vrije Universiteit amsterdam

GPS TECHNOLOGY
OPTIMIZING CAR NAVIGATION

BMI Paper

Elvis N. Ngah
Business Mathematics and Informatics
Vrije Universiteit Amsterdam

Date: August 15, 2006.

Supervisor: Dr. Sandjai Bhulai
Department of Mathematics
Stochastics Section

This page is intentionally left blank.

CONTENTS

This page is intentionally left blank.

Preface

Prior to completion, each student of the Master program in Business Mathematics and Informatics (BMI) of the Vrije Universiteit Amsterdam has to write a compulsory paper on a subject matching the student's interest. This paper is research oriented and has to cover all three areas of focus, i.e., Business, Mathematics, and Computer Science. In principle, the chosen topic should use Mathematics/Computer Science to tackle a given societal problem. During my study, I became interested in Communication Networks. With GPS, being an emerging technology within the telecommunication industry, which is commercially used in car navigation, I decided to take an insight on how car navigation could be optimized.

Acknowledgements

I would like to give special thanks to my supervisor, Dr. Sandjai Bhulai, for squeezing time out of his tight schedule to guide me through this paper. His productive remarks made me realize this work.

Abstract

GPS technology is a promising technology that has applications in many aspects of life, such as in agriculture, law, sports, the automobile industry, etc. In the automobile industry, GPS is used in many forms. In this paper, I explain how GPS technology could be used for optimal routing decisions on urban highways. In this way, we can reduce traffic congestion, eventually reduce travel times and air pollution, and even save money on fuel consumption. This task is accomplished by using the data collected by the car GPS navigation system, analyzing this data, and then using the data to model how optimal decisions can be made for a driver to reach his/her destination.

1 INTRODUCTION

Any system that can provide intelligent vehicle location and navigation information will let us avoid congested freeways and find more efficient routes to our destinations, saving millions of euros in gasoline and tons of air pollution. Travel aboard ships and aircrafts will be safer in all weather conditions. Businesses with large numbers of outside plants (e.g., railroads and utilities) will be able to manage their resources more efficiently, reducing consumer costs. A Global Positioning System (GPS) provides the answer to facilitate all these issues.

A GPS is a space-based radio positioning system that combines computer mapping techniques to provide 24-hour three-dimensional position, velocity, and time information to suitably equipped users anywhere on or near the surface of the earth (and sometimes off the earth). GPS is amongst the major developments of the wireless telecommunication industry. It is an important tool for map-making and land surveying and has become a vital global utility, which is indispensable for modern navigation on land, sea, and air.

A GPS is basically divided into 3 major components as shown in Figure 1.1. These are the user segment, the control segment, and the space segment. A description of these components is given in many papers. I decided to adopt the one given in [11].

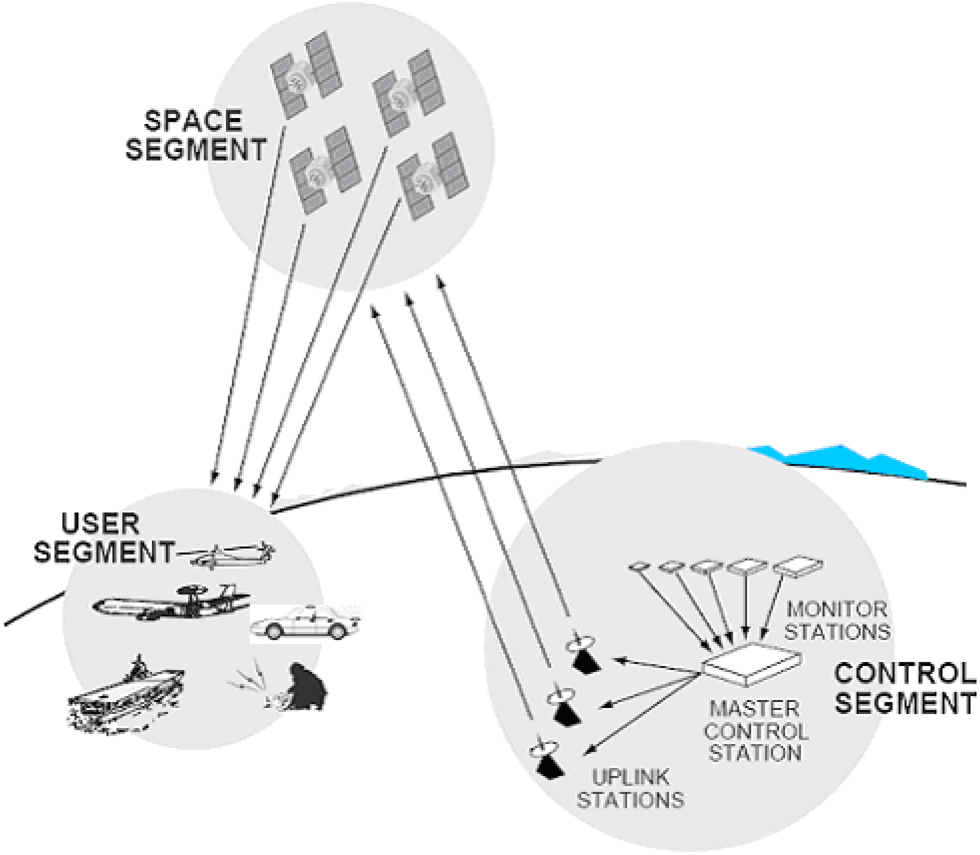


Figure 1.1 Components of a GPS.

The Space Segment

The space segment is composed of the GPS satellites[1] that transmit time and position in the form of radio signals to the user. The whole set of 24 satellites is called a *constellation*. The source of energy for the constellation is the sun. In the situation of a solar eclipse (in which case there is no solar energy), highly powered backup batteries are used [16]. As shown in Figure 1.2 above, the GPS constellation is arranged in 6 equally spaced orbital planes. Each of these planes thus consists of 4 satellites and is inclined at 55 degrees to the equator. In this configuration, there is worldwide coverage since at least four satellites are available from any location on the earth's surface at all times.

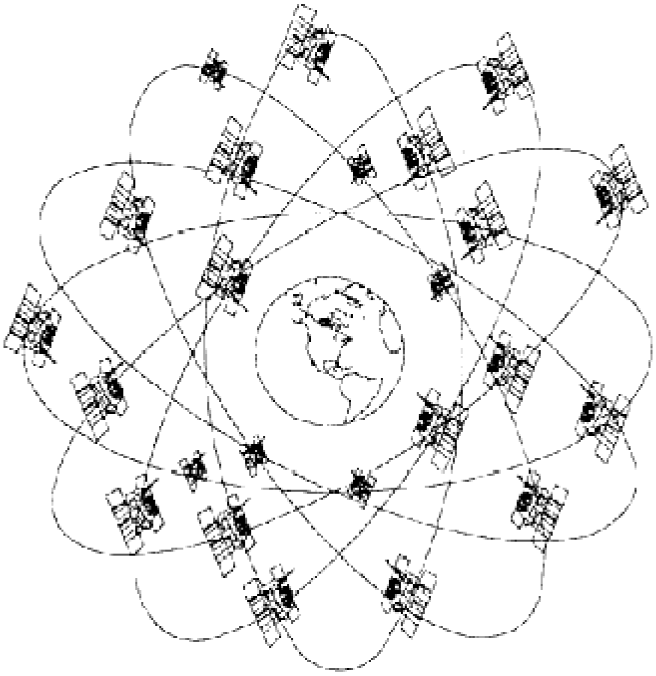


Figure 1.2 GPS constellation.

The Control Segment

The control segment is composed of all the ground-based facilities that are used to monitor and control the satellites. This segment is usually not observed by the user. This part of the system consists of tracking and uplink stations located around the world and a master control station located in Colorado, USA. These monitoring stations measure signals from the satellites, which are incorporated into orbital models for each satellite. The models compute precise orbital data (ephemeris) and clock corrections for each satellite. The Master Control station uploads ephemeris and clock data via the uplink stations to the satellites, which then send subsets of the orbital ephemeris data to GPS receivers over radio signals.

---

[1] A satellite could also be called a space vehicle (SV).

The User Segment

The user segment consists of the users and GPS receivers. A GPS receiver is a specialized radio receiver designed to listen to the radio signals being transmitted from the satellites. This process requires four satellites to compute the four dimensions: X, Y, Z (position), and time. With this ability, GPS has three main functions: navigation (for aircrafts, cars, ships, etc.), precise positioning (e.g., for surveying) and time and frequency dissemination (e.g., for telecommunications facilities). GPS receivers come in many different sizes, shapes, and price ranges. Some of them include integrated mobile phones and palm tops.

Before proceeding to the core of this paper, it is good to know some practical GPS applications. This is the topic of Section 2. Section 3 focuses on the car navigation system, which is the most well known commercial application of GPS. Section 4 explains how the car navigation system could use GPS to combat major traffic problems, which is a major issue on most urban highways. Then some conclusions and suggestions for future research are presented in Section 5.

2 APPLICATIONS OF GPS

2.1 AUTOMOBILE

The automobile industry is the largest user of GPS technology and uses it in the following ways.

2.1.1 Car Navigation

The most well-known commercial use of GPS is in car navigation systems. Before commencing a trip, a car driver just needs to feed in his destination in the GPS receiver's screen. The car GPS navigation system makes you find your way easily and quickly. While driving on an unfamiliar road or being caught by heavy traffic, the most convenient and fastest way to get to your destination is to use a GPS system. Through the satellites, car GPS navigation systems are able to show you other possible routes to your destination. Depending on the functional capabilities, the car GPS navigation system also gives you additional information, in which you might be interested. Such information may include the shortest or fastest road to your destination, the amount of fuel needed, the expected travel time, etc. With a car GPS navigation system, you will always enjoy your journey since you save time as well as you never get lost on the way to your destination.

2.1.2 Dynamic Vehicle Routing

Dynamic vehicle routing is seen in, for example, courier services (such as DHL) that have mini-vans, which go around a city collecting parcels and express packages. Requests for service arrive in real time at the courier's central dispatching office and are automatically relayed either by phone or some other device to a mini-van. This is done in the following way. A central computer or human scheduler carries out the processing of each request and decides which mini-van is selected to handle the request. This decision is based on GPS technology. Each mini-van is equipped with a GPS receiver, which is used by the scheduler to determine the positions of the different mini-vans. The scheduler then relays a request to the mini-van closest to the request. This saves not only vehicle miles, but the customer also gets serviced faster and the package is delivered to his destination sooner. The same scenario is applied to fleet management in a variety of industries, such as the taxi industry.

2.1.3 Tracking Rental Cars

Some car rental companies have installed GPS tracking on their rental cars to monitor mileage, speed, and location of their cars. These companies have levied heavy surcharges on customers who were found to be speeding or going outside the area covered by the lease agreement. This type of monitoring has been challenged, because the vehicle monitoring practices are usually not clearly explained to customers. In this sense it is clear that the rental companies are interfering with their customers' privacy.

2.1.4 Monitoring High Risk Auto Loans

Cars sold to some buyers with a poor credit history could be equipped with transmitting GPS units to monitor the whereabouts of the car and disable it in the event of a loan default.

2.2 AIRLINE SAFETY

"Boeing is proud of the outstanding safety record our industry has achieved, but we're never satisfied. We work literally day and night with our customers, pilots and government officials to make flying even safer."
– Alan Mulally (president and CEO, Boeing Commercial Airplanes)

There are two categories of airline safety, each using GPS technology to improve its performance. These are aviation safety and aviation security.

2.2.1 Aviation Safety

Aviation safety refers to the efforts that are taken to ensure that airplanes are free from factors that may lead to injury or loss. Of course, airplanes always have to be safe. If not, the manufacturer would not be in business for long. In order to make airlines safer, commercial airlines are working on new techniques called Radio NAVigation (RNAV) and Required Navigation Performance (RNP) [7]. These techniques use GPS technology [8] to improve approach and landing precision at airports that have limited ground-based navigation equipment [10]. With RNAV/RNP procedures, airlines can easily and effectively land in remote airports that often get socked in by stormy weather. The number of airlines that use GPS has increased over the years.

2.2.2 Aviation Security

Aviation Security is the intelligence gathering, pre-boarding procedures, and airport security that are necessary for a safe flight. It is related to passenger safety and not to airline safety. Since September 11, 2001, it is mainly aviation security that has been receiving urgent attention. The airline terror plot of August 10, 2006 in the UK has further increased this urgency. In some way, aviation security and aviation safety are interrelated.

Some commercial airlines allow passengers to use cell phones during flight. According to [10], engineers at the Carnegie Mellon University (USA) have found that cell phones, laptops, and other personal electronic devices can interfere with the airplane's critical electronics in contrast to what was previously thought. One of such devices is the airplane's GPS receiver, which is vital for a safe landing. It can be disrupted from normal operations. I therefore suggest that this opens up a very interesting topic for research, because of the fact that the use of GPS by airlines has been increasing significantly over the years while at the same time more airlines are considering the use of cell phones by passengers when an aircraft is in the air.

2.3 AGRICULTURE AND FARMING

GPS is being used in agriculture in the following ways.

2.3.1 Tractor Guidance

Farmers cannot put their tractors on autopilot. If they plow their fields with a recording GPS system, the tractor can be programmed to follow the same route – for cultivating, fertilizing, pest control, and harvesting. The programming of tractor routes greatly reduces farming costs.

2.3.2 Crop-duster Targeting

A crop-duster is a light airplane equipped for spraying crops with powdered insecticides and/or fungicides. Insects and fungi do not just attack a field according to a uniform distribution. Instead, they concentrate their attack on certain areas. Workers, strolling along the crops, can use a GPS to record the locations of insect and fungi problems. Instead of treating an entire field, the data can then be used by crop-duster pilots to selectively target the problem areas. This method results in a savings of time, fuel, insecticide, fungicide, and crop exposure to chemicals.

2.3.3 Tracking Livestock

The location of valuable animals on a large farm can be monitored by GPS transmitters attached to the animal's collar. When the animals are sent to the market, GPS transmitters can also be used to track their location.

2.3.4 Yield Monitoring

Estimates of yield variations across a property can be made using GPS. To do this, the property is divided into zones and the yield of each zone is estimated. During harvest, a GPS receiver is mounted on the harvesting machine. As the harvesting is being done on each zone, the yield is recorded and the GPS receiver determines the location of the zone. This information (yield and zone) is plotted on a map. The map can then be used to better understand the properties of the earth and for decision-making with regard to the next planting.

2.3.5 Soil Sampling

Collecting soil samples across a large area can be organized using GPS and mapping software. The sample locations can be way-pointed in the field and those waypoints are marked on the mapping software. Then, when the laboratory results are returned, the data can be plotted on the maps and decisions for soil treatment can be made for various parts of the property. This location information can save money and time by allowing variable rate applications and treating only those areas with a documented need.

2.4 LAW ENFORCEMENT

2.4.1 Tracking Convicted/Suspected Criminals

GPS bracelets can be placed on selected felons on parole to monitor their movements. For example, the system could monitor if criminals are staying away from the homes of their victims, are traveling to work each day, or are going near to schools. Such systems can be used to verify that certain restraining orders are being obeyed. GPS units have also been used to record and monitor the movements of crime suspects. Use of such information to aid in a conviction or an investigation has been challenged by defendants as an infringement of their privacy.

2.4.2 Online Crime Maps

Police departments in many densely populated cities use GPS to feed data into an online Geographical Information System (GIS). This online GIS allows the police to effectively create maps of the locations of different categories of crimes, which occur over a given period of time within a particular city. With this information, the public is kept well-informed. A clear example could be seen in some neighborhoods of Amsterdam with messages such as “Beware of pickpockets,” in the Center of Amsterdam or “Use it, Lose it,” in some parts of Amsterdam Zuid-Oost (South East) to indicate that you could lose your mobile phone to a thief upon using it.

2.4.3 Appealing Speed Tickets

A few individuals cited for speeding have produced GPS tracking information from their on-board GPS to appeal their ticket. Causes for this error could be that the police officer stopped the wrong car or his radar was malfunctioning.

2.5 SPORTS

2.5.1 Bicycle Racing

Ten riders in the 2004 Tour de France were equipped with GPS transmitters. This provided their location and speed at all times. In future years, all riders could be equipped with GPS and the progress of the race could be tracked on maps and topographic profiles could be made.

2.5.2 Athletic Training

Runners, cyclists, skiers, and other athletes who race across landscapes can use tiny GPS units while training to monitor their speed, distance covered, course difficulty, and more. Combined with a recording heart rate monitor the GPS units can provide valuable information about an athlete's condition and can be used to develop racing strategies as well.

2.5.3 GPS Golf

Many golf courses are installing GPS units on golf carts. These units can be used to estimate how far a ball has been hit and the distance to the green, and even to keep location-challenged golfers from getting lost on the links.

2.5.4 Soccer

Research is on the way on how GPS can be used by a coach to effectively know the positions of his players at each given moment in a football game. Maybe this could also be used as a device to determine if a critical situation, like an offside, has occurred or not.

3 CAR GPS NAVIGATION SYSTEMS

3.1 BRIEF REVIEW

A car navigation system combines GPS technology with a computer communication system to track a driver's progress on a digital map [4]. It provides services like the shortest route, fastest route, etc. on roadways, especially on motorways, which, on daily basis, are hardly free of congestion. This application is best termed *optimal car routing* and is the most widely used commercial application of GPS technology. Major companies like Motorola, TomTom, Nokia, Philips, Sony, etc. dominate this market.

CAR GPS combines telematics and telecommunications to do surveying and precise positioning. How this works is the content of the next section.

3.2 FUNCTIONING OF CAR GPS

While in continuous motion, each satellite in the GPS constellation continuously broadcasts radio signals in all directions. This information contains data about its orbit, equipment status, and the exact time. The car's GPS receiver consists of an antenna (usually mounted on the car or dashboard) and a computer with a display screen as shown in Figure 3.1 below.

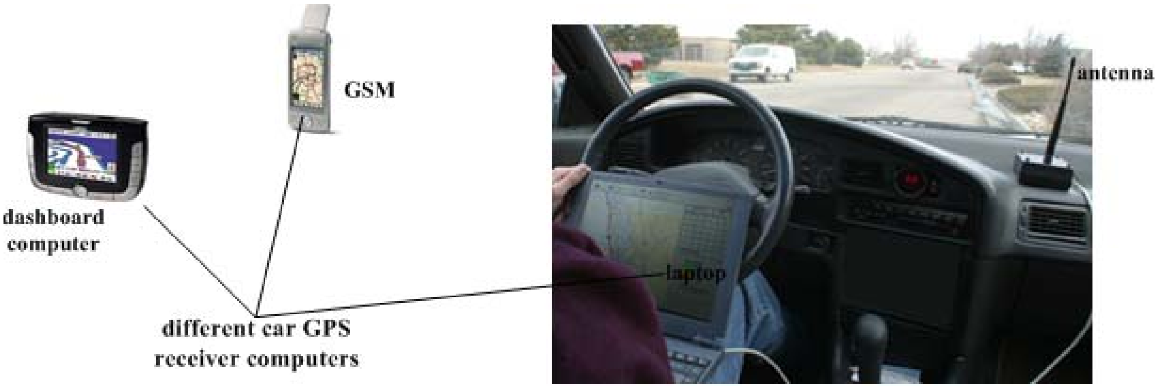


Figure 3.1. Components of the car GPS system.

The receiver computer has a digital road map in memory (in some cases, on a CD-ROM) that contains most roads in the country. Thus, all you need to do is just put in the details of your destination on the computer screen and then wait for the details the computer is going to give you.

3.2.1 Distance Calculation

The computer clock and the satellite clocks are synchronized. Since the center of gravity of the earth influences the rate at which the clocks change, the clocks are not synchronized with respect to each other but as a function of their velocity with respect to the center of the coordinate system, which is the center of the earth [12]. Packets from a satellite A contain a time stamp t_1 as part of the packets. Upon receiving these packets, the receiver clock reads its time t_2. The frequency of transmission is f and the wavelength of transmitted radio signals is λ. As a result of synchronization of the clocks, the computer determines the distance between the satellite and the car as follows:

The one-way delay of radio signals is: $t = t_2 - t_1$,
The speed of radio signals is: $v = f\lambda$,
The distance between satellite and car is: $speed * delay = f\lambda(t_2 - t_1)$.

3.2.2 Frequency Identification

Although all GPS satellites broadcast on the same frequency, the receiver computer is still capable to distinguish satellites using their broadcast frequency. The computer has identification codes for all the satellites stored in its memory. While transmitting, each satellite includes as part of the transmitted packets its identification code. The receiver computer then determines which transmission is from which satellite by matching the identification code with the ones stored in its memory. The computer uses this frequency identification together with the distance calculation above to determine its exact distance from that particular satellite. With frequency identification, GPS receivers that do not have antennas can still make reliable use of the tiny signals they receive.

3.2.3 Coordinate Determination

For a system with n coordinates, at least n lines need to intersect in order to specify the coordinates of a point. This is the principle that the GPS system uses. The origin of the GPS system is the center of the earth. Here, $n = 3$ (see Figure 3.2). Each satellite specifies a line. So, at least three satellites (lines) are needed to determine the coordinates (X, Y, Z) of the car. A fourth satellite is needed to specify the time.

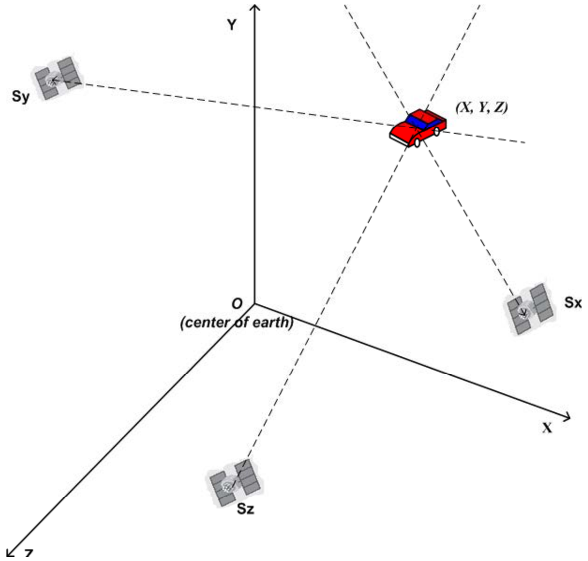


Figure 3.2 Three GPS satellites Sx, Sy, and Sz are used to geometrically determine the X, Y, and Z coordinates, respectively, of a car in the earth's 3D coordinate**.**

3.2.4 System Accuracy

The GPS satellites have very accurate clocks that use atomic vibrations (in the order of nanoseconds) as the fundamental unit of time. When the receiver clocks are considered, this is different. The receiver uses inexpensive quartz clocks. Since precise time measurements are a critical factor of the system, clock errors of about 0.001 seconds could cause a position error of about 320 km. The more satellites a receiver can see, the more accurate the measurements are and the fewer position errors one obtains. GPS receivers function with an accuracy of ±10 meters. This implies a clock accuracy of the order of nanoseconds.

3.3 GPS AND RELATIVITY

Relativity is essential for GPS clocks to work properly. Two theories come into play in order to determine the system accuracy.

- General Relativity (GR): According to GR, the clock of a moving object depends on the clock's position and velocity on the fields of nearby masses. Thus, the nearer the clock is to a massive object, the slower the clock ticks and vice versa.
- Special Relativity (SR): SR stipulates that the clocks of moving objects tend to tick slower than those of stationary objects. Thus, if an object moves with velocity v, then the rate at which its clock ticks while the object is in motion is given by

  $$f = \sqrt{1 - \frac{v^2}{c^2}} f_0, \tag{3.1}$$

  where c is the speed of light and f_0 is the rate at which the clock ticks when it is on a stationary object. In the situation where both objects are moving, v is the relative velocity between the two. The effect expressed by Equation (3.1) is referred to as *time dilation*.

The earth is a massive object compared to the size of each satellite. Thus, GR explains why the satellite clocks tick faster compared to the GPS receiver clocks on or near the surface of the earth. According to [24], a calculation using GR predicts that the clocks in each GPS satellite should get ahead of ground-based clocks by 45 microseconds per day. On the other hand, since the satellites are in continuous motion around the earth, SR predicts that the satellite clocks should fall behind ground-based clocks because of a slower ticking rate due to time dilation. The rate at which these clocks fall back is calculated to be 7 microseconds per day.

Combining these two effects would suggest that the satellite clocks are faster than ground-based clocks by 38 microseconds per day[2]. This presents a problem since GPS systems need a high degree of accuracy. In order to cancel the effect of GR, engineers slowed down the effect of the satellite atomic clocks before they were launched in such a way that the clocks would tick correctly when the satellites reached their proper orbit

---

[2] 45 (due to GR) – 7 (due to SR) = 38.

stations. From that time, they would start ticking at the same rate as the clocks on ground-based GPS stations. On the other hand, they counteracted the effect due to SR by equipping each satellite with a microcomputer, which performs the necessary corrective relativistic calculations.

4 MODELING ROAD TRAFFIC

4.1 MODELING OVERVIEW

Modeling the performance of a communication system involves either comparing different designs of the system or dimensioning the system [5]. In order to compare systems, we quantify the improvement obtained from a design option whereas in system dimensioning we determine the size for a given planned utilization. The benefit obtained from a performance modeling study has to outweigh its costs as well as the costs of the system.

After having defined the goal of an evaluation, we define the performance metrics and then proceed to the solution method. The solution method can either be one or a combination of

- Measurements of the real system. This sometimes involves the use of complex instruments.
- Discrete Event Simulation. A simplified model of the system is implemented in software. The performance of interest is then measured. This is as if it was measured in the real system. With simulations, rare and critical situations are easily seen [2].
- Mathematical Analysis. Here, the system is modeled mathematically and then numerical analysis is carried out. This is sometimes viewed as some form of simulation [5].

This paper treats the first and third solution methods.

4.2 PROBLEM DEFINITION

Traffic congestion in highways has proven to be a menace to society in many countries, such as the Netherlands and Belgium where cars are a major means of urban transport. Since 1980, mobility has enormously increased in the Netherlands. With its central location in Western Europe, Dutch highways serve as a transport hub to major European countries and as a result, despite the country's small size, there are more than 6 million private cars on Dutch roads [18].

In April 2006, the *Verkeers Informatie Dienst* (VID)[3] posted some interesting statistics on its website. According to these statistics, there have been 30% more traffic jams in March 2006 compared to March 2005. The majority of Dutch highways have 2 or 3 lanes. The biggest cause of road congestion is when a 3-lane road converges to 2 lanes. Although the given speed limit on Dutch highways is 120 km/h, more and more sections are being designated "80 km/h zones" to provide better traffic flow. However, this effort has put things on the wrong side. VID reports reveal that traffic jams have instead increased by 10% in these areas compared to the highways in the rest of the country. Only the A10 West yielded some fruitful results [19].

---

[3] VID = Dutch Traffic Information Service.

Although the government has introduced a technique called *spitsstroken*, where emergency lanes are turned into extra lanes during periods of heavy traffic, it turns out that these emergency lanes instead make things worse by increasing the length of traffic, especially, where they the lanes end.

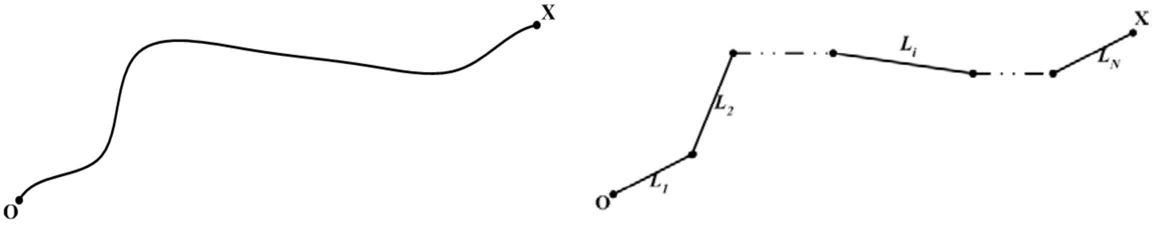
What we need to do is to design innovative solutions – solutions that make life easier for drivers by automatically selecting the right route and guiding the vehicle reliably to the desired destination. As a starting point for these solutions, we first need to specify what our performance measures are.

4.3 PERFORMANCE MEASURES

We cannot make measurements for road traffic when we do not know what to measure. First, we need to define our performance metrics and then try to find out which measurements we need to make.

Route Segmentation

In order to model traffic between two different locations, the complete route from source *O* to destination *X* is first determined. Once this route is determined, it is split into *N* different segments with segment *i* having length L_i, $i = 1, 2, \ldots, N$. Figure 4.1 below shows this. The measurements and analysis are then carried out on each segment.



a. Complete route from O to X

b. Division of route into N segments each of length L_i. For i, j in [1,N], L_i is not necessarily the same as L_j.

Figure 4.1 Segmentation of route from source O to destination X.

4.3.1 (Acceptable) Travel Time

The travel time t_L is the time it takes to traverse a segment of length *L*. This is either measured empirically or from simulation methods using some characteristics of the roadway segments.

A notion of travel time is the acceptable travel time t_{Lo}, which is the travel time, associated with a performance goal established for the transportation facility. Technical and non-technical groups have to come to an agreement to establish an acceptable travel time. Environmental issues, political concerns, community and economic activities influence such agreements. A more detailed analysis should differentiate the time period into peak and off-peak times and even the segment of the road in question into urban and suburban areas [1].

4.3.2 (Total) Delay

In computer networks, delay is the amount of time a packet takes to go from source to destination. In modeling road traffic, it has a slightly different meaning. It is the difference (d_L) between the travel time and the acceptable travel time on a road segment, i.e.,

$$d_L = t_L - t_{L_0} \,. \tag{4.1}$$

A notion used to measure the amount of traffic on a road segment is the *total delay.* This is the sum of delays for all vehicles traversing a segment during the time for which travel data are available. Suppose there are m vehicles traversing the segment at this time, then the *total delay* is given as

$$D_L = md_L \,. \tag{4.2}$$

The unit in which total delay is expressed is *vehicle-minutes*.

4.3.3 Segment Speed

The segment speed u is given as

$$u = \frac{L}{t_L}. \tag{4.3}$$

4.3.4 Delay Rate

The *delay rate* d_{rL} is the rate at which time is lost for a specified road segment. It is calculated as

$$d_{rL} = \frac{d_L}{L} = \frac{t_L - t_{L_0}}{L} \,. \tag{4.4}$$

4.3.5 Delay Index

More important is the dimensionless quantity called the *delay index.* It is an index used to compare congestion between different road segments. The greater the delay index, the less congestion is expected. It is given as

$$d_{RL} = \frac{d_L}{t_{L_0}} = \frac{t_L - t_{L_0}}{t_{L_0}} = \frac{t_L}{t_{L_0}} - 1 \,. \tag{4.5}$$

It is clear that when $t_{L_0} \le t_L$, congestion is least expected and vice versa.

Notice that it is sufficient to empirically determine t_L, and based on the agreed t_{L_0}, all other performance measures can be easily determined.

4.4 MAKING MEASUREMENTS

There exist two major techniques in making road traffic measurements [1]. These are called *roadside techniques* and *vehicle techniques*.

In roadside techniques, detection devices are placed at regular intervals along the route in question. These devices record the passing time of a vehicle at each of these points. Here, there is automatic vehicle matching where the license plate number of a vehicle is recorded as it passes each of the devices. These devices then relay the travel times to a central computer system. Travel times are calculated as the difference in time stamps between checkpoints. This is the technique used by most traffic management centers. However, it works on the assumption that vehicles do not make intermediate stops. This presents a major setback.

In vehicle techniques, the detection devices in the form of stopwatches, etc. are carried inside the vehicle. Two people are needed; a driver and a person making manual recordings. The recorder notes the time as the vehicle passes a landmark and the amount of time the vehicle spends in queues. This technique proves to be obsolete as it is labor intensive. Also, there is the possibility that landmarks could be missed.

The Use of GPS Technology

Over the years, researchers have been working on improving the two techniques mentioned above. In this course, GPS measurement techniques were developed (e.g., Guo and Poling, 1995; Zito et al., 1995; Quiroga and Bullock, 1999).

With GPS no humans or checkpoints are needed, thus fundamental errors are avoided. Only the GPS electronic devices are needed and there is a greater amount and variety of data obtained which could be used for purposes like queue lengths, delays, the average speed, models for planning, fuel consumption, and other performance measure computations [22].

Traffic management centers enhance GPS travel time measurements with an additional device called Differential GPS (DGPS). It receives signals from land-based differential correction stations. Figure 4.2 shows how DGPS is incorporated into GPS for better travel time data collection. The test vehicle has both a DGPS and a GPS antenna on its roof. The DGPS antenna connects to a DGPS, which transfers differential correction data to the GPS receiver. The GPS receiver uses this data to correct incoming signals from the GPS antenna. This corrected data is then output and stored into an in-vehicle portable computer. When the travel time run is completed, the portable computer information is downloaded to a data storage computer and can then be used for analysis.

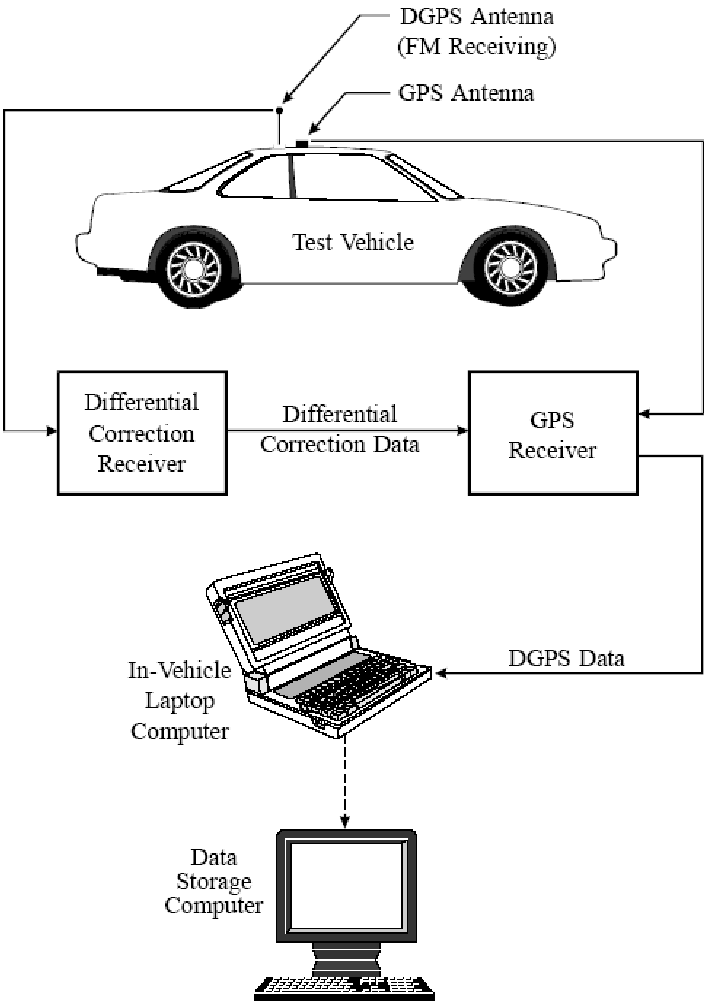


Figure 4.2 Equipment set up for GPS test technique[4].

4.5 SOLUTION APPROACH

4.5.1 Model Description

Each road segment consists of a number of links. Congestion is usually caused by irregularities (interruption) such as accidents, a vehicle break down, or when a link ends. In each road segment *L,* each vehicle occupies on average a small fragment *l.* According to [20], the best way to model highway traffic is to consider each link as a queueing model. In a link, each fragment *l* represents a server. Service begins as soon as a vehicle enters *l* and ends when the vehicle leaves it. Arrivals of vehicles into the link are assumed to be according to a Poisson process. A link can either be functional or broken. This phenomenon is described as a *batch service interruption*. Actually, when there is an interruption (accident, blockage, etc.) at some point on a link, other parts of the link are still functional. This is referred to as a *partial system breakdown*.

---

[4] Figure is adapted from reference [22].

Several queueing models have been brought forward to model road traffic. Heideman (1996) used an M/M/1 and an M/G/1 queueing model, Jain and Smith (1997) used an M/G/c/c queueing model, etc. All these models do not consider the fact that batch service interruptions occur. These interruptions are the major cause of congestion on roadways. Under batch service interruptions the best model to use is the M/M/c model [20].

Figure 4.3 shows a 3-link road. Each link represents an M/M/c queue. The parameters are described as follows:

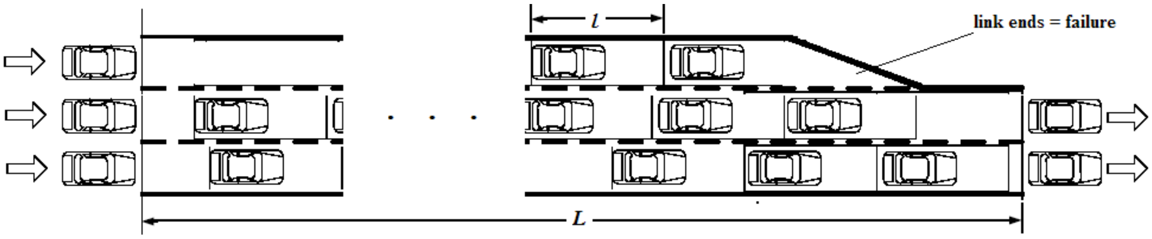


Figure 4.3 A three-link segment with interruption on one link.

- From the inter arrival times determined by the GPS measurements; the arrival parameter is determined, say λ. Arrivals occur according to a Poisson process.
- Also from GPS measurements, the average service time when the link is functional is determined, say $\beta = \frac{1}{\mu}$, with service rate, $\mu > 0$.
- When part of a link is interrupted, the average service time of all servers increases to $\frac{1}{\gamma}$, with $\gamma > 0$.
- The number of servers is $c = \frac{L}{l}$. Since c is required to be an integer value, we round c up or down to the nearest integer.

In doing a modeling study, some assumptions of the system in question have to be posed. For this model, it is assumed that

- The length of time in which a link is functional is exponentially distributed with mean $\frac{1}{\delta}$, with $\delta > 0$.
- The length of time in which a link is in interruption/failure is also exponentially distributed with mean $\frac{1}{\alpha}$, with $\alpha > 0$.
- There is sufficient buffer space in front of the link so that cars, which cannot get service, can wait in front of the link until a server is available.
- The queueing discipline is FIFO.

4.5.2 Markov Chain Analysis

The performance measures described in Section 4.3 obey the *insensitivity property*, i.e., they depend on the service time only through its mean. This reason, coupled with the fact that the underlying distributions are exponential, allows us to use a continuous time Markov chain.
Let for time $t > 0$, the random variables be defined as

$$X_1(t) = \text{number of cars present in a link at time t,}$$
$$X_2(t) = \begin{cases} 1 & \text{if the link is functional at time t,} \\ 0 & \text{if the link is in failure at time t.} \end{cases}$$

The process $X(t) = (X_1(t), X_2(t))$ has state space

$$I = \{(i, j) \mid i = 1, 2, \ldots, c \text{ and } j = 0, 1\}.$$

The one-state transition diagram is as shown in Figure 4.4.

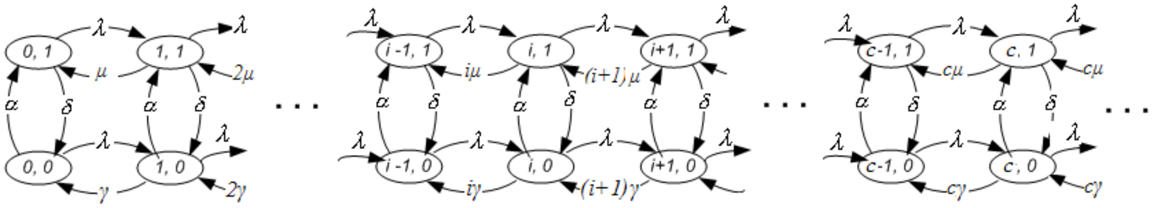


Figure 4.4 Transition rate diagram for the M/M/c queue with batch service interruption.

From the transition rate diagram we can derive the equilibrium distributions of the number of busy servers, $P_{(i,1)}$ and $P_{(i,0)}$, $i = 0, 1, 2, \ldots c$. These are given by the "*rate out = rate in*" condition as follows:

$$i = 0: \quad \begin{aligned} (\lambda+\delta)P_{(0,1)} &= \mu P_{(1,1)} + \alpha P_{(0,0)} \\ (\lambda+\alpha)P_{(0,0)} &= \gamma P_{(1,0)} + \delta P_{(0,1)} \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4.6)$$

$$0 < i < c: \quad \begin{aligned} (\lambda+\delta+i\mu)P_{(i,1)} &= (i+1)\mu P_{(i+1,1)} + \alpha P_{(i,0)} + \lambda P_{(i-1,1)} \\ (\lambda+\alpha+i\gamma)P_{(i,0)} &= (i+1)\gamma P_{(i+1,0)} + \delta P_{(i,1)} + \lambda P_{(i-1,0)} \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots(4.7)$$

$$i \geq c: \quad \begin{aligned} (\lambda+\delta+c\mu)P_{(i,1)} &= c\mu P_{(i+1,1)} + \alpha P_{(i,0)} + \lambda P_{(i-1,1)} \\ (\lambda+\alpha+c\gamma)P_{(i,0)} &= c\gamma P_{(i+1,0)} + \delta P_{(i,1)} + \lambda P_{(i-1,0)} \end{aligned} \quad \ldots\ldots\ldots\ldots\ldots(4.8)$$

The derivation of the equilibrium distribution in order to get a closed-form solution is done using the method of Generating functions.

4.5.3 The Method of Generating Functions

The generating function of a probability density function $P_{(i,k)}$ is given by

$$G_k(z) = \sum_{i=0}^{\infty} P_{(i,k)} z^i , \qquad |z| \le 1. \tag{4.9}$$

where

$$k = \begin{cases} 1 & \text{if the link is functional at time t,} \\ 0 & \text{if the link is in failure at time t.} \end{cases}$$

Thus, the generating function of the number of vehicles on the link is

$$G(z) = G_1(z) + G_0(z) = \sum_{i=0}^{\infty} P_{(i,1)} z^i + \sum_{i=0}^{\infty} P_{(i,0)} z^i \tag{4.10}$$

Multiplying both sides of the Equations (4.6), (4.7), and (4.8) by z^i and summing over *i* yields

$$\begin{aligned} &[\lambda(1-z) + \delta + c\mu(1-\frac{1}{z})]G_1(z) - \alpha G_0(z) = (1-\frac{1}{z})\mu \sum_{i=0}^{c-1}(c-i)z^i P_{(i,1)} \\ &[\lambda(1-z) + \alpha + c\gamma(1-\frac{1}{z})]G_0(z) - \delta G_1(z) = (1-\frac{1}{z})\gamma \sum_{i=0}^{c-1}(c-i)z^i P_{(i,0)} \end{aligned} \quad \ldots\ldots\ldots\ldots(4.11)$$

The Equations in (4.11), can be written in matrix form as

$$A(z)g(z) = b(z), \tag{4.12}$$

where

$$A(z) = \begin{pmatrix} \lambda(1-z) + \delta + c\mu(1-\frac{1}{z}) & -\alpha \\ -\delta & \lambda(1-z) + \alpha + c\gamma(1-\frac{1}{z}) \end{pmatrix},$$

$$g(z) = \begin{pmatrix} G_1(z) \\ G_0(z) \end{pmatrix},$$

$$b(z) = \begin{pmatrix} (1-\frac{1}{z})\mu \sum_{i=0}^{c-1}(c-i)z^i P_{(i,1)} \\ (1-\frac{1}{z})\gamma \sum_{i=0}^{c-1}(c-i)z^i P_{(i,0)} \end{pmatrix}.$$

Thus,

$$g(z) = A^{-1}(z)b(z) = \frac{adjA(z)}{\det A(z)} b(z).$$

Solving this equation and simplifying yields $g(z) = \begin{pmatrix} G_1(z) \\ G_0(z) \end{pmatrix}$,

where

$$G_1(z) = \frac{[\lambda(1-z)+\alpha+\delta+c\gamma(1-\frac{1}{z})]\mu\sum_{i=0}^{c-1}(c-i)z^i P_{(i,1)}}{\lambda^2 z^3 - \lambda(\lambda+c\gamma+\alpha+c\mu+\delta)z^2 + c(\lambda\gamma+\lambda\mu+c\gamma\mu+\alpha\mu+\gamma\delta)z - c^2\gamma\mu},$$

$$G_0(z) = \frac{[\lambda(1-z)+\alpha+\delta+c\mu(1-\frac{1}{z})]\gamma\sum_{i=0}^{c-1}(c-i)z^i P_{(i,0)}}{\lambda^2 z^3 - \lambda(\lambda+c\gamma+\alpha+c\mu+\delta)z^2 + c(\lambda\gamma+\lambda\mu+c\gamma\mu+\alpha\mu+\gamma\delta)z - c^2\gamma\mu}.$$

We now have an expression for the generating function of the number of vehicles on the link as $G(z) = G_1(z) + G_0(z)$.

Proposition 1 (Stability):

a) The M/M/c queue with batch service interruption is stable if the productivity

$$\rho = \frac{\lambda(\alpha+\delta)}{c\mu\alpha} < 1 \Leftrightarrow \frac{\lambda}{c\mu} < \frac{\alpha}{\alpha+\delta} \Leftrightarrow \lambda < \frac{\alpha}{\alpha+\delta}c\mu.$$

b) The M/M/*c* queue with partial system breakdown is stable if

$$\lambda < \frac{\alpha}{\alpha+\delta}c\mu + \frac{\delta}{\alpha+\delta}c\gamma \Leftrightarrow \frac{c(\alpha\mu+\delta\gamma)}{\lambda(\alpha+\delta)} < 1.$$

Proposition 2 (Analyticity):

$\forall\, z \in$ **IR**, $|z| \le 1$, $G(z)$ is analytic.

Let n(z) = numerator of G(z)

$$= [\lambda(1-z)+\alpha+\delta+c\gamma(1-\frac{1}{z})]\mu\sum_{i=0}^{c-1}(c-i)z^i P_{(i,1)} +$$
$$[\lambda(1-z)+\alpha+\delta+c\mu(1-\frac{1}{z})]\gamma\sum_{i=0}^{c-1}(c-i)z^i P_{(i,0)},$$

and d(z) = denominator of G(z)

$$= \lambda^2 z^3 - \lambda(\lambda+c\gamma+\alpha+c\mu+\delta)z^2 + c(\lambda\gamma+\lambda\mu+c\gamma\mu+\alpha\mu+\gamma\delta)z - c^2\gamma\mu,$$

i.e., $G(z) = \frac{n(z)}{d(z)}$.

Claim: $\exists\ z_0,\ |z_0| \le 1$ such that $n(z_0) = 0$.

Proof: We have
$d(z)=\lambda^2 z^3-\lambda(\lambda+c\gamma+\alpha+c\mu+\delta)z^2+c(\lambda\gamma+\lambda\mu+c\gamma\mu+\alpha\mu+\gamma\delta)z-c^2\gamma\mu$.
Let's define $g(z)=\lambda(\alpha+\delta)z-c(\alpha\mu+\delta\gamma)$.
It is clear that $g(z)$ has just one root $z_g=\dfrac{c(\alpha\mu+\delta\gamma)}{\lambda(\alpha+\delta)}<1$, according to Proposition 1 (b).

Also, we notice that on the boundary, $z = 1$, $|d(1)-g(1)|=0<|g(1)|$. Thus, by Rouché's Theorem, $d(z)$ also has one root z_0 in the unit circle i.e. $|z_0|\le 1$. By Proposition 2, z_0 is also a root of $n(z)$. This proves the claim. Thus,

$$n(z_0)=0. \tag{4.13}$$

Remark
For given values of α, δ, γ, λ, μ, and c, we simply determine z_0 by solving the cubic equation

$$d(z)=\lambda^2 z^3-\lambda(\lambda+c\gamma+\alpha+c\mu+\delta)z^2+c(\lambda\gamma+\lambda\mu+c\gamma\mu+\alpha\mu+\gamma\delta)z-c^2\gamma\mu=0.$$

One of the roots lies inside the unit circle. This root is z_0 i.e. $|z_0|\le 1$.

By definition,

$$G(1)=\sum_{i=0}^{\infty}P_{(i,1)}+\sum_{i=0}^{\infty}P_{(i,0)}=1.$$

But from (4.10)

$$G(1)=G_1(1)+G_0(1)=\frac{(\alpha+\delta)\left\{\mu\sum_{i=0}^{c-1}(c-i)P_{(i,1)}+\gamma\sum_{i=0}^{c-1}(c-i)P_{(i,0)}\right\}}{\alpha(c\mu-\lambda)+\delta(c\gamma-\lambda)}=1,$$

$$\Longleftrightarrow \qquad \mu\sum_{i=0}^{c-1}(c-i)P_{(i,1)}+\gamma\sum_{i=0}^{c-1}(c-i)P_{(i,0)}=\frac{\alpha(c\mu-\lambda)+\delta(c\gamma-\lambda)}{(\alpha+\delta)}. \tag{4.14}$$

To solve for $P_{(0,1)}$ and $P_{(0,0)}$, we need take $c=1\Rightarrow i=0$. Equation (4.14) thus yields

$$\mu P_{(0,1)}+\gamma P_{(0,0)}=\frac{\alpha(\mu-\lambda)+\delta(\gamma-\lambda)}{(\alpha+\delta)}. \tag{4.15}$$

Similarly, from Equation (4.13) we obtain

$$\left[\lambda(1-z_0)+\alpha+\delta+\gamma\left(1-\frac{1}{z_0}\right)\right]\mu P_{(0,1)}+\left[\lambda(1-z_0)+\alpha+\delta+\mu\left(1-\frac{1}{z_0}\right)\right]\gamma P_{(0,0)}=0 \tag{4.16}$$

Equations (4.15) and (4.16) represent two simultaneous equations in $P_{(0,1)}$ and $P_{(0,0)}$. Solving them gives

$$\boxed{P_{(0,1)} = \frac{-bd}{(a-b)\mu} \text{ and } P_{(0,0)} = \frac{ad}{(a-b)\gamma}} \tag{4.17}$$

with $a \neq b \Leftrightarrow \gamma \neq \mu$,

where,

$$a = \left[\lambda(1-z_0) + \alpha + \delta + \gamma\left(1 - \frac{1}{z_0}\right)\right],$$
$$b = \left[\lambda(1-z_0) + \alpha + \delta + \mu\left(1 - \frac{1}{z_0}\right)\right],$$
$$d = \frac{\alpha(\mu-\lambda) + \delta(\gamma-\lambda)}{\alpha+\delta}.$$

Notice that $P_{(0,1)}$ and $P_{(0,0)}$ are expressed in terms of the known transition rates α, δ, γ, λ, and μ.

The most important aspect is to derive the values of $P_{(0,1)}$ and $P_{(0,0)}$. From these, the rest of $P_{(i,1)}$ and $P_{(i,0)}$, i = 1, 2, … c-1, can be determined recursively, which are then used to solve Equation (4.12), thus obtaining the generating function $G(z)$ for the number of cars present on the link. The computational complexity increases with the value of c. The expected number of vehicles present on the link is then given by:

$$E(X) = \left.\frac{dG(z)}{dz}\right|_{z=1}. \tag{4.18}$$

In reality, most road links can serve a large number of vehicles. The computational complexity for large values of c gives rise to the need for simpler solutions. To this end, we use the M/M/∝ queue to approximate the M/M/c queue if c is large enough. Baykal-Gursoy et al.[5] (2004) stipulate that this approximation is valid if the acceptable relative error $e_{relative}$ is less than 15% where

$$e_{relative} = \frac{\left|E(X)_{M/M/c} - E(X)_{M/M/\infty}\right|}{E(X)_{M/M/c}} \times 100\%.$$

---

[5] The investigation of the use M/M/∝ system to approximate an M/M/c system for large values of c is given in [20]. This is done using INTEGRATION^{TM}, a traffic simulation package.

A derivation of the generating function G(z) for the number of vehicles on the link is given in [21]. From this, the expected number of vehicles, $E(X)$, on the link is determined using (4.18). This yields

$$E(X) = \frac{\lambda}{\mu} + \frac{\lambda\delta(\mu-\gamma)}{\mu^2(\alpha+\delta)}\left(1 + \frac{(\delta+\mu)(\mu-\gamma)}{\alpha\mu+\delta\gamma+\mu\gamma}\right). \tag{4.19}$$

Using Little's formula, the expected travel time t_L is given by

$$\boxed{t_L = \frac{E(X)}{\lambda} = \frac{1}{\mu} + \frac{\delta(\mu-\gamma)}{\mu^2(\alpha+\delta)}\left(1 + \frac{(\delta+\mu)(\mu-\gamma)}{\alpha\mu+\delta\gamma+\mu\gamma}\right)} \tag{4.20}$$

We have finally arrived at a closed-form solution. With this expression for t_L, all the other performance measures given in Section 4.3 can be easily determined as long as we know the agreed value for the acceptable travel time t_{Lo}.
Note that the parameters μ and γ are service rates. They are associated with the respective service times , $\beta = \frac{1}{\mu}$ and $\beta' = \frac{1}{\gamma}$.

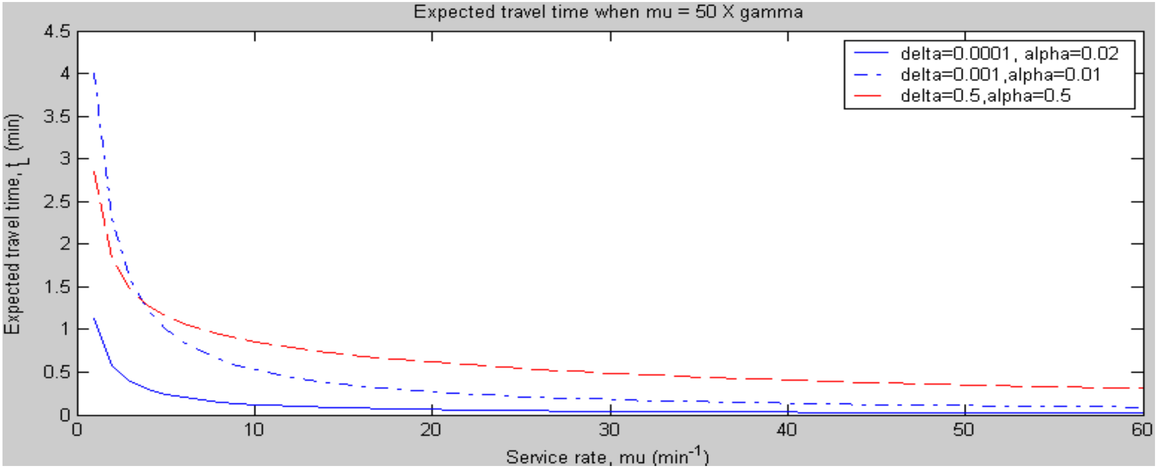
In the appendix are Matlab plots that show how t_L behaves as a function of some of the parameters. The given values of δ and α are in min^{-1}. From the plots, we notice that the expected travel time decreases with an increase in service rate (decrease in service time) and vice versa.

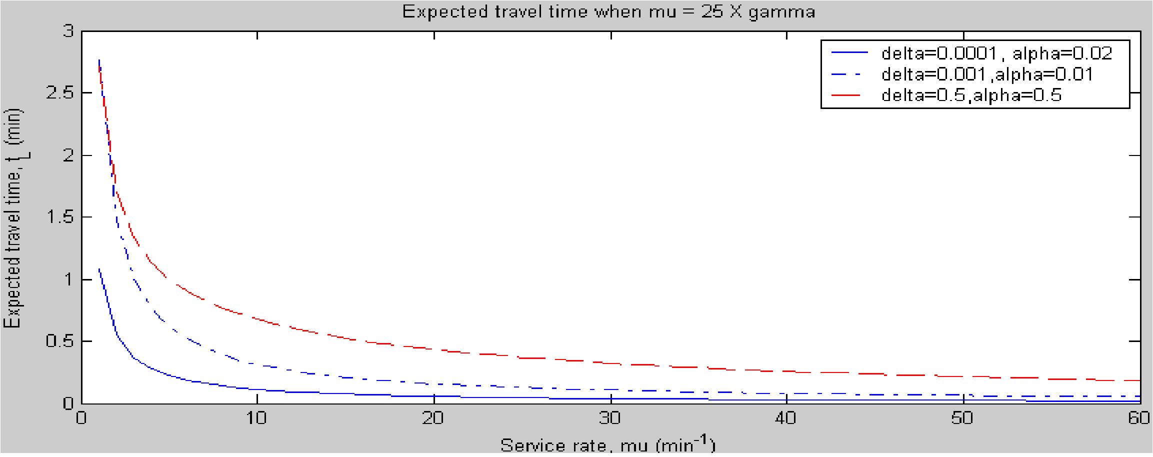
5 CONCLUSION AND FUTURE RESEARCH

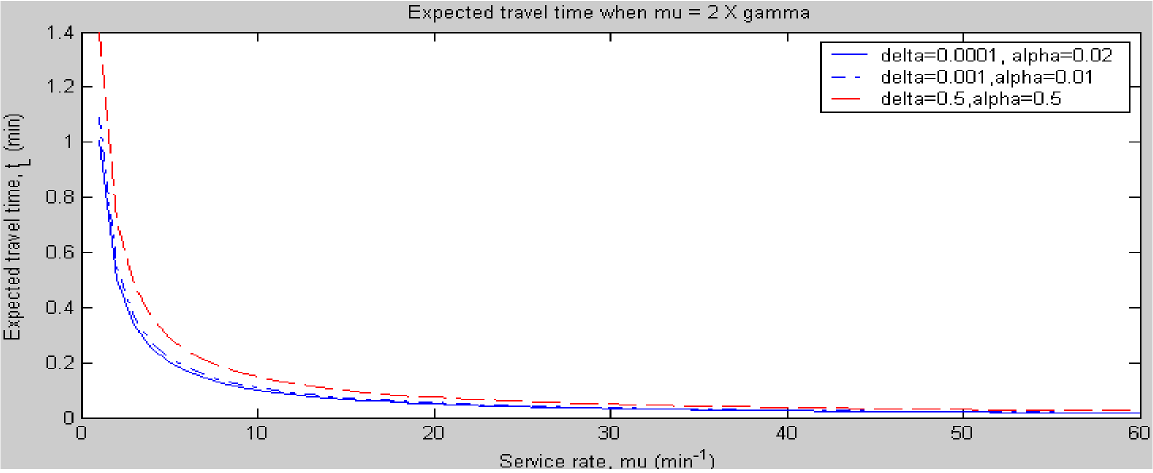
In this paper, I have presented the structure and some applications of GPS systems. The core of these applications is car navigation. The need for optimizing car navigation has been explained. How this optimization can be achieved has also been elaborated on. This involves defining some performance measures. The basis of all the performance measures are the travel time t_L and the acceptable travel time t_{Lo}. From these two quantities, all the other performance measures can be obtained. While t_{Lo} involves an agreement between technical and non-technical teams, t_L involves only technical teams. To this end, a closed-form solution has been obtained for t_L using stochastic analysis.

From the travel time and other performance measures, drivers would be able to know the optimal route to their respective destinations. For further research, I suggest that traffic management centers could use GPS technology for scheduling purposes as well. In the event where there is heavy traffic on a given route segment, the travel time increases. In addition to informing incoming cars at the route segment (usually in the form of electronic boards) about this traffic situation, they could also provide them other possibilities of reaching their respective destinations in relatively a little amount of time. In this way, fuel consumption, air pollution, etc., can be reduced.

APPENDIX **Plots of expected travel time t_L vs service rate μ.**







BIBLIOGRAPHY

[1] **Quiroga, C.A.:** "*Performance Measures and Data Requirements for Congestion Management Systems*", Transportation Research Part C, 2000.

[2] **Unsal, C., Sukthankar, R., and Thorpe C.:** *"Functional Sensor Modeling for Automated Highway Systems and Simulation",* SPIE Technical Conference on Intelligent Transportation Systems, 1997.

[3] **Tijms, H.C.:** *"A First Course in Stochastic Models",* Wiley, 2003.

[4] **Dommety G. and Jain, R.:** *"Potential Networking Applications of Global Positioning Systems".*

[5] **Le Boudec J.Y.:** *"Performance Evaluation of Computer and Communication Systems".* Lecture notes of Performance Evaluation course, EPFL, 2006.

[6] **Storey, B. and Holtom, R.:** *"The Use of Historic GPS Data in Transport and Traffic Monitoring",* Congestion Monitoring, Mott MacDonald, 2003.

[7] **Williams, J.:** *"Performance Based Navigation (RNAV and RNP): U.S. Strategic Plans",* Presentation to Eurocontrol/FAA Workshop "Implementing RNP", Toulouse, France, 2005.

[8] **Haggart, C.:** *"How GPS works: An Introduction",* http://www.american-tiger.com/index.asp?PageAction=Custom&ID=22.

[9] **Kielmann, T.:** *"Computer Network Measurements",* Lecture notes of Performance Analysis of Communication Networks course, VU Amsterdam, 2005.

[10] **Hanser, K. and Proux, J.:** *"Aviation Safety and Aviation Security",* http://www.boeing.com/commercial/news/feature/safety.html.

[11] **Motorola:** *"Motorola GPS Products",* Oncore Engineering Notes, 1998.

[12] **Hatch, R. R.:** *"Relativity and GPS",* http://www.egtphysics.net/GPS/RelGPS.htm.

[13] **Psaraftis, H.N.:** *"Dynamic Vehicle Routing",* Annals of Operations Research, 1995.

[14] **Dana, P.H.:** *"Global Positioning System (GPS) Time Dissemination for Real-Time Applications",* Kluwer Academic Publishers, 1997.

[15] **Levinson, D., Gillen, D., and Chang, E.:** *"Assessing the Benefits and Costs Intelligent Transportation Systems: The Value of Advanced Traveler Information Systems",* California PATH Research Report, 1999.

[16] *“GPS Navigation System”,* http://www.autoaccessorystore.com/resource/gps-navigation/gps-navigation.asp.

[17] **Thomas, R.D.:** *“Covert Surveillance GPS Car Tracking: Huge Selections!”* http://www.pimall.com/nais/tracking.html.

[18] **Arthur:** *“Traffic Jams on Dutch Highways Continue to Grow”,* Life in Holland. http://www.wolfstad.com/2006/04/traffic-jams-on-dutch-highways-continue-to-grow/.

[19] *“Veel drukkere eerste drie maanden, situatie 80 kilometer zones verslechterd”,* http://www.vid.nl/Nieuws/article/VID.2006.092.01.

[20] **Baykal-Gursoy, M., Xiao, W., and Ozbay, K.:** *“Modeling Traffic Flow Interrupted by Incidents”,* Rutgers University.

[21] **Baykal-Gursoy, M., Xiao, W.:** *“Stochastic Decomposition in M/M/∝ Queues with Markov Modulated Service Rates”,* Kluwer Academic Publishers, 2004.

[22] **Turner, S.M.; Eisele, W.L.; Benz, R.J., and Holdner, D.J.:** *“Travel Time Data Collection Handbook”,* Texas Transportation Institute, 1998.

[23] **Daigle, J.N.:** *“Queueing Theory with Applications to Packet Telecommunication”,* Springer, 2004.

[24] **Pogge, R.W.:** *“Real-World Relativity: The GPS Navigation System”,* http://www-astronomy.mps.ohio-state.edu/~pogge/Ast162/Unit5/gps.html.

[25] **Cohen, J.W., and Down, D.G.:** *“On the Role of Rouché's Theorem in Queueing Analysis”,* CWI Amsterdam, 1995.