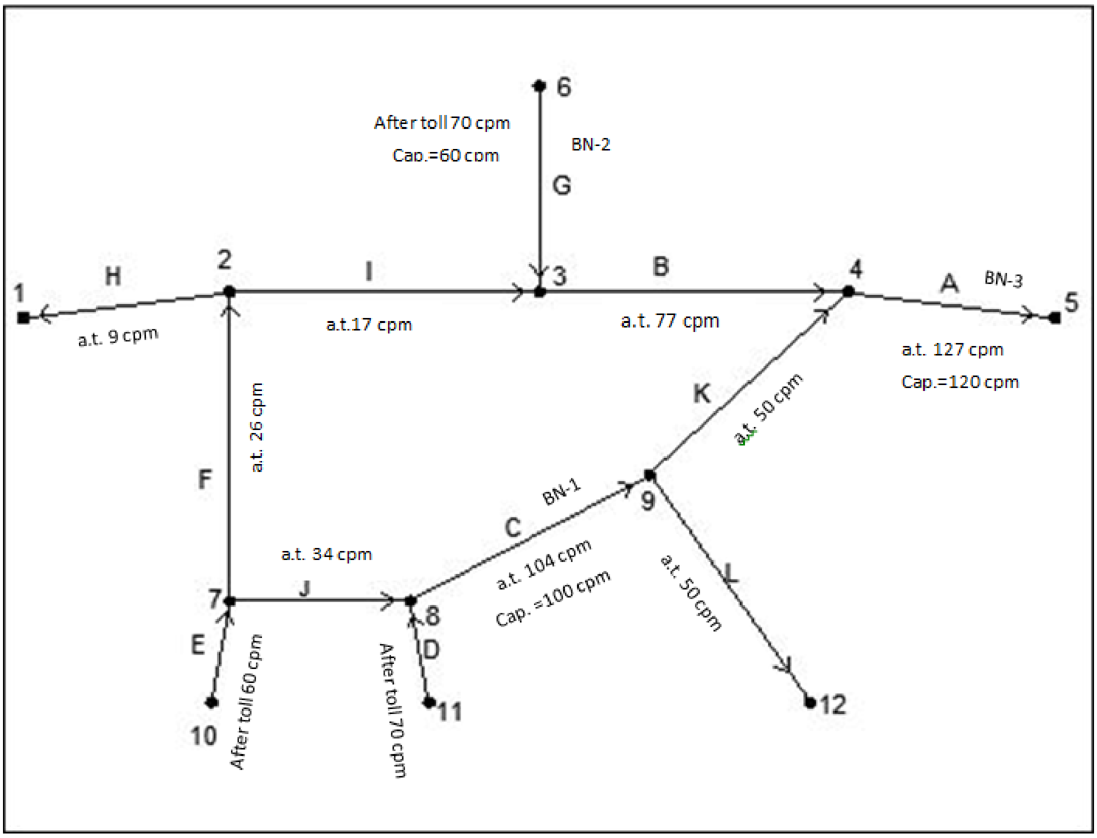

Figure 8 Equilibrium flows for Toll Scenario-1

After toll scenario-1, average queue waiting time in zone A is decreased to 1.75 minutes instead of 2.50 min while average queue waiting time are reduced to 1.00 min and 2.50 min respectively by zone C and zone G (obtained with updated simulation runs). Summary results are shown in Table 2.

Table 3 Toll Scenario-1 Equilibrium

Zone (Link)	Succ. of	Foll. by	Initial Flow	Capacity	Initial Flow After Toll	Eff. Inc. Flow	Eff. Outg. Flow	Average Waiting Time in Queue (in minutes)
A	B, K	-	160	120		127	120	1.75
B	G, I	A	100			77	77	0.00
C	J, D	K, L	120	100		104	100	1.00
D	-	C	80		70	70	70	0.00
E	-	F, J	70		60	60	60	0.00
F	E	H, I	30			26	26	0.00
G	-	B	80	60	70	70	60	2.50
H	F	-	10			9	9	0.00
I	F	B	20			17	17	0.00
J	E	C	40			34	34	0.00
K	C	A	60			50	50	0.00
L	C	-	60			50	50	0.00

According to scenario 2, tolls are implemented through the same links E, D and G. Assume that with higher toll fees incoming flow values are reduced to 50, 60, and 60 cpm respectively. Since outgoing flow through link E is decreased to 50 cpm, 30 cpm will pass through link J and 20 vehicles per minute will pass through link F. Since outgoing flow through link D is decreased to 60 cpm, the incoming flow at link C will become 90 cpm (because of equation Flow (C) = Flow (D) + Flow (J)). The queue at link C disappears completely because the traffic flow (90 cpm) is below to the capacity limitation (100 cpm), meaning that incoming flow through C equals outgoing flow. Consequently, revised flow values through links K and L will be 45 cpm. 7 cpm exit from the network by using link H, while 13 cpm pass through link I. Since incoming flow to link G is equal to its capacity, 60 cpm will pass through link G. Therefore, the queue at zone G is eliminated completely. As a result of revised flow through links F and G, 73 cpm pass through link B instead of 80 cpm. Consequently, 118 cpm will pass through link A. The queue at zone A disappears because the arrival rate is below to capacity limitation of link A.

Tolling according to scenario-2 has eliminated all bottlenecks and there will be free flow traffic during the morning rush-hours (see Figure 9).

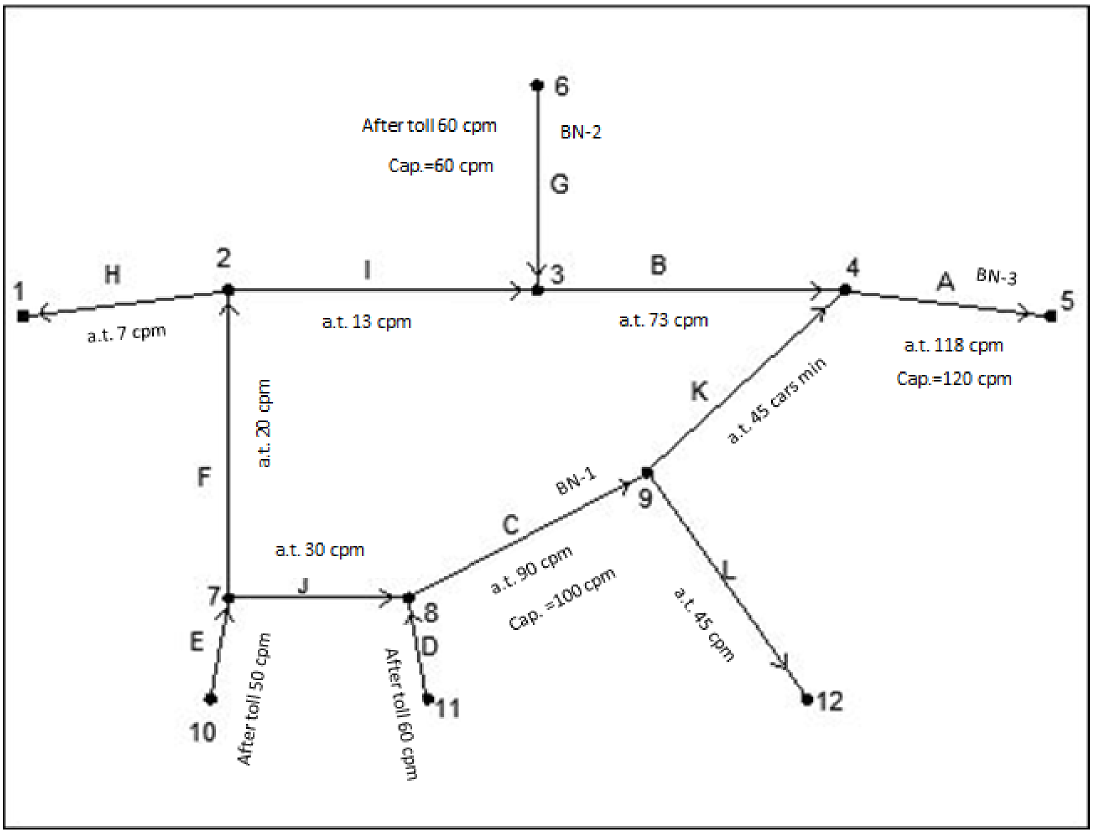

Figure 9 Equilibrium Flows for Toll Scenario-2. All bottlenecks have disappeared.

The updated results are given in Table 3. Although we have not made the discussion of monetary toll amounts necessary for achieving the results in Scenario-1 and Scenario-2 comparisons can be made by the total social and private costs of tolling against the cost of traffic congestion in the case of no toll.

Table 4 Toll Scenario-2 Equilibrium

Zone (Link)	Succ. of	Foll. by	Initial Flow	Cap.	Initial Flow After Toll	Eff. Inc. Flow	Eff. Outg. Flow	Average Waiting Time in Queue (in minutes)
A	B, K	-	160	120		118	118	0.00
B	G, I	A	100			73	73	0.00
C	J, D	K, L	120	100		90	90	0.00
D	-	C	80		60	60	60	0.00
E	-	F, J	70		50	50	50	0.00
F	E	H, I	30			20	20	0.00
G	-	B	80	60	60	60	60	0.00
H	F	-	10			7	7	0.00
I	F	B	20			13	13	0.00
J	E	C	40			30	30	0.00
K	C	A	60			45	45	0.00
L	C	-	60			45	45	0.00

3.3.2 A network example with a cycle

There are two differences of this network from the example studied in Section 3.3.1. First, this network contains a cycle and it will be seen that the existence of a cycle adds extra complications to reach an equilibrium (under both no-toll and toll scenarios) even when we consider only a-type effects. Second, we will consider the possibility of some drivers changing their commuting routes when tolls are applied (in Section 3.3.1 we only considered some drivers shifting to public transportation, changing their travel hours or car-pooling). The network with initial flows diagram is given in Figure 10.

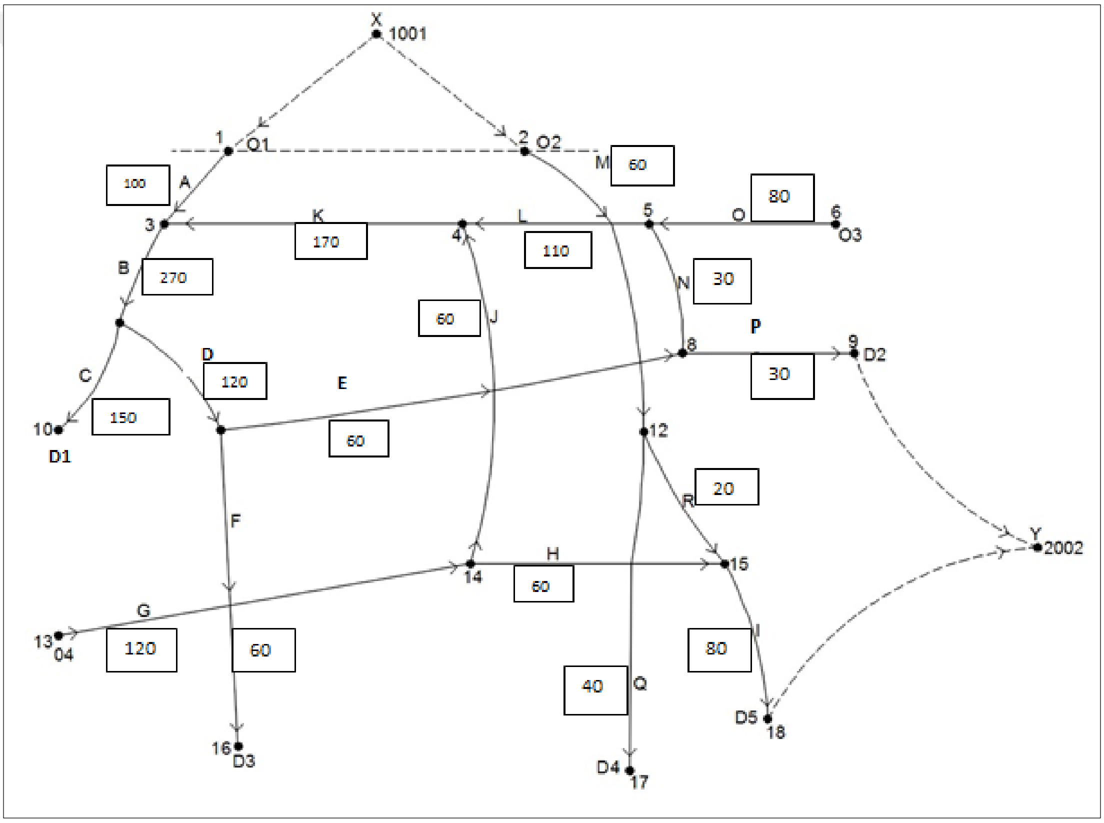

Figure 10 Initial flow values for another network

Similar to the network example in Section 3.3.1 the flows represent the actual values had there been free flow traffic. Because of possible bottleneck zones the initial flow values are not the actual values hence the no-toll equilibrium should be computed for morning rush hours. Nevertheless, the initial flow values fulfill the flow conservation equations as given by:

Flow (B) = Flow (A) + Flow (K) (4a)

Flow (B) = Flow (C) + Flow (D) (4b)

Flow (D) = Flow (E) + Flow (F) (4c)

Flow (G) = Flow (J) + Flow (H) (4d)

Flow (K) = Flow (J) + Flow (L) (4e)

Flow (L) = Flow (N) + Flow (O) (4f)

Flow (M) = Flow (Q) + Flow (R) (4g)

Flow (E) = Flow (N) + Flow (P) (4h)

Flow (I) = Flow (H) + Flow (R) (4i)

Let us consider that this network is an important junction on the highway system of a city for which node X represents a residential center and node Y represents a business center and some of the commuters use this junction every morning (assume that both nodes X and Y are distant from this junction). Because of the cycle B-D-E-N-L-K, it will be seen that to reach the equilibrium solution more than one round of computations might be necessary.

As explained in the earlier paragraph we consider the O-D pair (X-Y) in this example. We will assume that some of the drivers taking the route A-B-D-E-P can take the alternative route M-R-I because the origin (X) and destination (Y) of both routes are the same. On the other hand a commuter taking the route A-B-C does not have an alternative choice if he/she wants to make his journey using this junction.

The No-Toll Equilibrium

Similar to that in Section 3.3.1 we will develop the equilibrium flow values employing the processing algorithm given in Figure 6. It is assumed that the flow capacity of link B is 150 cpm and link C is 90 cpm. This means that the initial flow values cannot be the actual flow values during the morning rush hours. For the given initial flow values no other capacity values are violated (only the binding capacity values are illustrated in Table 4 hence only for links B and C). The derivation of the no-toll equilibrium is explained fully in the next paragraph.

Since there is a capacity limit of 180 cpm for link B a queue will form in that zone (for simplicity we will assume that a queue in zone B will not extend to secondary queues in zones A and K or beyond. Therefore, the outgoing from link B towards links C and D should be 100 and 80 cpm (instead of 150 and 120 cpm) respectively. While computing these values we reduce both quantities based on the ratio 150/270. Note that because of the capacity limit of 90 cpm through link C another queue formation takes place here as well (incoming flow is 100 cpm and outgoing flow is 90 cpm). Since flow rate through link D is 80 cpm, the revised flow values through links E and F will both be 40 cpm (instead of 60 cpm). 40 cpm exit via link F. The outgoing flow from link E toward links N and P will become 20 cpm (instead of 30 cpm). Outgoing flow through link N is 20 cpm which together with outgoing flow of 80 cpm through link O will result with incoming flow of 100 cpm through link L. At the same time, a flow of 120 cpm enter the network from node 13 through link G. This flow through link G is divided equally through links J and H. Therefore, a flow of 60 cpm through link J will be added to the outgoing flow through L to become 160 cpm incoming to link K. This means that the incoming flow through link B is actually 260 cpm (Flow (B) = Flow (A) + Flow (K)) instead of 270. The outgoing flow through link B is still 180 cpm. The remaining flow values through links C, D, F, E, N, P, L and K are unchanged. It was necessary to update some flow values twice because of the cycle contained in the network. All the rest of the flow values are same with those of the initial flows diagram (Figure 10).

50 cpm on link H and 20 cpm on link R are connected each other on node 15 as 70 cpm instead of 80 cpm. Hence, 70 cpm exit the network by using D5. Figure 11 shows the no-toll equilibrium for this network.

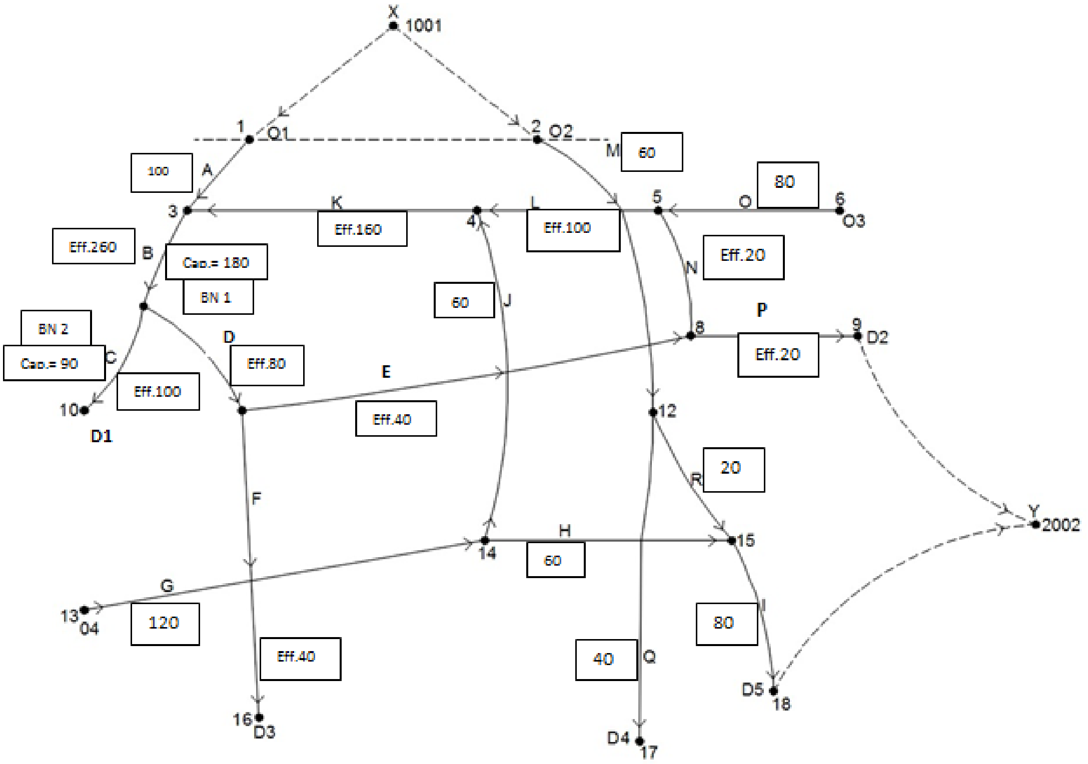

Figure 11 No-toll equilibrium for the new network

Similar to the example in Section 3.3.1 the average waiting times at bottleneck zones B and C can be computed with the single bottleneck simulation tool. Average queue waiting time is 13.33 minutes for link B, and 3.33 minutes for link C for the no-toll equilibrium. These results are summarized in Table 5.

Table 5 No-Toll Equilibrium

Zone (Link)	**Succ. of**	**Foll. by**	**Initial Flow**	**Capacity**	**Eff. Inc. Flow**	**Eff. Outg. Flow**	**Average Waiting Time in Queue (in minutes)**
A	-	B	100		100	100	0.00
B	A, K	C, D	270	180	260	180	13.33
C	B	-	150	90	100	90	3.33
D	B	E, F	120		80	80	0.00
E	D	N, P	60		40	40	0.00
F	D	-	60		40	40	0.00
G	-	J, H	120		120	120	0.00
H	G	I	60		60	60	0.00
I	H, R	-	80		80	80	0.00
J	G	K	60		60	60	0.00
K	J, L	B	170		160	160	0.00
L	N, O	K	110		100	100	0.00
M	-	Q, R	60		60	60	0.00
N	E	L	30		20	20	0.00
O	-	L	80		80	80	0.00
P	E	-	30		20	20	0.00
Q	M	-	40		40	40	0.00
R	M	I	20		20	20	0.00

Toll-Equilibrium – Scenario-1

In this scenario, we assume that an application of tolling will not result with drivers shifting to alternative routes (but shifting to alternative ways of transportation instead). To reduce the congestion in this network, consider a toll is applied through link A (call this toll-entry 1) and also through link G (call this toll-entry 2). Due to toll-entry 1, 60 cpm will enter through link A instead of 100 cpm. The effective incoming value is calculated as 220 cpm for bottleneck point on link B (160 cpm on link K plus 60 cpm on link A). However, the flow of link B still exceeds the capacity limitation of link B, so the effective outgoing value is 180 cpm on link B and therefore a queue formation will still take place (though with reduced average waiting time). Since the outgoing flow from link B has remained unchanged the incoming values for links C and D remain also unchanged meaning an identical queue formation for link C. Similar to the case of no-toll equilibrium, 40 cpm pass through links E and F. Flow rates through links N and P are again 20 cpm. Similarly, flow rate through link L is also unchanged (100 cpm).

Due to tolling at entry 2, 100 cpm will enter through link G instead of 120 cpm. At node 14, flow through link G is equally divided: 50 cpm through links J and H. At node 15, flows through links R and H are added and the flow passing through link I become 70 cpm.

Links J and L meet at node 4, therefore the incoming flow through K is computed as 150 cpm (instead of 160 cpm). Since links K and A meet at node 3, the incoming flow through B becomes 210 cpm. Because of the presence of a cycle in this network, the effective incoming value through link B has to be computed in two rounds (instead of 220 computed in first round). However, since the incoming flow through link B is still above the capacity limit the bottleneck is not eliminated but the average queue waiting time is smaller than that computed in the first round. The rest of the revised flow values remain unchanged (see Figure 12).

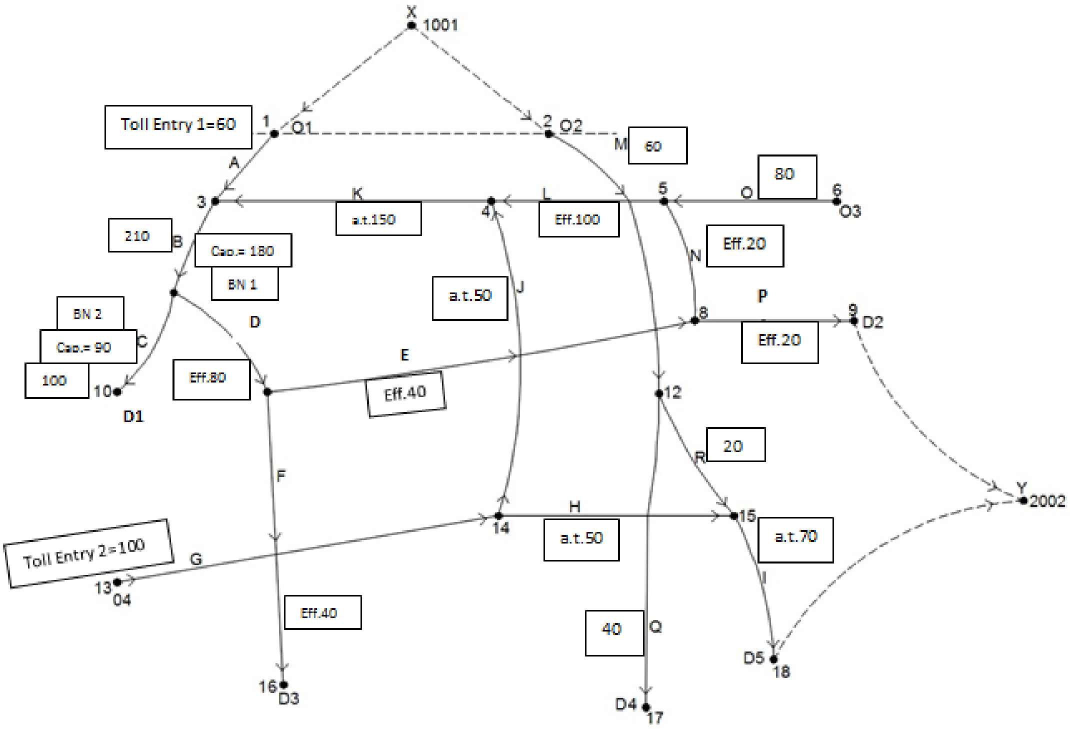

Figure 12 Equilibrium Flows for Tolling Scenario-1

Due to tolling average queue waiting time for link B is decreased to 5.00 minutes whereas the average queue waiting time for link C has not changed. The results are summarized in Table 6.

Table 6 Toll-Scenario-1 Equilibrium (Assuming no effect on alternative route of A-B-D-E-P)

Zone (Link)	Succ. of	Foll. by	Initial Flow	Capacity	Initial Flow After Toll	Eff. Inc. Flow	Eff. Outg. Flow	Average Waiting Time in Queue (in minutes)
A	-	B	100		60	60	60	0.00
B	A, K	C, D	270	180		210	180	5.00
C	B	-	150	90		100	90	3.33
D	B	E, F	120			80	80	0.00
E	D	N, P	60			40	40	0.00
F	D	-	60			40	40	0.00
G	-	J, H	120		100	100	100	0.00
H	G	I	60			50	50	0.00
I	H, R	-	80			70	70	0.00
J	G	K	60			50	50	0.00
K	J, L	B	170			150	150	0.00
L	N, O	K	110			100	100	0.00
M	-	Q, R	60			60	60	0.00
N	E	L	30			20	20	0.00
O	-	L	80			80	80	0.00
P	E	-	30			20	20	0.00
Q	M	-	40			40	40	0.00
R	M	I	20			20	20	0.00

Toll-Equilibrium: Scenario-2

In this scenario we will assume that toll-entry 1 through link A will cause some of the drivers using the route A-B-D-E-P to shift to alternative routes. As a result of tolling let some of the 40 drivers who stop using entry link A shift to alternative route M-R-I. Let the flow rate through link M increase by 30 cpm and let the flow rates through links R and I increase by 20 cpm and let the flow rate through link Q increase by 10 cpm. With these revised values the flow conservation equations are not violated. It is assumed that 10 drivers each minute (who originally used the entry link A) will switch to other transportation methods.

Based on tolling scenario-2 flow rate through link Q becomes 50 cpm causing a queue formation since its capacity is 40 cpm. Since flow rate through link R becomes 40 cpm due to flow conservation 90 cpm will pass through link I (a value just equal to its capacity).

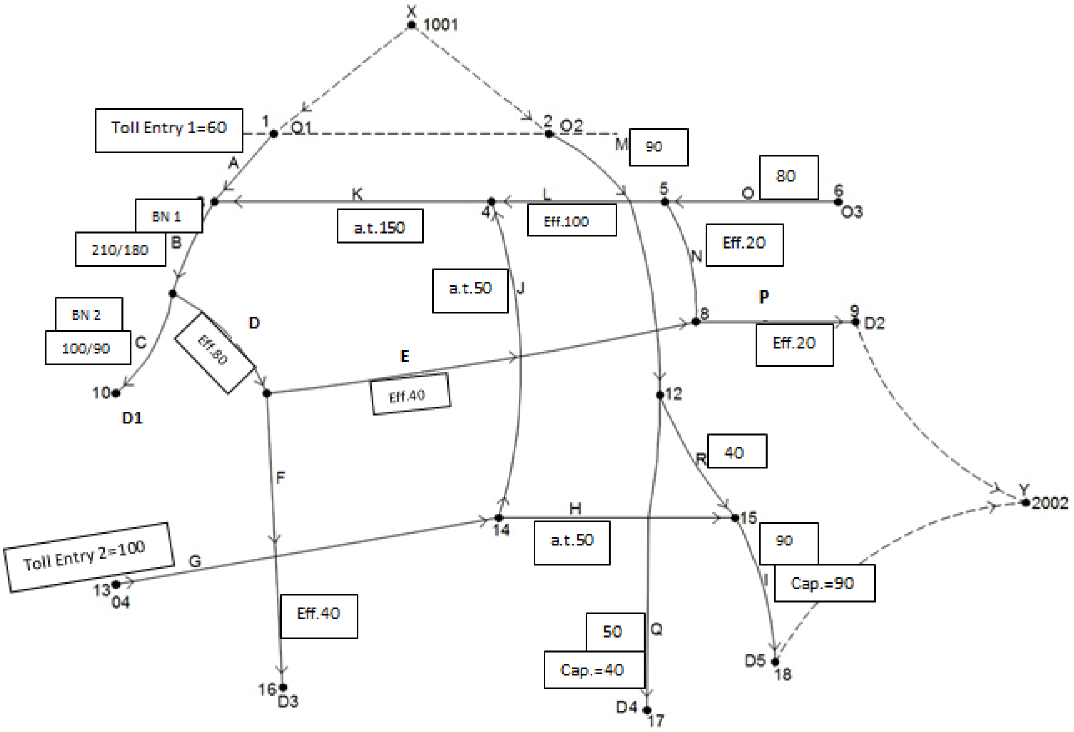

Figure 13 Equilibrium Flows for Tolling Scenario-2

The summary of Scenario-2 results is given in Table 7. This time due to increasing flow through link Q there is new queue formation at this zone. Average queue waiting times at zones B and C are same with those of Scenario-1.

Table 7 Toll-Scenario-2 Equilibrium (Assuming that some of the drivers who used route of A-B-D-E-P will shift to M-R-I and M-Q)

Zone (Link)	**Succ. of**	**Foll. by**	**Initial Flow**	**Capacity**	**Initial Flow After Toll**	**Eff. Inc. Flow**	**Eff. Outg. Flow**	**Average Waiting Time in Queue (in minutes)**
A	-	B	100		60	60	60	0.00
B	A, K	C, D	270	180		210	180	5.00
C	B	-	150	90		100	90	3.33
D	B	E, F	120			80	80	0.00
E	D	N, P	60			40	40	0.00
F	D	-	60			40	40	0.00
G	-	J, H	120		100	100	100	0.00
H	G	I	60			50	50	0.00
I	H, R	-	80	90		90	90	0.00
J	G	K	60			50	50	0.00
K	J, L	B	170			150	150	0.00
L	N, O	K	110			100	100	0.00
M	-	Q, R	60		90	90	90	0.00
N	E	L	30			20	20	0.00
O	-	L	80			80	80	0.00
P	E	-	30			20	20	0.00
Q	M	-	40	40		50	40	7.50
R	M	I	20			40	40	0.00

3.3.3 Derivation of the average waiting time for single bottleneck queues

In this subsection the average waiting time for the single bottleneck queue which has been computed as a simulation outcome will be derived analytically. We refer to Figure 4 to display the queuing structure. The parameters of the single bottleneck queue are:

a_1: Arrival rate during phase-1 (cars per minute)

t_1: Duration of phase-1 (minutes)

d: Departure rate from queue (cars per minute)

a_2: Arrival rate during phase-2 (cars per minute)

t_2: Duration of phase-2 (minutes)

To fulfill the queuing structure given in Figure 4 the parameters should satisfy the assumption (2) in Section 3.3. In our example at 7.30 AM the queuing starts, and reaches the maximum value at 8.30 AM. This interval is phase-1 and its duration will be denoted as t_1. The queue length increases linearly with a slope of $a_1 - d$. After 8.30 AM the arrival rate suddenly drops to a_2 so the queue length starts to decrease linearly with a rate of $a_2 - d$ (negative slope value). The time which the queue is completely dissolved depends on the value of a_2. If a_2 is relatively large it will take a relatively long time. If a_2 is relatively small the queue will disappear quickly. It will be seen that the average waiting time for the cars arriving during the queue formation depends on a_1,d, and t_1 regardless of the value of a_2.

Since time intervals in terms of seconds are not fine enough we have considered tertias (1 second = 60 tertias) to measure time between successive arrivals and successive departures during the simulation runs. We will consider similar units for the derivations although we keep arrival and departure rates as number of cars per minute.

At the peak level of the queue, a cumulative number of $a_1 t_1$ cars have arrived at the bottleneck zone. The number of queuing cars at that instant is equivalent to $(a_1 - d)t_1$. By the end of phase-2 the queue should diminish entirely, so the duration of phase-2 can be computed from:

$$(\boldsymbol{a_1} - \boldsymbol{d})\boldsymbol{t_1} = (\boldsymbol{d} - \boldsymbol{a_2})\boldsymbol{t_2} \tag{5}$$

Hence,

$$\boldsymbol{t_2} = \frac{(\boldsymbol{a_1} - \boldsymbol{d})\boldsymbol{t_1}}{\boldsymbol{d} - \boldsymbol{a_2}} \tag{6}$$

During phase-1 each new arrival waits slightly more than the previous one with a difference of (measured in tertias instead of minutes):

$$\frac{\mathbf{3600}}{\boldsymbol{d}} - \frac{\mathbf{3600}}{\boldsymbol{a_1}} = \frac{\mathbf{3600}(\boldsymbol{a_1} - \boldsymbol{d})}{\boldsymbol{d}\boldsymbol{a_1}} \tag{7}$$

The $\boldsymbol{i^{th}}$ arriving car during phase-1 has a waiting time of $\left(\mathbf{3600}\boldsymbol{i}\left(\frac{\boldsymbol{a_1}-\boldsymbol{d}}{\boldsymbol{d}\boldsymbol{a_1}}\right)\right)$ in tertias. Therefore, the average waiting time during phase-1 (in tertias) is:

$$\frac{\mathbf{1}}{(\boldsymbol{a_1}\boldsymbol{t_1} - \mathbf{1})} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{a_1}\boldsymbol{t_1}-\mathbf{1}} \left(\frac{\mathbf{3600}\boldsymbol{i}(\boldsymbol{a_1} - \boldsymbol{d})}{\boldsymbol{d}\boldsymbol{a_1}}\right) = \left(\frac{\mathbf{3600}(\boldsymbol{a_1} - \boldsymbol{d})\boldsymbol{t_1}}{\mathbf{2}\boldsymbol{d}}\right) \tag{8}$$

During phase-2 each new arrival waits slightly less than the previous one with a difference of:

$$\frac{\mathbf{3600}}{\boldsymbol{a_2}} - \frac{\mathbf{3600}}{\boldsymbol{d}} = \frac{\mathbf{3600}(\boldsymbol{d} - \boldsymbol{a_2})}{\boldsymbol{d}\boldsymbol{a_2}} \tag{9}$$

A similar derivation yields the average waiting time during phase-2 as:

$$\frac{\mathbf{1}}{(\boldsymbol{a_2}\boldsymbol{t_2} - \mathbf{1})} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{a_2}\boldsymbol{t_2}-\mathbf{1}} \left(\frac{\mathbf{3600}\boldsymbol{i}(\boldsymbol{d} - \boldsymbol{a_2})}{\boldsymbol{d}\boldsymbol{a_2}}\right) = \left(\frac{\mathbf{3600}(\boldsymbol{d} - \boldsymbol{a_2})\boldsymbol{t_2}}{\mathbf{2}\boldsymbol{d}}\right) \tag{10}$$

Paradox of the average waiting time

Intuitively one can say that the average waiting time in the single bottleneck queue should be smaller for a relatively small $\boldsymbol{a_2}$ value since the queue will disappear quickly and the duration of phase-2 will be much shorter. However, it did not turn out to be the case during our simulation results and the verification of the fact that whatever the value of $\boldsymbol{a_2}$ is the average waiting time does not change is simple. Let us substitute $\boldsymbol{t_2}$ from equation (6) in the right-hand-side of equation (10). We will obtain the right-hand-side of the expression in equation (8). This means that the average waiting time in phase-1 is always equal to the average waiting time in phase-2 regardless of the value of $\boldsymbol{a_2}$ (and hence the duration of phase-2) provided that assumption (2) is satisfied. Therefore, the combined phase-1 and phase-2 average waiting time is also identical.

To prevent round off errors due to discontinuity we have chosen $\boldsymbol{a_2}$ values so that the first car which does not wait in the queue witnesses the departure of the last car which has waited in the queue. This can be ensured if the peak waiting time value is divisible by the quantity given by equation (9). The maximum waiting time in the queue is experienced by the $(\boldsymbol{a_1 t_1})^{\boldsymbol{th}}$ car and equals to:

$$\frac{3600(a_1t_1)(a_1-d)}{da_1}=\frac{3600(a_1-d)t_1}{d} \qquad (11)$$

Therefore, the value below should be an integer

$$\frac{\frac{3600(a_1-d)t_1}{d}}{\frac{3600(d-a_2)}{da_2}}=\frac{(a_1-d)t_1a_2}{(d-a_2)} \qquad (12)$$

Note that the quantity in equation (12) is also equivalent to $\boldsymbol{t_2 a_2}$.

A summary of the average waiting times for several phase-2 arrival rates is given in Table 8 calculated from equations (6) and (10). As just said the $\boldsymbol{a_2}$ value in the table has been chosen to guarantee an integer quantity for the expression given by equation (12). The results in the table confirm the results of the simulation runs.

Table 8 For fixed arrival/departure rates during phase-1 ($\boldsymbol{a_1}$ = 80, $\boldsymbol{d}$ = 60 cpm) phase-2 duration and average waiting time for all vehicles crossing the bottleneck zone are given for various values of $\boldsymbol{a_2}$ (departure rate being unchanged).

a_2 (vehicles per minute)	Queue Formation start at	Queue Reaches peak at	Queue Disappears at	t_1 (min)	# of arrivals in phase-1	Total Waiting time of arrivals in phase-1 (min)	t_2 (min)	# of arrivals in phase-2	Total Waiting time of arrivals in phase-2 (min)	Average Waiting Time in phase_1 (min)	Average Waiting Time in phase_2 (min)
48	07:30:00AM	08:30:00AM	10:10:00AM	60.00	4799	47,990.00	100.00	4799	47,990.00	10	10
36	07:30:00AM	08:30:00AM	09:20:00AM	60.00	4799	47,990.00	50.00	1799	17,990.00	10	10
24	07:30:00AM	08:30:00AM	09:03:20 AM	60.00	4799	47,990.00	33.33	799	7,990.00	10	10
15	07:30:00AM	08:30:00AM	08:56:40 AM	60.00	4799	47,990.00	26.67	399	3,990.00	10	10
10	07:30:00AM	08:30:00AM	08:54:00 AM	60.00	4799	47,990.00	24.00	239	2,390.00	10	10

CHAPTER 4

CONCLUSION

Traffic congestion is an important problem for many cities and metropolitan areas and therefore congestion tolls have been the focus of academic research. Drivers have to pay significant amount of charges, however congestion tolling reduces overall time spent in traffic and makes the city life easier. For this purpose, expensive software have been developed by many companies to control and manage tolling systems in cities.

Our purpose has been to develop an alternative approach for congestion management, especially at busy traffic junctions where several commuter routes meet. We have slightly modified the deterministic single bottleneck model of Vickrey keeping the queue increase and decrease properties same and developed a simulation tool for the single bottleneck model to compute the average waiting time based on the severity of the bottleneck. The single bottleneck models have been treated by many researchers in the literature. However, the single bottleneck approach cannot be used plainly in a simple city network because changes in flow values will affect the other flow values in the entire network. Therefore, a step by step methodology has been explained in detail to determine equilibrium flow values in a congested network so that the single bottleneck approach could be used separately at bottleneck zones. In this way we were able to evaluate the results of alternative tolling scenarios. For this approach we have considered tolling through routes instead of cordon or zone type and since an analytical treatment of the supply-demand analysis was beyond the scope of this work we simply assumed that increasing tolls through a route would reduce the flow rate. We also developed an analytical derivation of the average waiting time computed by the simulation tool for the single bottleneck problem. As an indirect outcome of this derivation we have arrived at an interesting paradoxical result related to the average time of cars waiting in a queue. The average waiting time turns out to be independent of how fast the queue starts to dissolve and we believe that this result is very interesting from the perspective of discrete event simulation. We anticipate that this result can be generalized to similar queuing systems where the arrival and service rates are of stochastic nature.

Although we have considered simple networks in this study, the described methodology can also be applied for more complicated networks. We intend to extend our study along this direction.

REFERENCES

1. Akamatsu, T., Wada, K., Hayashi, S., 2015. The corridor problem with discrete multiple bottlenecks, *Transportation Research*, 7, pp.474-498.
2. An, Y., Zhang, Z., 2012. Congestion with heterogeneous commuters, *Economic Modelling*, 29, pp.557-565.
3. Anas, A., Lindsey, R., 2011. Reducing Urban Road Transportation Externalities: Road Pricing in Theory and in Practice, *Review of Environmental Economics and Policy*, 5(1), pp.66-88.
4. Arnott, R., Small, K., 1994. The Economics of Traffic Congestion, *American Scientist*, 82(5), pp. 446-455.
5. Arnott, R., de Palma, A., Lindsey, R., 1994. The welfare effects of congestion tolls with heterogeneous commuters, *Journal of Transport Economics and Policy*, 28 (2), pp. 139–161.
6. Arnott, R., de Palma, A., Lindsey, R., 1993. A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand, *The American Economic Review*, 83 (2), pp. 161–179.
7. Arnott, R., de Palma, A., Lindsey, R., 1994. The Welfare Effects of Congestion Tolls with Heterogeneous Commuters, *Journal of Transport Economics and Policy*, 28 (2), pp. 139–161.
8. Ben-Akiva, M., Palma, A., Kanaroglou, P., 1986. Dynamic Model of Peak Period Traffic Congestion with Elastic Arrival Rates, *Transportation Science*, 20(3), pp. 164-181.
9. Bloomberg, L., Dale, J., 2000. A Comparison of the VISSIM and CORSIM Traffic Simulation Models, *Institute of Transportation Engineers Annual Meeting.*

10. Braid, R.M., 1996. Peak-Load Pricing of a Transportation Route with an Unpriced Substitute, *Journal of Urban Economics*, 40, pp. 179-197.

11. Buckeye, K.R., 2014. An Overview of Congestion Charging on Roadways in the United States of America.

12. Chen, H., Nie, Y.M., Yin, Y., 2015. Optimal multi-step toll design under general user heterogeneity, *Transportation Research*, Part B, 81, pp.775-793.

13. Cohen, Y., 1987. Commuter welfare under peak-period congestion tolls: Who gains and who loses, *International Journal of Transport Economics*, 14 (3), pp. 239–266.

14. Daganzo, C.F.,Garcia, R.C., 2000. A Pareto Improving Strategy for the Time-Dependent Morning Commute Problem, *Transportation Science,* 34(3), pp. 303-311.

15. Danielis, R., Marcucci, E., 2002. Bottleneck road congestion pricing with a competing railroad service, *Transportation Research*, 38, pp.379-388.

16. D. Banister, “Critical pragmatism and congestion charging in London,” International Social Science Journal, vol. 55, no. 176, pp. 249–264, 2003.

17. Demirtaş, R.B. Kadıköy Merkez Bölgesinde Trafik Tıkanıklık Fiyatlandırması Potansiyeli Üzerine Bir Araştırma. MSc. Yıldız Technical University,2009.

18. Eliasson, J., 2014. The Stockholm Congestion Charges, *Centre of Transport Studies.*

19. Else, P.K., 1981. A Reformulation of the Theory of Optimal Congestion Taxes, *Journal of Transport Economics and Policy*, 15(3), pp.217-232.

20. G. Giuliano, “An assessment of the political acceptability of congestion pricing,” Transportation, vol. 19, no. 4, pp. 335–358, 1992.

21. Hendrickson, C., Kocur, G., 1981. Schedule delay and departure time decisions in a deterministic model, *Transportation Science*, 15 (1), pp. 62–77.

22. Hurdle, V.F., 1981. Equilibrium flows on urban freeways, *Transportation Science*, 15 (3), pp. 255–293.

23. Ingles, D., 2009. Road congestion charges: An idea whose time has come, *The Australia Institute*.

24. Kearns, S., 2014. Congestion Charging in London, *Transport for London.*

25. Kristofersson, I., Engelson, L., 2009. A Dynamic Transportation Model for the Stockholm Area: Implementation Issues Regarding Departure Time Choice and OD-pair Reduction, *Network and Spatial Economics*, 9, pp.551-573.

26. Leape, J., 2006. The London Congestion Charge, Journal of Economic Perspectives, 20(4), pp.157-176.

27. Lindsey, C.R., Berg, V.A.C., Verhoef, E., 2012. Step tolling with bottleneck queuing congestion, *Journal of Urban Economics*, 72, pp.46-59.

28. May, A.D., 1992. Road pricing: An international perspective, *Transportation* (Kluwer Academic Publishers), pp.313-333.

29. Moroni, S., 2014. Eco-Zone in Milan: Policy design, Enforcement and Impcats on Traffic and the Environment.

30. Palma, A., Lindsey, R., Monchambert, G., 2017. The economics of crowding in rail transit, *Journal of Urban Economics*, 101, pp.106-122.

31. Palma, A., Kilani, M., Lindsey, R., 2005. Congestion pricing on a road network: A study using the dynamic equilibrium simulator METROPOLIS, *Transportation Research*, 39, pp.588-611.

32. Phang, S.Y., Toh, R.S., 2004. Road Congestion Pricing in Singapore: 1975 to 2003, *Transportation Journal*, 43(2), pp.16-25.

33. Smith, M.J., 1984. The Existence of a Time-Dependent Equilibrium Distribution of Arrivals at a Single Bottleneck, *Transportation Science*, 18(4), pp. 384-394.

34. Vickrey, W.S., 1969. Congestion theory and transport investment, *The American Economic Review*, 59 (2), pp. 251–260.

35. Walters, A.A., 1961. The theory and measurement of private and social costs of highway congestion, *Econometrica*, 29 (4), pp. 676–699.

36. Watson, P.L., Holland, E.P., 1976. Congestion Pricing – the example of Singapore, Athens Center of Ekistics, 42(248), pp.14-18.
37. Xin, W., Levinson, D., 2015. Stochastic Congestion and Pricing Model with Endogenous Departure Time Selection and Heterogeneous Travelers, *Mathematical Population Studies*, 22(1), pp.37-52.
38. Xian, G.S., 2014. Singapore's Experience in Road Pricing.
39. Yüksel, H. Kentiçi ulaşımda talep yönetimi. MSc. Yıldız Technical University,1998.
40. Yüksel, H. Toplu taşımacılığın geliştirilmesi için bir tıkanıklık fiyatlandırması modeli önerisi. PhD. Yıldız Technical University,2004.
41. Yüksel, H., Yardım, M.S., 2007. Tıkanıklık ücretlendirmesi uygulamalarının Eminönü Bölgesi'nde bir trafik iyileştirme yöntemi olarak kullanılabilirliğinin araştırılması, "Projem İstanbul" Araştırma Projesi.
42. Yüksel, H., Yardım, M.S., Gürsoy, M., 2010. Eminönü İçin Bir Trafik Tıkanıklık Fiyatlandırması Modeli, *İMO Teknik Dergi*, pp.4995-5022.
43. Zhang, X., et. al., 2005. Integrated scheduling of daily work activities and morning-evening commutes with bottleneck congestion, 39, pp.41-60.

APPENDICES

APPENDIX 1

Arrival rate in phase-1: a_1 = 80 cars per minute

Arrival rate in phase-2: a_2 = 48 cars per minute

Departure rate: d=60 cars per minute

Arrival Time	Queue Length	Waiting Time (tertias)	Waiting Time (minutes)
07:30	20	1200	0,33
07:31	40	2400	0,67
07:32	60	3600	1,00
07:33	80	4800	1,33
07:34	100	6000	1,67
07:35	120	7200	2,00
07:36	140	8400	2,33
07:37	160	9600	2,67
07:38	180	10800	3,00
07:39	200	12000	3,33
07:40	220	13200	3,67
07:41	240	14400	4,00
07:42	260	15600	4,33
07:43	280	16800	4,67
07:44	300	18000	5,00
07:45	320	19200	5,33
07:46	340	20400	5,67
07:47	360	21600	6,00
07:48	380	22800	6,33
07:49	400	24000	6,67

07:50	420	25200	7,00
07:51	440	26400	7,33
07:52	460	27600	7,67
07:53	480	28800	8,00
07:54	500	30000	8,33
07:55	520	31200	8,67
07:56	540	32400	9,00
07:57	560	33600	9,33
07:58	580	34800	9,67
07:59	600	36000	10,00
08:00	620	37200	10,33
08:01	640	38400	10,67
08:02	660	39600	11,00
08:03	680	40800	11,33
08:04	700	42000	11,67
08:05	720	43200	12,00
08:06	740	44400	12,33
08:07	760	45600	12,67
08:08	780	46800	13,00
08:09	800	48000	13,30
08:10	820	49200	13,67
08:11	840	50400	14,00
08:12	860	51600	14,33
08:13	880	52800	14,67
08:14	900	54000	15,00
08:15	920	55200	15,33
08:16	940	56400	15,67
08:17	960	57600	16,00
08:18	980	58800	16,33
08:19	1000	60000	16,67
08:20	1020	61200	17,00

08:21	1040	62400	17,33
08:22	1060	63600	17,67
08:23	1080	64800	18,00
08:24	1100	66000	18,33
08:25	1120	67200	18,67
08:26	1140	68400	19,00
08:27	1160	69600	19,33
08:28	1180	70800	19,67
08:29	1200	72000	20,00
08:30	1188	71280	19,80
08:31	1176	70560	19,60
08:32	1164	69840	19,40
08:33	1152	69120	19,20
08:34	1140	68400	19,00
08:35	1128	67680	18,80
08:36	1116	66960	18,60
08:37	1104	66240	18,40
08:38	1092	65520	18,20
08:39	1080	64800	18,00
08:40	1068	64080	17,80
08:41	1056	63360	17,60
08:42	1044	62640	17,40
08:43	1032	61920	17,20
08:44	1020	61200	17,00
08:45	1008	60480	16,80
08:46	996	59760	16,60
08:47	984	59040	16,40
08:48	972	58320	16,20
08:49	960	57600	16,00
08:50	948	56880	15,80
08:51	936	56160	15,60

08:52	924	55440	15,40
08:53	912	54720	15,20
08:54	900	54000	15,00
08:55	888	53280	14,80
08:56	876	52560	14,60
08:57	864	51840	14,40
08:58	852	51120	14,20
08:59	840	50400	14,00
09:00	828	49680	13,80
09:01	816	48960	13,60
09:02	804	48240	13,40
09:03	792	47520	13,20
09:04	780	46800	13,00
09:05	768	46080	12,80
09:06	756	45360	12,60
09:07	744	44640	12,40
09:08	732	43920	12,20
09:09	720	43200	12,00
09:10	708	42480	11,80
09:11	696	41760	11,60
09:12	684	41040	11,40
09:13	672	40320	11,20
09:14	660	39600	11,00
09:15	648	38880	10,80
09:16	636	38160	10,60
09:17	624	37440	10,40
09:18	612	36720	10,20
09:19	600	36000	10,00
09:20	588	35280	9,80
09:21	576	34560	9,60
09:22	564	33840	9,40

09:23	552	33120	9,20
09:24	540	32400	9,00
09:25	528	31680	8,80
09:26	516	30960	8,60
09:27	504	30240	8,40
09:28	492	29520	8,20
09:29	480	28800	8,00
09:30	468	28080	7,80
09:31	456	27360	7,60
09:32	444	26640	7,40
09:33	432	25920	7,20
09:34	420	25200	7,00
09:35	408	24480	6,80
09:36	396	23760	6,60
09:37	384	23040	6,40
09:38	372	22320	6,20
09:39	360	21600	6,00
09:40	348	20880	5,80
09:41	336	20160	5,60
09:42	324	19440	5,40
09:43	312	18720	5,20
09:44	300	18000	5,00
09:45	288	17280	4,80
09:46	276	16560	4,60
09:47	264	15840	4,40
09:48	252	15120	4,20
09:49	240	14400	4,00
09:50	228	13680	3,80
09:51	216	12960	3,60
09:52	204	12240	3,40
09:53	192	11520	3,20

09:54	180	10800	3,00
09:55	168	10080	2,80
09:56	156	9360	2,60
09:57	144	8640	2,40
09:58	132	7920	2,20
09:59	120	7200	2,00
10:00	108	6480	1,80
10:01	96	5760	1,60
10:02	84	5040	1,40
10:03	72	4320	1,20
10:04	60	3600	1,00
10:05	48	2880	0,80
10:06	36	2160	0,60
10:07	24	1440	0,40
10:08	12	720	0,20
10:09	0	0	0
10:10			

APPENDIX 2

Arrival rate in phase-1: a_1 = 80 cars per minute

Arrival rate in phase-2: a_2 = 12 cars per minute

Departure rate: d=60 cars per minute

Arrival Time	Queue Length	Waiting Time (tertias)	Waiting Time (minutes)
07:30	20	1200	0,33
07:31	40	2400	0,67
07:32	60	3600	1,00
07:33	80	4800	1,33
07:34	100	6000	1,67
07:35	120	7200	2,00
07:36	140	8400	2,33
07:37	160	9600	2,67
07:38	180	10800	3,00
07:39	200	12000	3,33
07:40	220	13200	3,67
07:41	240	14400	4,00
07:42	260	15600	4,33
07:43	280	16800	4,67
07:44	300	18000	5,00
07:45	320	19200	5,33
07:46	340	20400	5,67
07:47	360	21600	6,00
07:48	380	22800	6,33
07:49	400	24000	6,67
07:50	420	25200	7,00

07:51	440	26400	7,33
07:52	460	27600	7,67
07:53	480	28800	8,00
07:54	500	30000	8,33
07:55	520	31200	8,67
07:56	540	32400	9,00
07:57	560	33600	9,33
07:58	580	34800	9,67
07:59	600	36000	10,00
08:00	620	37200	10,33
08:01	640	38400	10,67
08:02	660	39600	11,00
08:03	680	40800	11,33
08:04	700	42000	11,67
08:05	720	43200	12,00
08:06	740	44400	12,33
08:07	760	45600	12,67
08:08	780	46800	13,00
08:09	800	48000	13,33
08:10	820	49200	13,67
08:11	840	50400	14,00
08:12	860	51600	14,33
08:13	880	52800	14,67
08:14	900	54000	15,00
08:15	920	55200	15,33
08:16	940	56400	15,67
08:17	960	57600	16,00
08:18	980	58800	16,33
08:19	1000	60000	16,67
08:20	1020	61200	17,00
08:21	1040	62400	17,33

08:22	1060	63600	17,67
08:23	1080	64800	18,00
08:24	1100	66000	18,33
08:25	1120	67200	18,67
08:26	1140	68400	19,00
08:27	1160	69600	19,33
08:28	1180	70800	19,67
08:29	1200	72000	20,00
08:30	1152	69120	19,20
08:31	1104	66240	18,40
08:32	1056	63360	17,60
08:33	1008	60480	16,80
08:34	960	57600	16,00
08:35	912	54720	15,20
08:36	864	51840	14,40
08:37	816	48960	13,60
08:38	768	46080	12,80
08:39	720	43200	12,00
08:40	672	40320	11,20
08:41	624	37440	10,40
08:42	576	34560	9,60
08:43	528	31680	8,80
08:44	480	28800	8,00
08:45	432	25920	7,20
08:46	384	23040	6,40
08:47	336	20160	5,60
08:48	288	17280	4,80
08:49	240	14400	4,00
08:50	192	11520	3,20
08:51	144	8640	2,40
08:52	96	5760	1,60

08:53	48	2880	0,80
08:54	0	0	0,00

APPENDIX 3

Arrival rate in phase-1: a_1 = 120 cars per minute

Arrival rate in phase-2: a_2 = 80 cars per minute

Departure rate: d=100 cars per minute

Arrival Time	**Queue Length**	**Waiting Time (tertias)**	**Waiting Time (minutes)**
07:30	20	720	0,20
07:31	40	1440	0,40
07:32	60	2160	0,60
07:33	80	2880	0,80
07:34	100	3600	1,00
07:35	120	4320	1,20
07:36	140	5040	1,40
07:37	160	5760	1,60
07:38	180	6480	1,80
07:39	200	7200	2,00
07:40	220	7920	2,20
07:41	240	8640	2,40
07:42	260	9360	2,60
07:43	280	10080	2,80
07:44	300	10800	3,00
07:45	320	11520	3,20
07:46	340	12240	3,40
07:47	360	12960	3,60
07:48	380	13680	3,80
07:49	400	14400	4,00
07:50	420	15120	4,20

07:51	440	15840	4,40
07:52	460	16560	4,60
07:53	480	17280	4,80
07:54	500	18000	5,00
07:55	520	18720	5,20
07:56	540	19440	5,40
07:57	560	20160	5,60
07:58	580	20880	5,80
07:59	600	21600	6,00
08:00	620	22320	6,20
08:01	640	23040	6,40
08:02	660	23760	6,60
08:03	680	24480	6,80
08:04	700	25200	7,00
08:05	720	25920	7,20
08:06	740	26640	7,40
08:07	760	27360	7,60
08:08	780	28080	7,80
08:09	800	28800	8,00
08:10	820	29520	8,20
08:11	840	30240	8,40
08:12	860	30960	8,60
08:13	880	31680	8,80
08:14	900	32400	9,00
08:15	920	33120	9,20
08:16	940	33840	9,40
08:17	960	34560	9,60
08:18	980	35280	9,80
08:19	1000	36000	10,00
08:20	1020	36720	10,20
08:21	1040	37440	10,40

08:22	1060	38160	10,60
08:23	1080	38880	10,80
08:24	1100	39600	11,00
08:25	1120	40320	11,20
08:26	1140	41040	11,40
08:27	1160	41760	11,60
08:28	1180	42480	11,80
08:29	1200	43200	12,00
08:30	1180	42480	11,80
08:31	1160	41760	11,60
08:32	1140	41040	11,40
08:33	1120	40320	11,20
08:34	1100	39600	11,00
08:35	1080	38880	10,80
08:36	1060	38160	10,60
08:37	1040	37440	10,40
08:38	1020	36720	10,20
08:39	1000	36000	10,00
08:40	980	35280	9,80
08:41	960	34560	9,60
08:42	940	33840	9,40
08:43	920	33120	9,20
08:44	900	32400	9,00
08:45	880	31680	8,80
08:46	860	30960	8,60
08:47	840	30240	8,40
08:48	820	29520	8,20
08:49	800	28800	8,00
08:50	780	28080	7,80
08:51	760	27360	7,60
08:52	740	26640	7,40

08:53	720	25920	7,20
08:54	700	25200	7,00
08:55	680	24480	6,80
08:56	660	23760	6,60
08:57	640	23040	6,40
08:58	620	22320	6,20
08:59	600	21600	6,00
09:00	580	20880	5,80
09:01	560	20160	5,60
09:02	540	19440	5,40
09:03	520	18720	5,20
09:04	500	18000	5,00
09:05	480	17280	4,80
09:06	460	16560	4,60
09:07	440	15840	4,40
09:08	420	15120	4,20
09:09	400	14400	4,00
09:10	380	13680	3,80
09:11	360	12960	3,60
09:12	340	12240	3,40
09:13	320	11520	3,20
09:14	300	10800	3,00
09:15	280	10080	2,80
09:16	260	9360	2,60
09:17	240	8640	2,40
09:18	220	7920	2,20
09:19	200	7200	2,00
09:20	180	6480	1,80
09:21	160	5760	1,60
09:22	140	5040	1,40
09:23	120	4320	1,20

09:24	100	3600	1,00
09:25	80	2880	0,80
09:26	60	2160	0,60
09:27	40	1440	0,40
09:28	20	720	0,20
09:29	0	0	0,00

APPENDIX 4

Arrival rate in phase-1: a_1 = 120 cars per minute

Arrival rate in phase-2: a_2 = 20 cars per minute

Departure rate: d=100 cars per minut**e**

Arrival Time	**Queue Length**	**Waiting Time (tertias)**	**Waiting Time (minutes)**
07:30	20	720	0,20
07:31	40	1440	0,40
07:32	60	2160	0,60
07:33	80	2880	0,80
07:34	100	3600	1,00
07:35	120	4320	1,20
07:36	140	5040	1,40
07:37	160	5760	1,60
07:38	180	6480	1,80
07:39	200	7200	2,00
07:40	220	7920	2,20
07:41	240	8640	2,40
07:42	260	9360	2,60
07:43	280	10080	2,80
07:44	300	10800	3,00
07:45	320	11520	3,20
07:46	340	12240	3,40
07:47	360	12960	3,60
07:48	380	13680	3,80
07:49	400	14400	4,00
07:50	420	15120	4,20

07:51	440	15840	4,40
07:52	460	16560	4,60
07:53	480	17280	4,80
07:54	500	18000	5,00
07:55	520	18720	5,20
07:56	540	19440	5,40
07:57	560	20160	5,60
07:58	580	20880	5,80
07:59	600	21600	6,00
08:00	620	22320	6,20
08:01	640	23040	6,40
08:02	660	23760	6,60
08:03	680	24480	6,80
08:04	700	25200	7,00
08:05	720	25920	7,20
08:06	740	26640	7,40
08:07	760	27360	7,60
08:08	780	28080	7,80
08:09	800	28800	8,00
08:10	820	29520	8,20
08:11	840	30240	8,40
08:12	860	30960	8,60
08:13	880	31680	8,80
08:14	900	32400	9,00
08:15	920	33120	9,20
08:16	940	33840	9,40
08:17	960	34560	9,60
08:18	980	35280	9,80
08:19	1000	36000	10,00
08:20	1020	36720	10,20
08:21	1040	37440	10,40

08:22	1060	38160	10,60
08:23	1080	38880	10,80
08:24	1100	39600	11,00
08:25	1120	40320	11,20
08:26	1140	41040	11,40
08:27	1160	41760	11,60
08:28	1180	42480	11,80
08:29	1200	43200	12,00
08:30	1120	40320	11,20
08:31	1040	37440	10,40
08:32	960	34560	9,60
08:33	880	31680	8,80
08:34	800	28800	8,00
08:35	720	25920	7,20
08:36	640	23040	6,40
08:37	560	20160	5,60
08:38	480	17280	4,80
08:39	400	14400	4,00
08:40	320	11520	3,20
08:41	240	8640	2,40
08:42	160	5760	1,60
08:43	80	2880	0,80
08:44	0	0	0,00