Simplified traffic responsive signal control method for developing large cities

DISSERTATION

submitted to

Graz University of Technology,
Faculty of Civil Engineering

for the degree of
Doktor der technischen Wissenschaften (Dr. techn.)

presented by

Fatih Gündogan (M.Sc.)

born February 10th, 1977 in Alaca, Turkey

Supervisor: Univ.-Prof. Dr.-Ing Martin Fellendorf, TU Graz
Reviewer: Assoc. Prof. Dr. Y. Sazi Murat, Pamukkale University, Turkey

Oral examination: September 19th, 2012

Acknowledgements

This dissertation was completed during my activities as a doctoral student at the Institute for Highway Engineering and Transport Planning at Graz University of Technology in Austria.

I would like to acknowledge many people for their contribution to the successful completion of this dissertation study. First, I wish to express my deep gratitude towards my supervisor Professor Martin Fellendorf, for his valuable guidance, stimulating suggestions and invaluable support through my doctoral program. I also wish to thank my other supervisors Professor Yetiş Şazi Murat from Pamukkale University, in Turkey and Professor Gerd Sammer from University of Natural Resources and Life Sciences in Vienna for agreeing to act as supervisors.

This doctoral study was supported by Istanbul Metropolitan Municipality and its affiliate company ISBAK Inc. I owe thanks to the Mayor Dr. Kadir Topbas in the name of the municipality and to Mr. Kasim Kutlu the general manager of ISBAK Inc.

I am also grateful to Professor Kurt Fallast, for his stimulating discussions about transportation planning in Austria during my stay at the institute. My special thanks are extended to all my colleagues at the Institute for Highway Engineering and Transport Planning for their continuous support and mutual exchange of expertise.

I am also thankful to Gertrud Mawid for her excellent administrative support at the institute and to Peter Kollegger for his continuous support at the institute library.

I also owe thanks to my former manager Yusuf Kotil and former colleagues for their great effort and support by collecting the data in Istanbul and special thanks to my former colleague Abdullah Karadağ for his great support especially in programming languages and other technical ideas.

A special thanks to my colleague and also my friend Mustafa Erciyas for his continuous support with his ideas and discussions.

My deepest gratitude and appreciation go to my parents for their continuous support in during my life. Finally, I would like to thank to my wife for her encouragement, endless support and patience during my doctoral study,

…and to my daughters Elif and Esma who probably will not remember our stay in Austria.

Fatih Gündoğan

Abstract

Due to the explosion of population growth in urban areas the number of large and mega cities are increasing mostly in the developing or emerging countries. The large cities beef not only about traffic problems, but the traffic problems come to the forefront. Lack of road infrastructure, inadequate public transport facilities and uncoordinated transport and urban planning can be found frequently in large cities of emerging countries. Transport problems are getting even worse by lacking traffic education and undisciplined driving behavior.

Traffic signals are indispensable control method of urban areas. Most probably, this will not change even if the cities grow further. However, if the traffic signals are not correctly operated to cope with the increasing and fluctuating traffic demand, the governments produce neither economically nor environmentally solutions (i.e. replacement of signalized intersections with multistory interchanges). Therefore, the municipal authorities need an effective and feasible traffic signal control system to cope with the growing and fluctuating traffic demand and to prevent irreversible decisions about road infrastructures.

This study focuses on traffic signal control in developing large cities and consists of two parts. The first part describes the current situation of traffic signal control in developing large cities. For this purpose, two preliminary studies were conducted during this dissertation. The first preliminary study compares current traffic signal control in developing large cities with developed ones. The results indicate that there is perceptible difference both in quality and quantity in traffic signal control.

The second preliminary study deals with the usage of capacity in developing large cities. Therefore, saturation flow measurements are conducted in several districts of Istanbul. The intersections are grouped in three categories according to their geometric standards. The objectives of this preliminary study were to analyze the saturation flow variations in different districts and to examine several impact factors on saturation flow rate. According to results, the existing difference on saturation flow rates between intersections increases with increasing number of minibuses. Minibuses are special public transport mode without regular stops. If signalized intersections have many pedestrian activities and a large number of minibus, the intersection capacity drops by half.

The second part of this dissertation consists of strategies for developing large cities to deal with signal control related problems. A multi-layer traffic signal control architecture is proposed. The upper level of system contains data management and an optimization platform. The lower level provides real-time traffic signal control for predetermined road networks. A pattern recognition method is applied using multi-layer perceptrons to recognize the present traffic state and associate it with predefined traffic states. A suitable predefined fixed time signal plan is selected based on the traffic flow pattern. The system requires only few detectors for the traffic state recognition. In comparison with the existing traffic adaptive systems the proposed system recognizes the traffic state with much less system detectors. Since one of the main advantages of the artificial neural networks is its ability to learn, the system can continuously learn with the new data that is collected during the operation. If a new traffic state appears, the user will be informed and new timing plans are prepared for the new traffic state.

The system is applied in two case studies. First the system is examined and the parameters of the algorithm are adjusted on a hypothetical three-intersection arterial network. Here, TRANSYT was used for the optimization of timing plans of each scenario. Sensitivity analysis was conducted for ranking of detectors to reduce the number of detectors further. Secondly, a real arterial network in

historical district of Istanbul is used to test the proposed system. Since detectors are not available on this network, the data are collected manually. The macroscopic planning tool VISUM was used to calculate flow paths and to optimize the timing plans. Both case studies indicate noticeable improvements in traffic flow and reduction in delay and travel times.

Kurzfassung

Aufgrund des Bevölkerungswachstums in städtischen Agglomerationen steigt die Anzahl von Groß- und Megastädten in Entwicklungs- und Schwellenländern. Zwar sind Verkehrsprobleme nicht die einzigen Probleme von Großstädten, sie haben jedoch einen hohen Stellenwert. Groß- und Megastädten in Entwicklungsländern, wie mangelnde Straßeninfrastruktur, unzulängliche ÖPNV-Infrastruktur oder Probleme zwischen Stadt- und Verkehrsplanung sind häufig anzutreffende Defizite in den Groß- und Megastädten von Entwicklungs- und Schwellenländern. städtische Planungsprobleme oder unzulängliche und unbequeme ÖPNV-Infrastruktur stellen eine hohe Belastung für den Verkehr dar. Zusätzlich verstärken mangelnde Verkehrserziehung und undiszipliniertes Fahrverhalten der Verkehrsteilnehmer die Verkehrsprobleme.

Lichtsignalanlagen sind unentbehrliche Steuerungsmethode im städtischen Verkehr. Das wird sich höchstwahrscheinlich auch in Zukunft nicht ändern, selbst wenn die Städte weiter wachsen. Um wegen der mangelhaften Lichtsignalsteuerung verursachten volkswirtschaftlichen Schaden zu reduzieren, versuchen die Stadtverwaltungen in den Schwellenländern weder wirtschaftlich noch städtebaulich vertretbare Lösungen, wie Umbau der mit Lichtsignalanlage gesteuerten Kreuzungen zu einem niveaufreien Knotenpunkt anzuwenden. Deswegen mit der wachsenden und schwankenden Verkehrsnachfrage zurechtzukommen, brauchen die Stadtverwaltungen effektive Lichtsignalanlagensteuerung- bzw. Verkehrsmanagementsysteme.

Die vorliegende Arbeit beschäftigt sich mit Lichtsignalsteuerung in Großstädten, mit besonderem Augenmerk auf Entwicklungs- und Schwellenländer. Die Arbeit besteht aus zwei Teilen: Der erste Teil beschreibt die gegenwärtige Situation der Lichtsignalsteuerung in Großstädten. Für diesen Zweck wurden zwei Vorstudien durchgeführt. Die erste Studie vergleicht den Zustand von Lichtsignalsteuerungen einiger Großstädte von Schwellenländern mit Städten in Industrieländern. Die Ergebnisse zeigen an, dass es wahrnehmbare Unterschiede in Qualität und Menge der Lichtsignalsteuerungen gibt.

Die zweite Vorstudie untersucht die Nutzung der Kapazität von Lichtsignalanlagen in Großstädten von Entwicklungsländern. Zu diesem Zweck wurden Untersuchungen der Sättigungsverkehrsstärke in mehreren Bezirken von Istanbul durchgeführt. Die Knotenpunkte werden entsprechend ihrer geometrischen Standards in drei Gruppen unterteilt. Die Ziele hier sind, die Schwankungen von Sättigungsverkehrsstärke in verschiedenen Bezirken zu bestimmen und jene Faktoren, welche besonders die Sättigungsverkehrsstärke beeinflussen, zu prüfen. Den Ergebnissen zufolge, nimmt die Sättigungsverkehrsstärke ab, wenn die Anzahl von Kleinbussen steigt. Kleinbusse sind einer Sonderform des ÖPNV und verkehren weder nach festem Fahrplan noch mit genau festgelegten Haltestellen. Wenn hohe Fußgängeraktivitäten und zahlreiche Kleinbuslinien zusammenkommen, halbiert sich die Kreuzungskapazität.

Im zweiten Teil der Arbeit werden Lichtsignalsteuerungsstrategien für Entwicklungsgroßstädte entwickelt. Das Modell besteht aus zwei Ebenen: Die obere Ebene dient der Datenverwaltung und Optimierung, während die untere Ebene eine echtzeitfähige Lichtsignalsteuerung für vorherbestimmte Verkehrsnetze liefert. Für die untere Ebene wird ein Mustererkennungsverfahren mit Hilfe von künstlichen neuronale Netzen (KNN) angewandt, um die tatsächliche Verkehrssituation zu erkennen und geeignete Festzeitprogramme auszuwählen. Das System verwendet nur wenige Detektoren für die Erkennung der Verkehrssituation. Im Vergleich zu vorhandenen verkehrsabhängigen Lichtsignalsteuerungsverfahren, erkennt das vorgeschlagene System die

Verkehrssituation mit nur einigen Systemdetektoren. Da einer der Hauptvorteile der künstlichen neuronale netzen ist die Lernfähigkeit, kann das System stetig mit den neuen Daten lernen. Wenn eine neue Verkehrssituation erscheint, wird der Benutzer informiert sein, und neue Signalzeitpläne für die neue Verkehrssituation werden vorbereitet.

Das System wurde durch zwei Fallstudien analysiert. Zuerst wurde das System auf einem hypothetischen Straßenzug mit drei signalisierten Knotenpunkten geprüft und die Parameter des Algorithmus wurden eingestellt. Hier wurde TRANSYT zur Optimierung von Signalzeitenplänen für jedes Szenario verwendet. Weiterhin wurde ein reales arterielles Netz in einem historischen Bezirk von Istanbul verwendet, um das vorgeschlagene System zu testen. Da auf diesem Netz keine Detektoren verfügbar sind, wurden die typischen Verkehrsbelastungen erhoben. Das makroskopische Planungsinstrument VISUM wurde benutzt, um Routenbelastungen zu berechnen und die Signalzeitenpläne zu optimieren. Beide Fallstudien zeigen deutliche Verbesserungen bezüglich des Verkehrsflusses und der Verminderung der Anzahl der Halte, der Wartezeit und der Gesamtreisezeit an.

Table of Contents

1 Introduction

1.1 Background

Not only in developing world but also in developed countries, urbanization exists as one of the prevailing issues of this century. The global proportion of urban population raised from 29% in 1950 to 49% in 2007 and it is expected in 2025 about 60% of world population will live in urban areas [UN, 2008]. As a consequence of urbanization, number of large cities increases. The population of large cities grows and their boundaries expand. Apart from effects on social and economic issues, urbanization has a significant effect on increase of traffic congestion.

One of the biggest tasks of city governments of large cities both in developing and developed world is congestion mitigation. However, large cities in developing countries have also to consider increasing in population, in number of passenger cars and fast expanding municipal area. Traffic congestion can be mitigated by employing transportation demand management strategies, traffic management strategies or by new road and interchange building. Despite developing authenticated traffic and travel demand management strategies, unfortunately, the cities exclusively in the developing countries spend their budget funds on construction of new automobile roads and interchanges in urban area. Well-designed and updated traffic signal operations can also be good opportunity for congestion mitigation beside traffic demand management strategies.

Since its first installation over hundred years ago in London, the development of signal controllers has been in line with the general technological development. This development is directed according to territorial or regional experience. Therefore, similar but diverse approaches in traffic signal control strategies exist in several regions. Generally, in practice signal control strategies are considered in three main groups: fixed time, traffic actuated and traffic responsive/adaptive control. A well-designed and coordinated pre-timed signal plan can work well enough as long as the traffic situation is not changed. However, frequent occurring incidents and big events can change the repeated traffic pattern in large cities. As a simplified definition, traffic actuated control queries demand and extends green time according to current demand, while in traffic adaptive control, demand is determined (reactive) or forecasted (proactive) and the timing plan is optimized via an online mode.

However, both real-time approaches require a high-density sensor mesh unit and a powerful system hardware, which increases the installation and maintenance costs. Additionally, know-how for first parameter settings and further adjustments are other requirements of these systems.

General problem for large cities in traffic signal control is the fluctuating traffic. This is due to stochastic demand variations, traffic accidents or big events. The traffic signals should react against these variations in traffic demand, while bearing in mind that to keep the system in service. Another problem is that the maintenance is difficult in those cities, where much road works and digging occurs without coordination between municipal units.

This study focuses on traffic signal control in developing large cities, describes the situation and problems based on preliminary studies, and presents a low-cost traffic responsive system in order to contribute to the solution of traffic congestion problem in those cities. Although the proposed method is based on experiments with microscopic traffic flow simulation, the algorithm itself is ready for real world implementations.

1.2 Objectives

Most of the existing traffic adaptive and traffic responsive systems are more sophisticated and quite expensive. They require relatively quite high sensor meshes to select or to generate timing plans for current traffic situations and also require too many parameter settings during implementation in the field.

The main objective of this study is to develop a simplified traffic signal control method for developing large cities using relative fewer detection points. In order to achieve this goal, the following set of objectives has to be achieved:

1. Investigate the current situation of developing large cities in traffic signal control according to signal equipment, architectures and control methods in comparison with developed large cities. Additionally, to investigate traffic related problems (infrastructure and behavior of road users) and its impact on current traffic situation in those cities.
2. Develop a framework that consists of planning and operation tasks as well as sort-term and long-term traffic control strategies for developing large cities.
3. Develop a simplified traffic responsive control method for developing large cities. The method should consider the user knowledge and also hardware/software situation in those cities. The method should not bring additional cost for hardware and software to ensure a traffic responsive control. Therefore should also use less number of detection point. Then, generally, it is difficult to keep the detectors in service. The system should also be able to learn itself from detected traffic situations.
4. Evaluate the proposed approach using microscopic traffic flow simulation in comparison with pre-timed and optimized control in a real-traffic network.

1.3 Overview of the study

This dissertation is organized as follows. Chapter 2 presents the basic elements of traffic signal control. The traffic signals are described in detail according to their types, methods and strategies. Various commercial offline and online signal control software and methods are also summarized in this chapter. The chapter also provides an overview of previous research projects, and studies that used artificial intelligence (AI) in traffic signal control and traffic state estimation.

Chapter 3 provides a literature review on pattern recognition, which is used for proposed signal control method. Pattern recognition problems can be solved using several approaches such as template matching, statistical, syntactical or neural networks. The neural network approach that used in this study is described in detail.

Chapter 4 presents the findings and highlights of preliminary studies about current situation in developing large cities. This chapter consists of three sections. The first section reports the results of an interview about the state of praxis control strategies and equipment that has been conducted within transportation departments of large cities. The second section presents the results of saturation flow measurements that are taken, to observe the variation in different traffic conditions as well as to see the saturation flow difference between developed and developing cities. Finally, traffic control related problems in developing large cities and their reasons are discussed in this chapter.

In Chapter 5, the framework for traffic signal control in developing countries is presented. This framework includes short and long-term measures for developing large cities as well as planning and operation levels.

Chapter 6 presents the development of a real-time traffic responsive control method for operation level. The general steps of implementation of pattern recognition method using artificial neural networks are presented. These steps include; data collection, feature selection, clustering, determination of parameters, determination of system architecture, training and the test of the network.

Chapter 7 presents the simulation environment and the results of traffic simulation studies. This chapter describes first, the communication method between microscopic traffic flow simulation and the proposed controller. After that, the small hypothetical arterial network with various scenarios is studied. In this network, several parameters are tested to increase the system performance of neural network controllers. Finally, the proposed controller is also tested using the real network of a district in Istanbul. Two scenarios are also examined for this network to test the performance in several conditions.

Chapter 8 concludes the thesis with a summary and provides a recommendation for the future studies.

2 Traffic Signal Control

2.1 Introduction

Primarily, traffic signals are installed to control conflicts between opposing vehicular or pedestrian traffic movements and thus improve road safety, to regulate the movements and therefore improve the intersection capacity [Slinn et al., 2003]. Traffic signal control is an important operational measure for urban road traffic management, in particular as it has become much more difficult to provide sufficient road space despite the growing traffic demand [FGSV, 2010].

The history of traffic signal control started in Cleveland if the first traffic light with gas-signal lights in London as a temporarily trial is excluded. First, in 1913 in Cleveland an electrical light bulb of red and green was in operation. Then in 1918 both in Detroit and in New York, a manually controlled three-color control system came into operation [Lapierre and Obermaier 1988].

Since its first installation over hundred years ago, the development of signal controllers has been in line with the general technological development. This development has been directed by territorial or regional experience. Consequently, similar but diverse approaches in traffic signal control strategies exist in several regions. Generally, traffic signal control strategies are considered in three main groups: fixed time, traffic actuated and traffic responsive/adaptive control. However, some guidelines examine signal control strategies more in detail.

In this chapter, first, the main definitions, variables and parameters of traffic signal control are described and control strategies are also presented in detail. As an interim conclusion pros and cons are discussed at the end of this chapter.

2.2 Basic definitions of signal control

This section gives some important definitions about signalized intersections. A simple intersection and its signal timing plan with the basic elements are presented in Figure 1. The following paragraphs describe the basic elements of signal timing plan.

- ***Signal group:*** The signal group is a set of one or more lanes that receive always the same signal indication.

- ***Phase:*** The phase is a set of signal groups that do not have conflict and can be served in the same time interval

- ***Intergreen time:*** The intergreen time is the interval between the end of one green time of a signal group and the start of the green time of the next signal group.

- ***Phase sequence:*** The phase sequence is the order of phases that will be applied in a complete cycle. In fixed time control this will be determined according to minimum value of the sum of intergreen times.

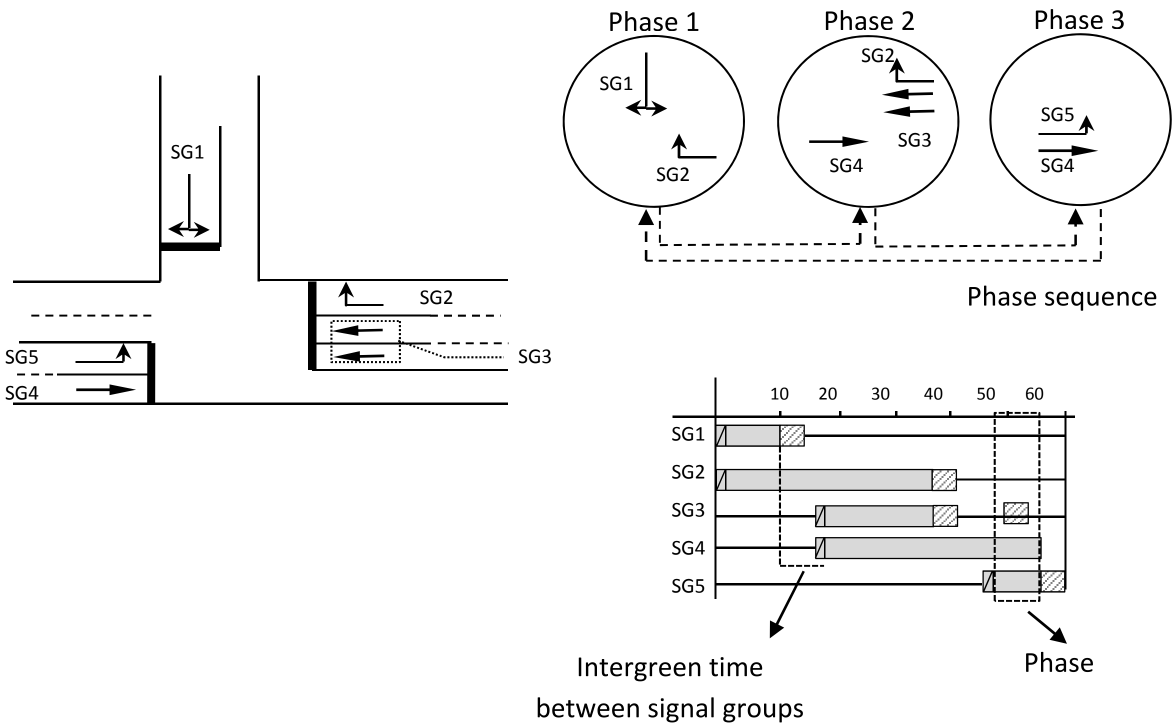


Figure 1: Basic elements of traffic signal control

Besides infrastructural and geometrical elements of a signalized intersection, there are some variables that affect intersection capacity and traffic flow on the network. These parameters are explained in following paragraphs:

- **Cycle Length**

A signal cycle is one complete rotation through all of the indications provided and cycle length is the time that it takes to complete one full cycle of indications [Roess et al. 2004]. Cycle length is one of the significant parameter that affects the quality of traffic flow at signalized intersections. From point of view of traffic engineering the cycle length must be long enough that each approach could be served in one cycle. Thus, the theoretical cycle length is the sum of the lost time and the required green time of each approach [Schnabel and Lohse 1997]. However, too long cycle time is not recommended due to the average waiting time and acceptance of pedestrian. In praxis, cycle length should be between the range of 60s and 90s. A cycle length over 120s is not recommended; otherwise, for example the average delay for pedestrians will be longer [FGSV, 2001b].

As Webster's graph on delay-cycle length relationship in Figure 2 shows, the average vehicle delay at intersections also increases as well with longer cycle time. Webster and Cobbe [1966] recommended an optimum cycle time to minimize average vehicle delay:

$$C = \frac{1.5 * L + 5}{1 - \sum_{i=1}^{p} q_{crit,i}/q_{Si}} \quad (1)$$

where

C = Optimum cycle length (sec)

L = Total lost time

$q_{crit,i}$ =critical volume of phase i

q_{Si} = Saturation flow for critical lane group of phase i

This simple formula for calculating the optimum cycle time is used in other guidelines with small adjustments, for example Akcelik [1982], FGSV [2001b], TRB [2000].

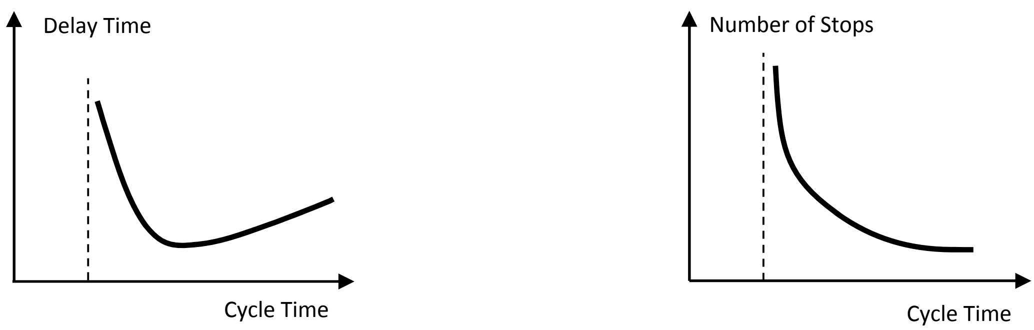


Figure 2: Cycle time and number of stops versus delay time [source: Boltze, 1988]

By coordination of adjacent signalized intersections the same cycle length is required and therefore cycle lengths are calculated for each intersection separately based on saturation degree. After that, the longest cycle length is chosen for arterial or network control [Schnabel and Lohse 1997].

$$C = \frac{L}{1 - \sum_{i=1}^{p} q_{crit,i}/(g_i * q_{Si})} \quad (2)$$

where:

L = Total lost time

$q_{crit,i}$ =critical volume of phase i

g_i = Saturation degree of critical lane group of phase i

q_{Si} = Saturation flow for critical lane group of phase i

- **Green Splits**

Green time is the time interval in that the vehicles are allowed to pass the intersection. A required green time for each approach is calculated as follows [FGSV, 2001b]:

- Actual flow rate divided to saturation flow rate to find flow ratio for each lane group.
- Critical flow rate is determined for each phase.
- Required green time is calculated using following formula:

$$G_i = \frac{\frac{q_{crit,i}}{q_{si}}}{\sum_{i=1}^{p}\left(\frac{q_{crit,i}}{q_{Si}}\right)}(C - L) \quad (3)$$

Where

C = Optimum cycle length (sec)

L = Total lost time

G_i = Green time lane group of phase i

$q_{crit,i}$ =critical volume of phase i

q_{Si} = Saturation flow for critical lane group of phase i

However, according to RILSA [FGSV, 2010] the minimum green time for vehicle traffic flows should be $t_{gr,min}$= 10 sec and generally 15 sec. green time for main approaches is recommended. At intersections with very low traffic volume minimum green time can be reduced to 5 sec [FGSV, 2010].

- **Saturation Flow**

Saturation flow is one of the main variables of capacity and cycle length calculations. The saturation flow is the number of vehicles per hour that can be discharged within one green hour from a signal approach. The base value of saturation flow can vary from one country to another depending on different traffic rule and regulations, geometric standards, and driving behavior.

The base (or ideal) saturation flow rate is 1900(pc/h/ln) on HCM 2000 [TRB, 2000], from 2000 (pc/h/ln) to 3000 (pc/h/ln) in German Traffic Guidelines HBS 2001 [FGSV, 2001b] and from 1800 (pc/h/ln) to 2000 (pc/h/ln) in Austrian Traffic Guidelines RVS [FSV 2009]. Saturation flow can be calculated using adjustment factors, or measured directly. HCM 2000 [TRB, 2000] uses a long formal that includes all factors together, while German Guideline [FGSV, 2001b] prefers only two factors that have the greatest impact.

$$q_{si} = q_{s,st} * f_1 * f_2 \quad (4)$$

The greatest impact on saturation flow has the heavy vehicle ratio. Saturation flow decreases by increasing the ratio of heavy vehicles. Additionally, lane width, turn direction, grade, pedestrian activities and area type also has an effect on the saturation flow rate. Generally, the adjustment factors, according to their negative impacts lay between 0.8 and 1.0 [FGSV, 2001b].

As shown in Figure 3, the saturation flow rate is determined in the middle interval of green period. Generally, in the field measurement, the saturation flow rate is observed after the fourth queuing vehicle passes the intersection stop point.

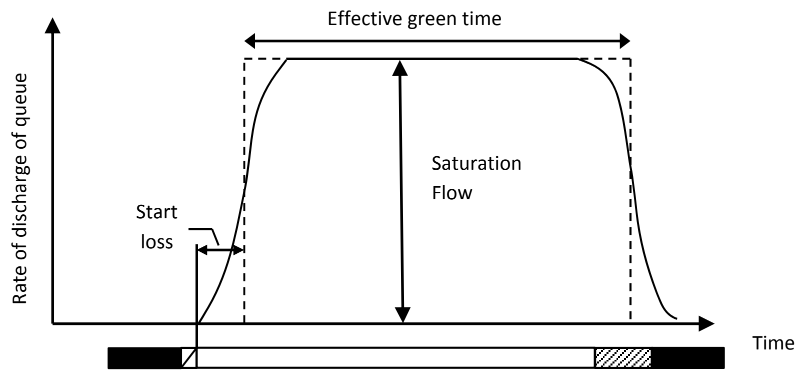


Figure 3: Basic model of saturation flow [source: based on Akcelik, 1982]

- **Offset Time**

Generally, where traffic signals are close enough together so that vehicles arrive at the downstream intersection in platoons, it is necessary to coordinate their green times so that vehicles may move efficiently through the set of signals. In signal coordination, all signal controllers must have the same cycle length, to ensure that the beginning of green occurs at the same time relative to the green at the upstream and downstream intersections. However, if a critical intersection has such a high demand, that it may require a double cycle length. This is rarely done and only when no other solution is feasible [Roess et al. 2004]. Coordination is primarily used to reduce the travel time of vehicles on the road network, thereby reducing fuel consumption and emissions [FGSV, 2010].

The time difference between the two green initiations (i.e. the difference between the time when the upstream intersection turns green and the downstream intersection turns green), is referred to as the signal offset. In Figure 4, the signal offset is defined as (t_2-t_1) [Roess et al. 2004].

Fundamentally, coordination or green wave is applied in two different methods: simultaneous and progressive. The simultaneous system displays the same signal simultaneously at all signals on the coordinated arterial. A simultaneous system is suitable for short intersection spacing (up to 100m). In the progressive system, the green periods are always postponed at succeeding intersections by exactly the arithmetical travel time from stop-line to stop-line [FGSV, 2010]. Progressive systems are also subdivided in simple, forward, flexible and reverse progressive systems, according to the time difference (positive-negative) between green initiations [Roess et al. 2004].

Time-distance planning has to take into account the demands of different road user groups, like private traffic, public transport, pedestrians and cyclist and also emergency vehicles from fire brigades, police and ambulances, too. Temporally and locally differentiated compromises between various road user groups have to be found which do not discriminate against any of the groups [FGSV, 2010].

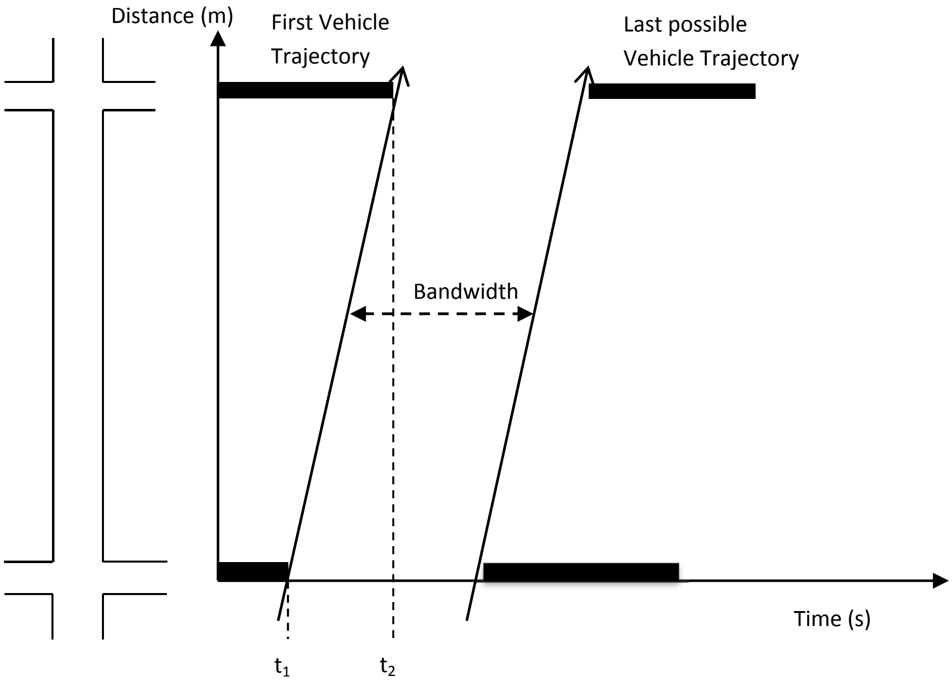


Figure 4: Vehicle Trajectory, Offset Time and Bandwidth on a Time-space Diagram [source: based on Roess et al. 2004]

2.3 Performance of Signalized intersections

In order to assess the performance of the intersections directly or indirectly measurable target values have to be defined [FGSV, 2010]. These assessment parameters are used for comparison of different strategies or junction types, or to assess the system performance with different parameter setting, like cycle length or split. Generally, the following assessment parameters are used in in urban traffic control [FGSV, 2010]:

- Delay

Delay can be defined as the difference between uninterrupted and interrupted travel times through the intersection or intersection groups. This total delay includes the delay due to deceleration and acceleration of vehicle and the delay while the vehicle is idling in a stationary queue [Akcelik, 1981]. Delay should be minimized in order to [FGSV, 2010];

- save time for road users,
- reduce macro-economic loss,
- reduce exhaust emissions,

- Number of Stops

The number of stops can be determined in online or in offline. The number of stops at intersections should be minimized in order to [FGSV, 2010];

- improve driving comfort,
- reduce exhaust and noise emissions,
- reduce fuel consumption etc.

- Speed

Speed can be measured online by detectors if available or with portable equipment on side. The objective is to increase the speed to the local allowed speed [FGSV, 2010].

2.4 Signal Control Methods

Traffic signals can be controlled either time-dependent (mostly known fixed-time control) or traffic-dependent. The traffic-dependent is classified into three groups: traffic actuated, traffic responsive plan selection and traffic adaptive control.

2.4.1 Time-dependent control (Fixed-time control)

The oldest and most widely used traffic signal control method is the fixed time control. Control parameters such as cycle length, green splits and offsets are pre-determined and can only be changed manually or pre-defined by a point of time. Generally in the praxis for each pre-determined or measured traffic situations (morning peak, afternoon peak etc.) timing plans are prepared and transition times between timing plans are pre-defined. Therefore this control method is also called time-of-day method (TOD). A well-designed fixed time signal timing plan can work well enough as long as the pre-measured or pre-determined traffic situation is not changed.

Major advantages of this control method are the simple design, the robustness in operation, and the low construction and operation costs, since this control method does not require any detectors for the intersection approaches. In addition, fixed time controlled intersections on an arterial can be well coordinated without any software support [Priemer, 2011].

However, particularly in urban areas, commonly occurring incidents such as maintenance operations or big events change the daily traffic patterns. Therefore, already in 1928 the first controller was installed in Baltimore, USA, which was operated using real-time traffic data. This system is now called a traffic actuated control [Lapierre and Obermaier, 1988].

2.4.2 Traffic-dependent control

While in fixed time control a relationship between “time” and “timing plan” is established based on historical data, traffic-dependent control methods consist of a closed-loop between “traffic demand”

and "timing plan". Generally, traffic dependent control methods are considered in three groups: Traffic-actuated, traffic responsive plan selection (TRPS) and adaptive traffic control systems (ATCS).

- **Traffic actuated control**

In a traffic actuated control method, the control strategies have a processing logic by directly querying the logical (e.g. green time extension) and temporal (e.g. predefined maximum green time is over, or not) state conditions. Real time evaluation and an optimization are not required in this type of control. Green splits are variable depending on the frame plan and stochastic traffic demands on the approaches (or signal groups) and cycle lengths are chosen according to traffic demand in the signalized intersection. This method can be applied in a single intersection either in semi-actuated or fully actuated mode. In semi-actuated control, detectors generally exist only on minor streets and this approach receives green time and also green extension depending on traffic demand, whereas all approaches (or signal groups) are furnished with detectors in fully actuated mode.

In arterial or network coordination, either the intersections are operated locally as actuated mode and coordination of intersection are provided centrally, or coordinated comes forefront and semi-actuated adjustment on intersections are provided with detectors on respective intersection.

Traffic actuated control provides advantages like directly adaptating to the stochastic microscopic fluctuation and simple realization of temporal demand (i.e. public transport priority). On the other hand, some disadvantages of this system are installation and maintenance costs and its working with high demand as a fixed time [Friedrich, 1999].

- **Traffic responsive plan selection**

Traffic responsive mode was developed to ensure implementing appropriate timing plans in response to actual traffic conditions in the network. Actual traffic demand at strategic locations in the network is measured and the control system software compares the measured traffic demand to established thresholds to determine which timing plans to implement. Subsequently, the traffic responsive system implements a timing plan that is (theoretically) best suited to accommodate the current traffic demand. Generally, the cycle lengths, phase sequences, and green time allocations at each intersection remain fixed until a new timing plan is implemented, according to traffic demand [Balke et al., 1997].

Even if the idea of a plan selection mode could be implemented for isolated intersections, this method in praxis is generally applied to interconnected systems. This control method is located between traffic actuated and traffic adaptive control and can also be confused with both. In traffic actuated control traffic detectors are used to measure microscopic variations (seconds) at intersection approaches, whereas in traffic responsive mode the system detectors are used to measure macroscopic and relative long-term variations (minutes) on a study area.

Traffic responsive systems have similar but different structures in the United States and Central Europe. The system; termed Traffic Responsive Plan Selection (TRPS) in United State, uses a computational channel (CC) to utilize a set of pattern selection (PS) parameters to select a timing plan based on the threshold values set for each of these parameters [Abbas and Sharma, 2006]. While in Germany, named Signal Plan Selection and timing plans are selected using structured query after asking logical, temporal and other conditions [FGSV, 2010]. In the German system, it can also be considered as traffic actuated control method at the macroscopic level. These two systems are depicted in Figure 5.

The main advantage of this system is to adapt current traffic situation and select a best suitable timing plan for respective situation. However, current controllers have a restricted capacity to load timing plans in the plan library. Another aspect in this control method is the transition frequency and stability of the coordination.

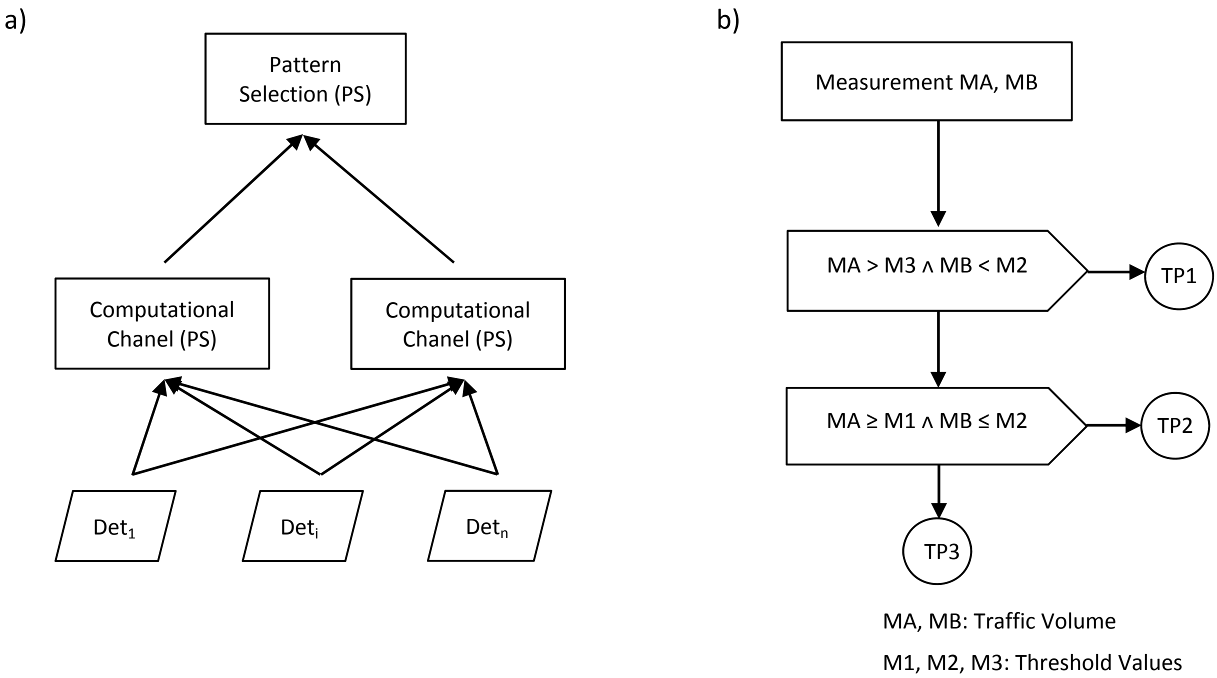


Figure 5: Traffic responsive plan selection mode 5a) in United States [source: Abbas et al., 2009] 5b) in Germany [source: Schnabel and Lohse, 1997]

- **Traffic adaptive control**

The development of traffic adaptive systems goes back to 1980. The objective of these researches was to quick response to the change in traffic demand and so to reduce the cost, like delay or number of stops. Already in 1983, the traffic adaptive systems are classified into three generations. The first generation of urban traffic control systems (UTCS) used pre-stored timing plans, which are previously calculated based on historical data. The second generation of UTCS calculates timing plans in every five minutes based on surveillance data. The third generation systems were conceived to implement and evaluate a fully responsive, online traffic control system. These are similar to the second generation systems. However the third generation systems revise the timing plans in shorter interval (3 to 5 minutes) [Gartner, 1983; Martin et al., 2003].

Compared with traffic actuated and traffic responsive plan selection systems, which responds to actual traffic situation directly using control logic, a traffic adaptive system uses a traffic flow model, impact model and an optimization model to make a decision for switching the timing plans (see Figure 4) [FGSV, 2010]. The traffic situation is replicated in a dynamic traffic model with data that is continuously collected on cross sections. The detailed analysis of the traffic situation with different traffic parameters allows the classification of traffic conditions and timely detection of traffic disruption. [FGSV, 2010]

To avoid confusing adaptive system with others besides the traffic flow and impact model an online adjustment or optimization model can be considered as the main features of adaptive control methods. Definitions about online and offline optimization are described later.

In the praxis, some adaptive control systems use domain-constrained or time-constrained optimization, whereas some others use only rule-based adjustments. In domain-constrained optimization, the optimization search is limited to avoid high fluctuations of signal timings to prevent negative transition effects (i.e. SCOOT), while the time-constrained optimization is constrained by time and/or structural boundaries set by local controller policies (i.e. RHODES, OPAC, BALANCE, and MOTION) [Stevanovic, 2010].

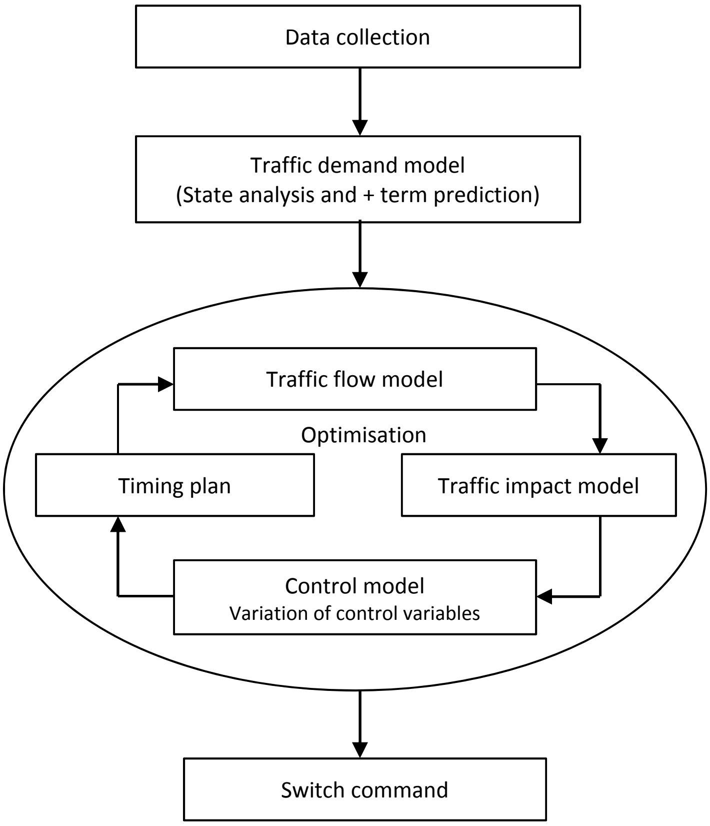


Figure 6: Traffic adaptive control method [source: FGSV, 2010]

The main advantages of traffic adaptive system are not only green time adapted to the short-term fluctuation on traffic demand, but also whole frame plans can be generated again [Priemer, 2011]. Reduction in the number of stops and delays are proved in several comparative studies, for example the research project AMONES in Germany [Boltze et al. 2011]. However, besides its high installation and maintenance costs, addiction of high density sensor mesh and troublesome calibration process are some important points to consider before implementation particularly in those cities, which have relatively less experience in real-time traffic engineering applications. Adaptive control methods vary themselves according to the experience of the countries where they have been developed.

Some adaptive systems support automatically creating of subsystems by adding new intersections (i.e. SCATS, SCOOT, UTOPIA), but some others need reconfiguration by the users (i.e. BALANCE, MOTION, OPAC, RHODES) [Stevanovic, 2010].

The adaptive control systems are summarized in Table 1 (on page 24) and are described in section 2.7.2 in detail.

2.5 Signal control scale

Traffic signals or signalized intersections can be distinguished in two groups according to their spatial range [Al-Mudhaffar, 2008]:

- (isolated) Intersection signal control
- Coordinated traffic signal control

In isolated signal control, the signal timing decisions are based solely on the traffic demand in the approaches to the intersection, while in coordinated control the signal timing plans are based on taking into consideration other adjacent signalized intersections to which the intersection controller is connected in order to facilitate passage of the signalized system [Al-Mudhaffar, 2008]. Coordinated control can also be considered in two groups; arterial (or open network) and network control [Gordon and Tighe, 2005; FGSV, 2010].

Coordination of signals in urban areas is generally beneficial. However, it must be investigated and decided, in each case, whether better performance will be achieved with a coordinated or isolated operation [Gordon, Tighe, 2005].

2.6 Signal control architecture

According to their system and communication architectures the interconnected signal systems can be distinguished in two groups:

- Distributed systems
- Central systems

With distributed control, signal-timing plans that are required for the intersections are directly kept in the intersection signal controller (see Figure 7). Regional traffic computer synchronizes and monitors the intersection controllers. In this control hierarchy, requirement on data communication is relatively low [FGSV, 2001a].

Distributed systems require powerful local-intersection controllers, since the power of the system is inherent in the local controller, these controllers must have all the features desired for signal control at the intersection. However, robustness, easy expandability and general communication costs are some prevailing advantages for distributed systems [Riedel, 2003].

In central control, the signal timing plans are processed only in the regional computer. The regional computer sends second-by-second to each connected intersection controllers signal states of each signal groups. However, some fixed parameters like inter-green time, all-red time etc. are kept in intersection controller directly (see Figure 8) [FGSV, 2001a].

In central control, a central computer makes control decisions and directs the actions of the intersection controllers. Therefore the intersections require only a standard controller and interfacing unit and do not perform any software functions. However, these systems require a reliable communication networks and this increases the cost of the system. Additionally, with this system architecture expending of the system with new intersections is difficult [Riedel, 2003].

However, according to Riedel [2003], the difference between these architectures have blurred in recent years, since most modern centralized systems have many of the important features of distributed systems, and most distributed systems offer many of the most useful central control features [Riedel, 2003].

In addition to these two system architectures, there is another system that is a mixture of these both systems: semi-central architecture. The functions are distributed in intersection control and regional computer according to tasks. The intersection controller monitors intergreen time, transition time and the minimum green time of signal groups, while the regional computer conducts regional signal plan selection and additional control tasks. If there is a problem in communication system each intersection control returns to the time-of-day mode [FGSV, 2001a]. The system architectures of central and distributed systems are depicted in Figure 7 and 8.

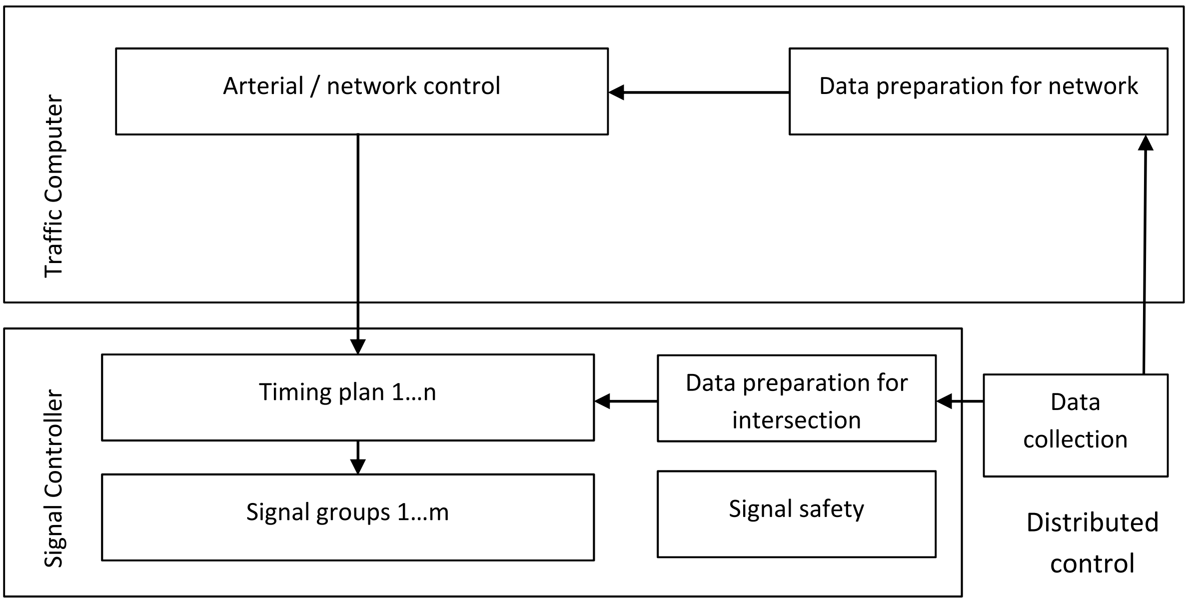


Figure 7: System architecture of distributed control [source: FGSV, 2001a]

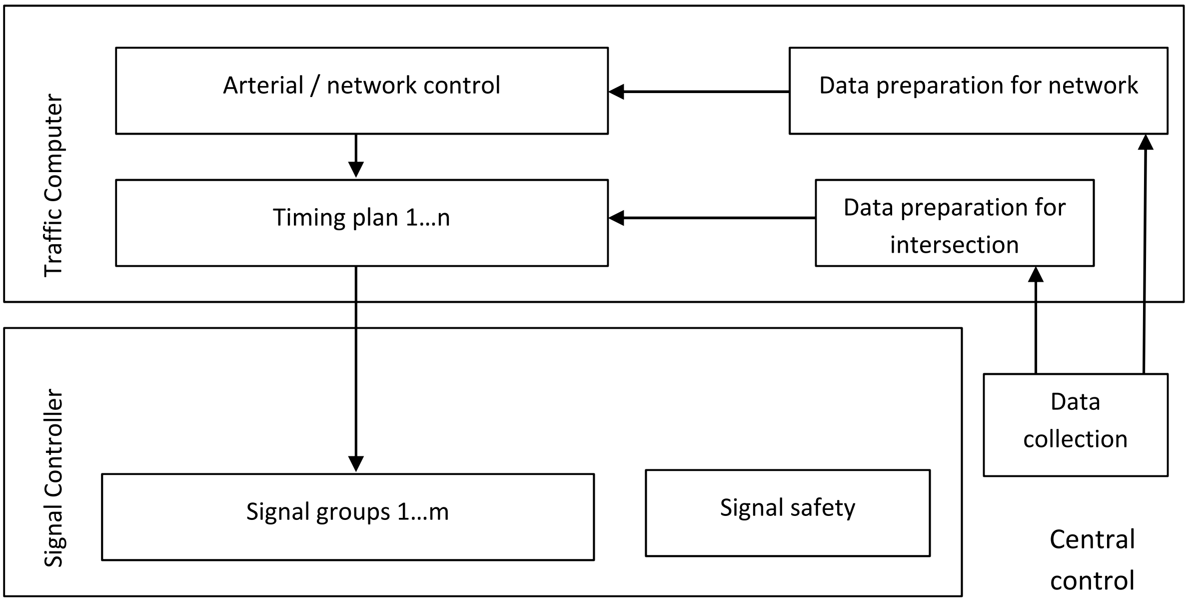


Figure 8: System architecture of central control [source: FGSV, 2001a]

2.7 Multi-level traffic signal control

Most of modern traffic control strategies are developed hierarchically and generally consist of up to three independent levels; strategic, tactical and local level. For example SCATS has two levels structure [Martin et al. 2003] and MOTION [Mück, 2009], BALANCE [Friedrich, 1999], RHODES [Mirchandani and Head 2001] have “three level” control architecture. However, control period differ from one system to another, but in general lower level reacts to short-term changes (in seconds) and upper level long-term changes in traffic flow (several minutes to hours). An example of three levels architecture is depicted in Figure 9.

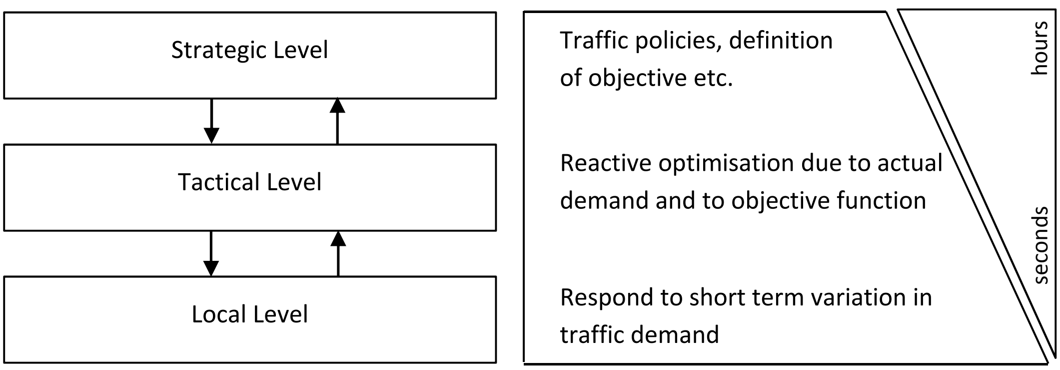


Figure 9: System architecture of a three level control system [source: based on Friedrich, 1999]

Transport policy and the main control criteria are defined for city or region on strategic level. The street categories (regional and trans-regional major roads, arterial urban street etc.) and intermodal facilities are determined in this level. In general, the control and the calculated traffic state is monitored and parameters of the respective objective functions for the different control systems are customized according to the transport policy for the metropolitan area. Thus, the consistency of the objectives is reached for different control systems and facilities (interurban, urban, private and public transport). Integration of current traffic control system is extremely significant for optimizing the capacity of all transport systems and for implementing the common transport policy in the most efficient way [Friedrich and Schüttle, 2000].

Aggregated traffic flow information is used and the origin-destination relations for (sub-) networks are estimated and short to medium term demand forecast is accomplished on a tactical level. Frame plans for signal controller are prepared and sent to the local level [Friedrich and Schüttle, 2000].

On the local level, the signal control responds to short-term variation in traffic demand according to a predefined objective function. Stochastic changes and public transport priority can be accomplished by local actuated control according to predefined frame plans. Microscopic traffic data is collected, aggregated and transmitted to the tactical level [Friedrich and Schüttle, 2000].

As aforementioned, these hierarchical control systems, in general, are introduced as commercial systems and the tasks of each layer and also communication interval between layers may vary from one system to another.

2.8 Optimization in signal control

The term optimization generally refers to searching for the best solution. Optimization consists of three essential elements; an objective function, unknowns or variables and constraints. According to the problem the optimization can mean a minimization of costs (time, money etc.) or a maximization of benefits. In some cases both function can be used for the same problem [Karaboga, 2004].

Variables are essential in optimization problems, since without variables the definition of objective function and of the problem constraints cannot be possible [Vössner, 2010].

The constraints allow the unknowns to take on certain values but exclude others. Constraints are not essential. However, it makes sense in some problems to constrain the search space. For example, in the manufacturing problem, it does not make sense to spend a negative amount of time on any activity, so all the “time” variables should have non-negative values [Vössner, 2010].

Generally, optimization in traffic signal control considers minimizing travel time, number of stops and delays. The main variables of traffic signal control like cycle length, green splits and offset time are optimized to minimize the cost function.

Optimization in signal control can be dated back to the early sixties, when the new developments were appearing in the field of signal coordination. The main idea of this development was to take into account the delay and number of stops. In the macroscopic consideration the traffic density during passing a cross-section is considered as a time function. The vehicle platoons and the traffic density are considered by the downstream intersection controller. The offset time affects both the delay time and the number of stops at an intersection [Zackor et al., 1991].

Several mathematical and heuristic algorithms are used in the optimization of traffic signals in offline as well as online mode. The following section describes a worldwide known offline and online signal optimization methods.

2.8.1 Offline Signal Optimization Models

The following paragraphs summarize worldwide known optimization models. All these models consist of a traffic model, which depicts traffic flow at an intersection or in the network and an optimization model, which seeks the best timing plans for the intersections.

- **SOAP**

SOAP (Signal Optimization Analysis Packages) was developed to optimize signal timing plans for isolated intersections by evaluating a wide range of alternatives, including fixed-time or multiphase actuated control. SOAP determines optimum timing plans for intersections using three computational functions: design, analysis and evaluation [Lin, 1999]. However, SOAP can only analyze and optimize isolated intersections [Martin et al., 2003].

- **MAXBAND**

MAXBAND (Maximal Bandwidth Traffic signal Setting optimization Program) is a macroscopic tool for optimization of coordinated arterials. Progression speed and cycle length are varied and the weighted sum of green bandwidth is maximized for both directions. Several predefined phase sequences can be considered, particularly, lead-lag operation for left turns from main road [Fellendorf, 1991]. The objective of optimization in MAXBAND is to enlarge the total green band width. The main advantage of this tool is that the opportunity to produce several cycle lengths and progression speeds [Wietholt 2009].

- **PASSER II**

PASSER II (Progression Analysis and Signal System Evaluation Routine) is also a macroscopic system to maximize the bandwidth of green wave signal coordination. Additionally to MAXBAND, it provides optimization of green splits. Passer II has a simple traffic model, which represents vehicle queues on the arterial [Fellendorf, 1991]. In the new version of PASSER II, the Webster's delay time calculation for the estimation of platoon dispersion is built upon. The main advantages of this tool are that the flexibility of the variable cycle time and green band, as well as consideration of more phases for different control strategies [Wietholt, 2009].

- **SIGOP**

SIGOP (Traffic Signal optimization Program) was developed in 1976 and was based on successive approximation method for optimizing the signal timing plans. The traffic model considers traffic flow platoons, as well as platoons of turning vehicles. SIGOP uses the sum of delays and the weighted number of stops in the performance index for optimization. For an optimization of n-intersections, the problem is considered as n-dimensional problem and divided into n-sub problems and so the signal control variables are optimized. The main advantage of SIGOP is that the green time and offset are simultaneously optimized and hence a real optimum can be found. However, it is very questionable whether the selected solution of sub-problems actually represents the optimum for the whole system [Zackor et al. 1991].

- **SYNCHRO**

SYNCHRO is currently one of the widely used optimization software. Synchro optimizes the timing plan by minimizing the performance index (PI) that consists of delay, number of stops and the number of vehicles affected by the queue. SYNCHRO optimizes the traffic signals in two levels; the intersection level and the network level. The intersection level optimizes the green splits and cycle lengths. The network level, if required, can divide the system into sub-networks and optimizes the offsets, phase sequences and common cycle for selected zones [Martin et al., 2003]. The main advantages of SYNCHRO are the user friendly interface and the similar input and output windows to Highway Capacity Manual (HCM) to whom who may familiar with the definitions of HCM.

- **TRANSYT**

TRANSYT (Traffic Network Study Tool) is the most worldwide known network optimization model, which was developed by Robertson in 1969. TRANSYT is an offline model for determining and studying optimum fixed time, coordinated, traffic signal timings in any network of roads for which the average traffic flow is known. The traffic model calculates a performance index in monetary terms, which is a weighted sum of all vehicle delay and stops. An optimizing routine systematically alters the signal offsets and/or allocation of the green times to search for the timings which reduce the performance index to a minimum value [TRL, 2011]. After 1975 in the United Kingdom developed TRANSYT model spread and branch off first in United States as TRANSYT-6B and TRANSYT 7F and in Germany as TRANSYT-6D and TRANSYT-ISBAC [Boltze, 1988]. The model used models are TRANSYT 7F in USA and TRANSYT 13 in UK.

- **VISUM**

The well-known transport planning software VISUM also introduced optimization tool in its new version. Within the scope of the intersection capacity analysis the green times, cycles and offsets can be optimized. The offset optimization is based on the traffic assignment results. This enables optimization of large networks automatically without any manually preset direction of coordination. In VISUM, optimization of networks is done according to delay minimization. The intersection capacity analysis (ICA) assignment method takes into account of the interaction between optimization and traffic assignments [PTV, 2012]. The offset optimization model of VISUM is based on [Möhring et al., 2006]. This model uses mixed-integer linear program (MIP) to minimize the total occurring delay of vehicles in the network. The road network is considered as a directed graph and convex link performance functions (LPFs) are introduced for each arc. The lost time due to poor synchronization of the signals is measured according to arrival time of vehicle platoons [Möhring et al., 2006].

The main advantages of this tool are that both the demand/supply data and optimization issue can be done in one network and particularly in large network with less data can also be optimized [PTV, 2012].

2.8.2 Online Signal Optimization Models

Online signal optimization models have to react directly to the current traffic situation by optimizing the timing plans in real-time. In the following paragraphs, several worldwide-implemented online models are reviewed. These models have several characteristics in common. Except SCATS, all other methods have a traffic model and also have an online optimization module. These models are also known as “traffic adaptive systems”.

Despite some confusion in traffic adaptive systems and traffic responsive systems, it can be assumed that online optimization is the common feature of traffic adaptive systems. For example, SCATS is assumed in this category despite its lack of traffic model or SCOOT calculates or determines cycle and green times incrementally. Therefore some newer publications try to categories these systems again. However, in this study, all these systems are considered as traffic adaptive systems. The following paragraphs describe some well-known traffic adaptive systems.

- **SCOOT**

SCOOT (Split, cycle and Offset Optimization Technique) is one of the first worldwide used industrial optimization techniques of adaptive systems. SCOOT adjusts signal timing parameters in frequent, small increments to match the latest traffic situation. The main idea of SCOOT based on cyclic flow profiles (see Figure 8). This is similar to offline method TRANSYT and will also be fundamental to some other adaptive systems, which were developed later. The vehicle counts and occupancy data are stored in the SCOOT computer in the form of ‘cyclic flow profiles’ for each approach of the signalized intersections. These ‘cyclic’ profiles consist of histograms, which record how the traffic flow varied during one cycle time of the upstream signal [Hunt et al., 1981].

The detectors in SCOOT are placed downstream of each intersection. As shown in Figure 10, SCOOT estimates queues. The current traffic state is recorded and vehicle arrivals are known from cyclic flow profile. Thus, vehicle arrivals during the red time are added onto the back of a queue, which usually builds up until the next green time [Hunt et. al., 1981]. In general the idea of SCOOT is based on TRANSYT, however in this case the time interval of calculation is shorter and the system works as a closed loop.

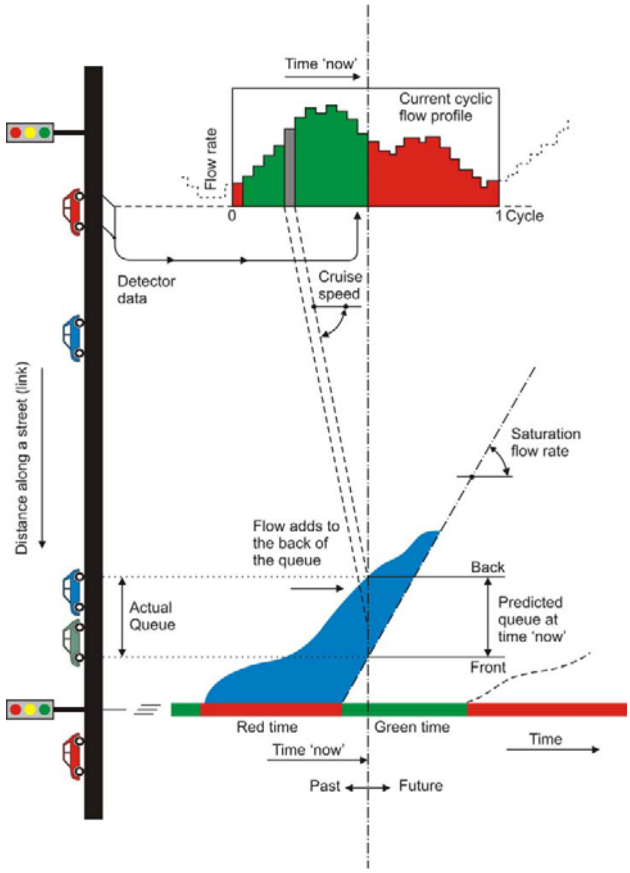


Figure 10: Traffic flow model in SCOOT [Hunt et al. 1981]

- **SCATS**

SCATS (Sidney Coordinated Adaptive traffic System) was developed and first installed in Sydney in Australia. Now, SCATS is used in more than 50 cities worldwide. Similar to SCOOT, SCATS adjust cycle time, splits and offset in real time to respond to the current demand. SCATS uses downstream detectors to calculate the cycle and split times. Unlike SCOOT and other adaptive methods, SCATS does not have any traffic model [Martin et al., 2003].

SCATS optimizes splits and cycles according to an objective function. Although the objective function is determined according to delay and number of stops, the degree of saturation of intersection approaches is used as a main parameter in the optimization. It controls the traffic in two levels: strategic and tactical levels, which are also used in newer adaptive systems. In the tactical level the regional computers of SCATS control up to 10 intersections. These regional computers are connected to the main system [Martin et al., 2003, Zackor et al, 1991].

- **RHODES**

RHODES (Real-time Hierarchical Optimized Distributed and Effective System) decompose optimization problem into level and uses three levels architecture to solve it. The heights level has the “dynamic network loading model”. The slow-varying characteristic of traffic is captured in this level. The middle level is the “network flow control model” that allocates the green time based on the different demand patterns. The lowest level has the “intersection control model”. The required phase change patterns are determined in this level. Each level has two components: estimation/prediction and control. The RHODES consists of several algorithms for prediction, control and optimization. The PREDICT predicts traffic flow at intersection level, while the APRES-NET predicts platoon arrivals at network level. The REALBAND algorithm optimizes the timing plans. RHODES enables also public transport priority and has a link to microscopic traffic flow simulation tool CORSIM [Mirchandani and Head, 2001; Martin et al., 2003].

- **MOTION**

MOTION (Method for the Optimization of Traffic signals in online controlled Networks) is an adaptive control, which also includes multi-layer traffic control. Optimal cycle length, state sequence and coordination are defined in tactical level in every 5 to 15 minutes. The green times of each signal groups are readjusted with consideration of private traffic and public transport at the local level in every second. MOTION can also estimate a local origin-destination matrix with the turn-rates. The optimization occurs in three steps. In the first the split-optimization occurs. The cycle length is optimized in the second step, and the optimization of the coordination takes place in the last step [Mück, 2008].

The general process in MOTION consists of four models. The traffic state estimation module estimates the traffic state based on measured traffic counts and occupancy data. It includes the estimation of queue lengths, turning rates and traffic volume. The parameter selection module activates predefined parameters for corresponding traffic condition. If no control strategy is activated, the default parameters are selected. Cycle and split optimization module works by optimizing the timing plans, which is based on estimated traffic situation and the predefined parameter set. The estimated queue lengths are also considered as the additional number of vehicles in the process to timing plan calculation. The offset optimization module realizes green wave in the control area based on the parameter set, the estimated turning rates and the cycle time and splits of each intersection [Ge and Poschinger, 2011].

The MOTION also offers also congestion and incident management for the control center. The parameters (i.e. discharge time, weights for delays and stops) are determined differently for the splits and offset time determination for several situations [Siemens, 2002]. Public transport priority and the consideration of adjustments for local fluctuation in traffic demand at intersections are some advantages of MOTION. The system is used in over 15 cities in Germany and Europe.

- **BALANCE**

Another German adaptive control system BALANCE (Balancing Adaptive Network Control Method) was developed firstly in 1994 and further in 1998 during several European research projects. Finally, Friedrich [1999] composed the whole algorithm in his dissertation in 1999 [BMVBS, 2011].

BALANCE has also a three level architecture. On the strategic level, the transportation politics and the target function of the system is defined. Medium term traffic forecast and frame signal plan optimization is realized in tactical level (MacroBALANCE). The local level (MicroBALANCE) considers to the short-term stochastic traffic flow variations. Unlike SCOOT the queue lengths or delay times are based not directly on the measurement (cyclic flow profiles), but on a model approach. This enables the optimization of coordination despite of failure in local components [Firedrich, 2000].

BALANCE analyzes and estimates the traffic state both in the tactical and local level. The frame of estimation (origin-destination matrix estimation) is done at a tactical level and after that the local variations in traffic state (i.e. effect of change in green time to the traffic state) are analyzed at a local level. In consideration of computer power BALANCE offers a simplified method for local traffic state estimation. BALANCE uses Hill-Climbing algorithm for the optimization of traffic signals. The system is implemented and tested in several German and other European cities [Friedrich, 1999].

- **UTOPIA/SPOT**

UTOPIA (Urban Traffic Optimisation by Integrated Automation) was developed as a hierarchical and distributed architecture, which consist of a higher level (center) and a lower level. The higher level is responsible for setting the overall control strategies like long-term forecasting and network optimization, while the traffic signal control is implemented by means of the SPOT (System for Priority and Optimisation of Traffic) software at a lower level. The weighted and selective priority of public transport vehicles at signalized intersections is made able with SPOT at the lower level. The system has a modular architecture, which is significant to extend the system and add additional intersections. Each local control unit is further subdivided into two main parts; the observer and the controller. The observer updates the estimations of traffic states based on collected data. The observer also performs also the estimation of slow time-varying intersection parameters such as travel times, turning percentages and saturation flows. The controller part determines the suitable timing plans for actual situation at the intersection. It performs an optimization on the 'time horizon' consisting of the next 120 sec., and is repeated every 6 sec. [Mauro and Di Taranto, 1984; Di Taranto 2007].

SPOT uses delay and number of stops as cost and in order to give priority for public transport vehicles higher unit costs are used for those vehicles [Al-Muhaffar, 2006]. The weights of costs are updated at the network level. The local level has also an important feature for local oversaturation problems. If the observer detects local unpredictable change in demand or in network parameters on the single link, the controller does two actions successively. The first action consists of relaxing constraints on green times both in the actual intersection and the receivers. Additionally, the queue weights of the critical links are increased. If the first action does not succeed, the second action is performed. During the critical period a new cost element is evaluated. In order to obtain a smoothed traffic for each demand situation in the whole control area, the area control chooses the suitable functional and parameters for the local level [Mauro and Di Taranto, 1984].

The system was further developed through the addition of the public transport fleet management system FLASH. The FLASH generates the forecasts that used by UTOPIA for public transport priority. UTOPIA/SPOT is implemented more than 30 European cities, controlling large areas (like Turin, Rome and Milan) and smaller networks (like Oslo, Bergamo, Cremona and Eindhoven) [Di Taranto 2007].

- **OPAC**

OPAC (Optimized Policies for Adaptive Control) is a real-time signal timing optimization algorithm. OPAC was originally developed at the University of Massachusetts, Lowell. This model consists of a distributed control strategy with a dynamic optimization algorithm that calculates signal timings to minimize a performance function of total intersection delay and stops. OPAC has three layer architecture system. The local layer implements rolling horizon procedure and continuously calculates optimal switching sequences for the projection horizon. The coordination layer optimizes the offset for each intersection one per cycle and the synchronization layer calculates the network-wide virtual-fixed-cycle for every few minutes [Gartner et al. 2002].

OPAC consists of two models: uncongested model and congested model. In the uncongested model, the timing plans can be determined either due to obtaining the fixed-time plans in offline or calculating a virtual cycle dynamically. In the congested model, the system considers the saturation flow rate and maximizes the number of vehicles that can pass through the intersection. This model also considers the critical links of networks. Except the cycle time calculation, OPAC is not controlled by a central computer. Thus, OPAC can still work despite a communication failure with the control center [Martin et al., 2003].

The basic principles, detector location, action frequency and other significant features of above mentioned traffic adaptive system can be found in Table 1.

Table 1: Adaptive Systems and basic principles [source: based on Stevanovic, 2010]

Adaptive System	BALANCE	MOTION	OPAC	RHODES	SCATS	SCOOT	UTOPIA
Detection	NSL	NSL	MB&SL	MB&SL	SL,NSL, MB	US&SL	US&SL
Action	P &R	P&R	P	P	R	P&R	P
Adjustment	TCO	TCO	TCO	TCO	RA	DCO	TCO
Timeframe	5 min	5-15 min	Phase/Cycle/5 min	Second by second	Cycle	Cycle/5min	3 sec-cycle
Level	C/L	C/L	C/L	C/L	C/L	C/L	C/L
Model	Yes	Yes	Yes	Yes	No	Yes	Yes
Timings	S,CI,O,PS	S,CI,O,PS	S,CI,O	S	S,CI,O	S,CI,O,PS	S,PS
Flexi Region	No	No	No	No	Yes	Yes	Yes
Vehicle Actuated	Yes	Yes	No	No	Yes	Yes	Yes
Transit Priority	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Detection: SL- Stop Line NSL-Near Stop Line MB- Mid Block US- Upstream
Action: P- Proactive R-Reactive
Adjustment: RA-Rule-based Adjustment DCO-Domain-constrained Optimization TCO-Time-constrained Optimization
Level: L-Local C-Central
Timings: S-Split CL-Cycle Length O-Offset PS-Phase Sequencing

2.9 Artificial Intelligence in Traffic Control - a literature review

Artificial intelligence (AI) is increasingly used in many controlling, classification and optimization problems as well as traffic signal control. Artificial intelligence can be defined as the study and construction of agent programs that perform well in a given environment, for a given agent architecture [Russel and Norvig 2003]. Generally, genetic algorithms are used for optimization, fuzzy logic for controlling and artificial neural networks are used for estimation problems. In this section the studies are summarized which have used one of these three artificial intelligence algorithms for traffic control.

Abbas and Sharma [2006] used a multi-objective non-dominated sorting genetic algorithm to select timing plans. Traffic states were clustered and each cluster was associated to one predefined timing plan. They also developed a new performance measure and degree of detachment to address the clustering requirements. Stevanovic et al. [2007] also used genetic algorithm and developed an optimization tool for signalized networks, named VISGAOST. The VISGAOST optimizes cycle, offset and splits as well as phase sequences, and uses VISSIM as the evaluation environment. Braun [2008] used genetic algorithm based GALOP for the traffic signal optimization. The Hill climbing algorithm of BALANCE has been replaced with GALOP optimization algorithm. The larger and more complex network, the better performance or advantages indicates the GALOP [Braun, 2008].

Favilla et al. [1993] presented a fuzzy logic traffic controller with an adaptive module for isolated intersections. Adaptive module is used to adjust the membership functions according to the traffic condition to optimize the performance of the controller. They used the statistical and fuzzy adaptive method and compared them to each other. Chiu and Chand [1994] presented a distributed method for traffic signal control using fuzzy logic. Using this approach, signal-timing parameters at a given intersection are adjusted as functions of the local traffic condition and of the signal timing parameters of adjacent intersections. Hence, they adjusted cycle, split and also offset for the coordination. Kaur and Konga [1994] developed a fuzzy logic controller for isolated intersections. They used traffic demand during red, traffic demand during green and cycle length as input, and change probability as output. Kaur and Konga [1994] compared the system against conventional pre-timed controller and human experts.

Nittymaki et al. [2001] developed a fuzzy control approach named ITCARI for arterial intersections. The system is evaluated using microscopic simulation software HUTSIM in comparison with vehicle actuated and other two approaches. They also installed the system at an intersection in Tampere for testing. Murat [2001] developed a fuzzy logic control for isolated intersections. The model consists of two sub-systems; fuzzy logic signal time controller and fuzzy logic phase sequencer. Maximum queue length during red interval, number of arrivals in green and remaining green time are selected as input of the system and decision of green time changing is considered as output of the fuzzy controller. Murat also developed own simulation model to evaluate the proposed fuzzy signal control model. Yulianto [2003] designed and evaluated a traffic signal controller based on fuzzy logic for an isolated intersection with four approaches. The proposed method uses maximum queue length and average occupancy rates which are collected in previous cycle in order to estimate the required green time for each approach. Yulianto [2003] focused on mixed traffic conditions in developing countries and modeled this situation in VISSIM for evaluation of the system. Liu and Hsu [2006] designed a fuzzy controller for isolated intersection and implemented the system in VISSIM using vehicle-actuated programming (VAP) module and compared with fixed time and vehicle actuated controller. They also compared fuzzy model both with and without an adaptation mode.

Khan [1996] used modular neural network for detection of operational problems on urban arterials. The proposed model aimed to detect reduced capacity, excess demand, and detector malfunctions. The algorithm was also tested using the data in Los Angeles and Anaheim in California. The main advantage of this algorithm is its modular architecture. The modular architecture enables to detect different types of problems under different operating conditions. Chang [1999] developed a practical approach of combining both advanced neural networks and conventional error correction techniques to estimate traffic control patterns based on real-world traffic measures. The work by Chin et al. [1999] uses neural network and simultaneous perturbation stochastic approximation (SPSA) technique for a system-wide traffic responsive control. The approach was also analyzed by a realistic simulation of a nine-intersection network within the central district of New York City. The study by Wen et al. [2001] uses probabilistic neural network (PNN) to solve incident detection problem on freeways using a pattern recognition method. Wen *et al*. [2001] divided the freeway into sections and used speed, occupancy and traffic counts for each section as input parameters for incident detection.

Sharma et al. [2005] proposed a methodology for optimal configuration of traffic responsive plan selection system. They used artificial neural networks for this methodology and compared them against the state of art Bayesian-based approach for traffic responsive plan selection system. They tested the methodology using the field data in Texas. The classification accuracy reached to 94% (it was 92% in Bayesian-model approach) and after some change in the current architecture of the signal control system it increased to 98%. Murat [2006] used two artificial intelligence approaches to model the vehicle delays at intersections. Murat developed neuro-fuzzy delay estimation (NFDE) model and artificial neural networks delay estimation (ANNDE) model and compared these models with analytical models such as Highway Capacity Manual (HCM) model and Akcelik's model. Both models showed better performance than the analytical models especially for high traffic volumes and over-saturated traffic conditions.

Abdelaziz [2008] used neural network to define optimum threshold value for TRPS Control and analyzed the system for a large arterial network using VISSIM. Von der Ruhren [2006] used pattern recognition and artificial neural networks for short-term traffic state estimation in urban areas. His modular approach clusters the traffic states firstly; in local level at detection points and then in upper level in the control center.

Asar et al. [2008] presented an artificial neural network based traffic signal controller. They used the Particle Swarm Optimization technique both in the learning of the network as well as generating training cases for neural network. They could eliminate clustering problem in a traffic responsive plan selection system. However, the system is tested at an isolated intersection. Shen [2008] used dynamic neural networks for short-term travel time prediction based on traffic data collected by radar traffic detectors. Shen also developed a travel time estimation method and compared it with existing methods. Shen additionally compared the performance of the system with different neural network topologies with different memory setups.

Li et al. [2010] used recurrent neural network to network-wide traffic control. The system is based on a traffic flow model, which comprises two recurrent neural networks. Two back-propagation networks are related to the most essential parameters of traffic signal control: green times, splits and offsets. Balaji et al. [2010] presented a distributed multi-agent-based traffic signal control for optimizing green timing in an urban arterial road network to reduce the total travel time and delay experienced by vehicles. They used the reinforcement learning method to adjust the parameters like weights, thresholds etc. They also evaluated the system by simulating of 29 intersections in Singapore using PARAMICS. Lu et al. [2011] developed and evaluated a multi-agent adaptive traffic signal control system based on swarm intelligence and the neural-fuzzy actor-critic reinforcement learning (NFACRL) method. They used two scenarios to evaluate the method and compared the new NFACRL-Swarm method with its NFACRL counterpart both an isolated intersection as well as in coordinated arterial.

2.10 Conclusion

Traffic signal control has a long history of more than a hundred years. The countries particularly in the developed world, pay attention to the signal control algorithms and developed their own experiences and regional factors. Some countries preferred only car traffic in the networks, whereas some others privileged public transport vehicles and researched for more microscopic solutions. Several studies proved that adaptive control methods reduce travel time and delays. However, according to an expert interview on North American cities show that only 57% of adaptive control users would install the same system again. They mentioned several reasons (i.e. technical support, cost-benefit ratio, maintenance costs etc.) not to install the same system again [Stevanovic, 2009]. Additionally, these systems require high-density sensor meshes and show better performance only if the parameters are well-calibrated and good maintenance is done. Also, Brilon et al. [2007] report problems of adaptive system with detector malfunctions and criticize the requirements on detectors and costs of these systems.

Under these circumstances, the developing large cities have to ask themselves, whether these long theoretical and practical experiences of developed countries in signal control is directly applicable in their cities.

The main objective of this study is to research the current situation in developing large cities and to develop a simplified control method that optimize timing plans in offline, but react the current situation in real-time. The next chapter reports the current situation in signal control of developing large cities and their traffic related problems.

3 Pattern Recognition and Artificial Neural Networks - Theoretical Background

3.1 Pattern

The origin of the term "pattern" relies on the old French term "patron" and in common usage has several meanings. For example it can refer to [Oxford, 2012]:

- a repeated decorative design,
- a model or design used as a guide in needlework and other crafts
- an excellent example for others to follow

In machine learning, an attractive description of the term "pattern" coined by Watanabe [1985]: "a pattern is the opposite of chaos; it is an entity, vaguely defined, that could be given a name". Besides fingerprint, shape or handwritten characters, pattern can also be measurements of an acoustic waveform in a speech recognition problem; measurements on a patient for diagnosis; measurements on weather variables (for forecasting or prediction). The examples indicated that the term 'pattern', in its technical meaning does not necessarily refer to structure within images [Webb, 2002]. Niemann [2003] defined the pattern in machine learning as the following: "patterns are measured physical quantities or functions of a certain problem". These quantities or functions can be measured with one or more sensors. This can be formulated as in (5):

$$f(x) = \begin{bmatrix} f_1(x_1, \ldots\ldots, x_n) \\ f_2(x_1, \ldots\ldots, x_n) \\ \vdots \\ f_m(x_1, \ldots\ldots, x_n) \end{bmatrix} \tag{5}$$

where,

$f(x)$ = a pattern

x = measured variable

n = number of variable

m = number of sensors or components

Figure 11 represents a simple example of a pattern of traffic flow as a function over time at one sensor. Since here only traffic count at only one sensor represented, the m=1 and n=1. Using more features the pattern can be classified more preciously and more sensor points can help to represent the situation in a region.

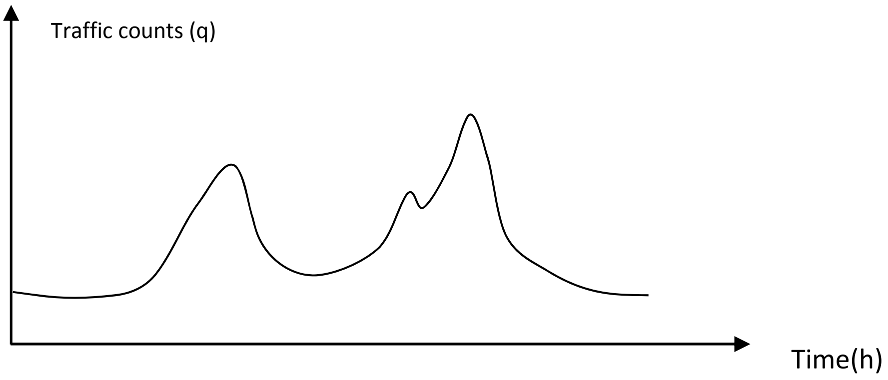


Figure 11: Time series of traffic count at one count station is an example of pattern

Also, Niemann (2003) considers the pattern in two groups: simple and complex pattern. A (simple) pattern has features that are characteristic for its membership to a class. These features build a reasonably compact area in feature space for the pattern of a class. The areas that belong to different classes are reasonably separated each other. Also, a (complex) pattern has simple components that have a certain relationship with each other and the pattern can be divided into these components. Of course, in a (complex) pattern, any arrangement of the components cannot build a pattern. A structure is required, which provides also more patterns using relative less number of components [Niemann, 2003]. Figure 12 represents that not each arrangement of components create a reasonable patterns [Niemann, 2003].

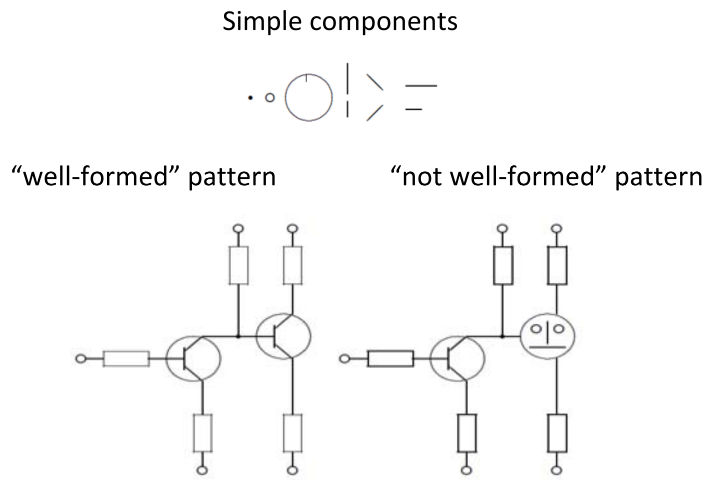


Figure 12: Well-formed or not well-formed pattern can be created from same components [Niemann, 2003]

Pattern recognition method is a useful method to analyze these data structures and to classify the data to the pattern automatically.

3.2 Pattern Recognition

Pattern recognition is a scientific discipline in machine learning. The goal of pattern recognition is the classification of objects into a number of categories or classes. These objects referred to as pattern and can be images or signal waveforms or any type of measurements that needs to be classified. [Theodoridis and Koutroumbas, 2003]. The history of pattern recognition goes back to the 1960s. However, before 1960 it was mostly the output of theoretical research in the area of statistics [Theodoridis and Koutroumbas, 2003]. Pattern recognition is an interdisciplinary subject, covering developments in the areas of statistics, engineering, artificial intelligence, computer science, psychology and physiology, among others [Webb 2002]. Pattern recognition algorithm comprehends a wide range of information processing problems such as speech recognition, classification of handwritten characters, fault detection and medical diagnosis [Bishop, 1995]. In this chapter several approaches to the pattern recognition will be briefly introduced such as, template matching, statistical method, syntactical method and artificial neural network approach. Generally, according to learning algorithm, pattern recognition models are considered in two groups [Webb 2002]:

- supervised learning
- unsupervised learning

In supervised learning, a set of training data with associated labels has been provided, while in unsupervised learning, a set of data is not labeled and the user seeks to find groups in the data and the features that distinguish one group from another [Webb 2002].

Pattern Recognition system is applied in two modes and each mode involves three modules. The provided data is separated into two groups and with the first part of data the system is trained first, subsequently; the system is validated using other part of data. In the *pre-processing* of training mode, the pattern of interest is segmented from the background, noise is removed; the pattern is normalized and other operations that will contribute in defining a compact representation of the pattern. The role of the *feature extraction*/*selection* module in training mode is to find the appropriate features for representing the input patterns and to train the classifier to partition the feature space. This module concerns about the best number of features to use. Parameters in the pre-processing and feature extraction strategies are optimized using feedback path. In the validation or classification mode, the input pattern is assigned to one of the pattern classes under consideration based on the measured features [Jain et al. 2000]. The detailed pattern recognition process for statistical approach described by Jain et al. [2000] is illustrated in Figure 13.

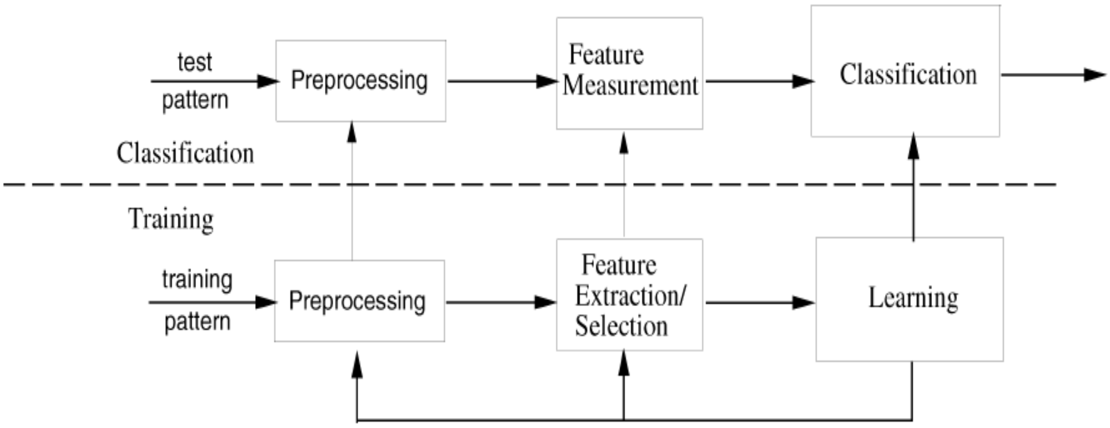


Figure 13: Model for statistical pattern recognition [source: Jain et al., 2000]

In the literature there are four widely used pattern recognitions techniques; template matching, statistical, syntactic and neural network approaches. The main features of these approaches are listed in Table 2 and described briefly in the following paragraphs.

Table 2: Pattern Recognition Models [source: Jain et al. 2000]

Approach	**Representation**	**Recognition Function**	**Typical Criterion**
Template Matching	Samples, pixels, features	Correlation, distance measure	Classification error
Statistical	Features	Discriminant function	Classification error
Syntactic or structural	Primitives	Rules, grammar	Acceptance error
Neural Networks	Samples, pixels, features	Network function	Mean square error

- **Template Matching**

Template matching is one of the simplest and earliest approaches to pattern recognition. The main idea is to determine the similarities between two entities (points, curves, or shapes) of the same type. All allowable pose and scale changes of the pattern are taken into account by the matching it against the stored template. The measure of the similarity may be optimized based on the available training set. Generally, the template is trained from the training set. This approach is actually computationally demanding, however the availability of today's faster processors has now made this approach more feasible [Jain et al., 2000]. However, because of the large variations often encountered in the examples, this approach is not the most effective method to pattern recognition [Jain and Duin, 2004].

- **Statistical Pattern Recognition**

In the statistical method, each pattern is represented in terms of d features or measurements and is viewed as a point in an n-dimensional space. The goal is to choose those features that allow pattern vectors belonging to different categories to occupy compact and disjoint regions in an n-dimensional feature space. The effectiveness of the representation space (feature set) is determined by how well patterns from different classes can be separated [Jain et al., 2000]. Statistical pattern recognition approach may estimate the statistical distribution of the features from the total samples. At any coordinate or point in the representation space, one could estimate the likelihoods of it belonging to one group or other; depending upon which likelihood is higher, one could determine the class of an entity [Jain and Duin, 2004].

- **Syntactic pattern recognition**

In syntactic pattern recognition, a set of rules is developed to determine the structure of a pattern class in terms of primitive features of the patterns. This set of rules is converted into a grammar [Hamilton et al., 1992] Thus; a large collection of complex patterns can be described by a small number of primitives and grammatical rules. The grammar for each pattern class must be inferred from the available training samples [Jain et al., 2000]. The main advantage of this method is to recognize highly structured and complex patterns. However, when they have to accommodate large amounts of noise, syntactic systems can become very complex [Hamilton et al., 1992].

- **Artificial Neural Networks**

Artificial neural networks, as suggested by its name, attempt to apply the models of biological neural systems to solve practical pattern recognition problems. Neural network approach has become popular for solving pattern recognition problems [Jain and Duin, 2004]. The main characteristics of this method are their ability to learn complex nonlinear input-output relationships and their sequential training procedures, and their ability to adapt themselves to the data [Basu et al., 2010]. Artificial neural network approach is explained in the following section in detail.

3.3 Neural Network Approach

3.3.1 Biological Neuron and Artificial Neuron

A biological neuron contains a cell body (*soma), axon* and the *dendrites* (see Figure 14). Signals or impulses are received by neuron from other neurons through dendrites (receivers) and signals generated by soma are transmitted along the axon (transmitter). An axon contains strands or sub-strands. At the end of an axon are located the synapses, which have an elementary structure and are functional unit between two neurons. When the signal inputs reach to the synapses, the neurotransmitters are released and diffuse across the synaptic gap and reach to the receiver of other neuron. The synapses can learn from the activities in which they participate. Hence, the effectiveness of the synapses can be adjusted [Jain, et al., 1996].

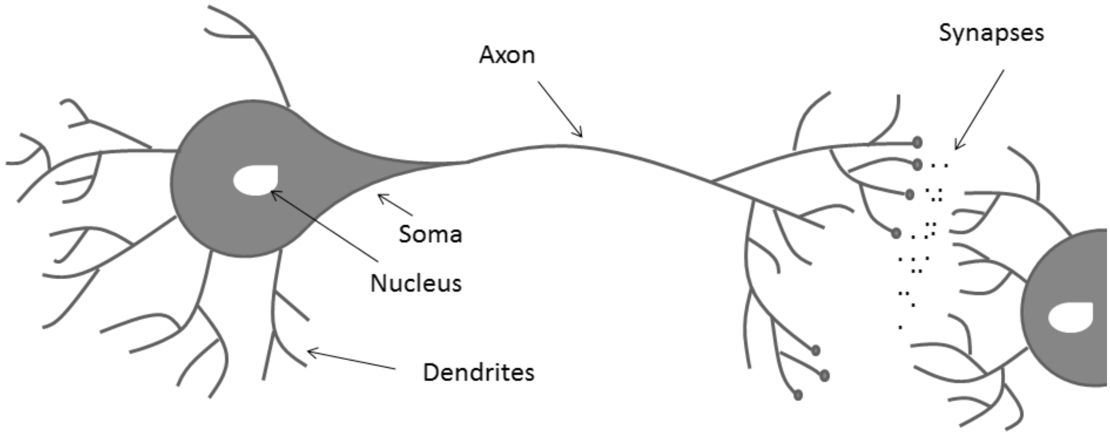


Figure 14: The structure of a biological neuron

Billions of nerve cells compose biological neural networks. Biological neural networks enable the understanding of the human's behaviors and environment. They learn relationships between events using their perception and understanding mechanism that is developed from the five sense organs that received information. Artificial neural networks are developed from these features of biological neurons [Öztemel, 2006]. A simplified non-linear model of an artificial neuron is depicted in Figure 15. In this artificial neuron, the weights represent the synapsis that connects the biological neurons. Similar to the body (soma) that sums the incoming signals, the sum function in artificial neural networks sums the incoming information and gives the result. The activation function limits the permissible amplitude range of the output signal to some finite value [Haykin, 2005]. A non-linear activation functions provide the non-linearity to the network.

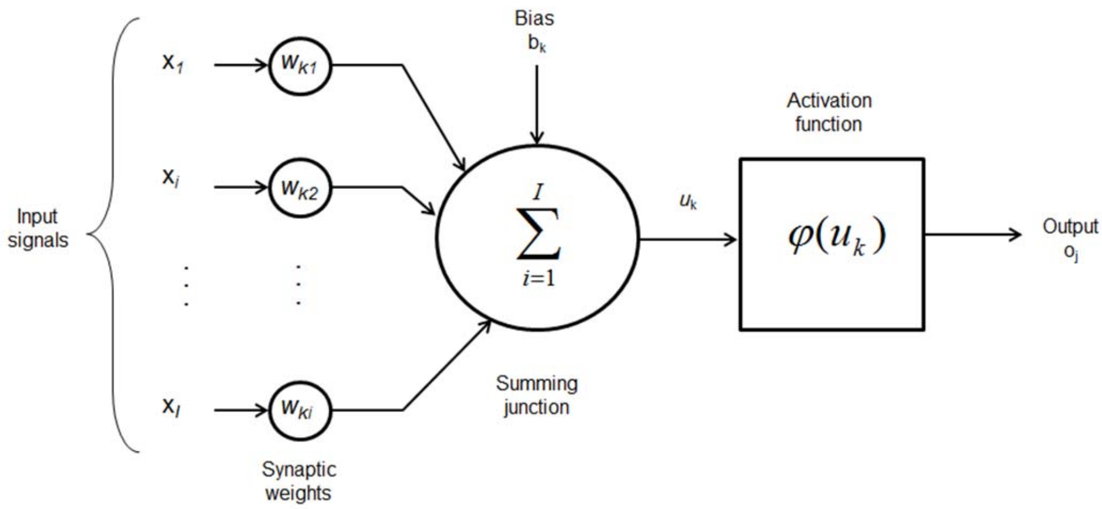


Figure 15: Nonlinear model of an artificial neuron [source: Haykin, 2005]

An artificial neuron involves a set of synapses connecting links, which is characterized by a weight of its own. A signal x_i at the input of synapse *i* connected to neuron k is multiplied by the synaptic weight w_{ki}. A sum function is used for the input signals, which are weighted by the respective synapses of the neuron. For limiting the amplitude of the output of the neuron an activation function is required. An external parameter bias can also be used to increase or decrease the network input of the activation function depending on whether it is positive or negative, respectively [Haykin, 2005]. Since in this example, the output has only one unit, here the j=1. The output o_j can be found by writing the following equations:

$$u_k = \sum_{i=1}^{I} w_{ki}.x_i \tag{6}$$

and

$$o_j = \varphi(u_k + b_k) \tag{7}$$

where:

x_i = input signals

w_{ki} = synaptic weights of neuron

u_k = linear combiner output

b_k = bias

φ = activation function

o_j = output signal of the neuron

Haykin [2005] offers the following description for neural networks in machine learning: *"A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:*

1. Knowledge is acquired by the network through a learning process.

2. Interconnection strengths known as synaptic weights are used to store the knowledge."

The history of artificial neural networks goes back to the 1890, where the first studies about human brain are conducted. Direct works about artificial neural networks are started in the 1940's. We can divide the artificial neural network (ANN) studies in two groups; before and after 1970. In 1960, some scientists argued that ANN cannot generate a solution for non-linear problems. Hence unproductive period of ANN started and financial support of it is interrupted. After 1970, with the studies of Kohonen and Anderson about associative memory and development of back-propagation model ANN became again significant [Öztemel, 2006].

Artificial neural networks find applications in different fields such as modeling, time series analysis and signal processing control or pattern recognition [Haykin, 2005]. Because of their important properties and benefits as follows [Hammerstrom 1993]:

- *Adaptivity:* Neural networks learn from training data. Thus they find solutions from the data presented to them before.
- *Generalization:* they can produce reasonable outputs to data that only broadly resembles the data they were trained on before. They can also work with imperfect or incomplete data, providing a measure of fault tolerance.
- *Nonlinearity:* Neural networks are nonlinear and are very suitable for real world applications that are generally nonlinear.
- *Parallelism:* neural networks have a highly parallel structure. They can execute independent operations simultaneously.

3.3.2 Artificial Neural Network Architecture

Artificial neural networks are considered as directed graphs in which artificial neurons are nodes and directed edges (with weights) are connections between neuron outputs and inputs [Jain et al., 1996].

Fundamentally, depending on their architectures the neural networks can be considered in two classes [Jain et al., 1996]:

- Feed forward networks
- Recurrent/feedback networks

Feed-forward networks have one directional graph while recurrent networks have bidirectional graphs because of feedback connections (see Figure 14) [Jain et al., 1996].

Different architectures yield different network behaviors. Feed-forward network produce only one set of output values rather than a sequence of values from a given input. Therefore, the feed-forward networks can be considered as static networks, while recurrent/feedback networks are dynamic systems. When a new input pattern is presented to the network, the neuron outputs are calculated. The input to each neuron can be modified due to the feedback paths. This leads the network to enter a new state [Jain et al. 1996].

The following paragraph gives an overview of recurrent/feedback networks and the next section deals with feed-forward neural networks in detail.

- ***Recurrent/Feedback Networks:***

Recurrent neural networks differ from other networks in that they consist at least one feedback loop, besides forward loops and each neuran can also consists of a self-feedback loop [Haykin, 2005]. Especially, in modelling and training of dynamic systems feedback loops are important to consider time delays [Öztemel, 2006]. Generally, recurrent networks can be considered in two groups; full recurrent and partial resurrent networks [Unadkat et al., 2001]. Competitive networks, adaptive resonance theory (ART) and, Kohonen's and Hopfield's models are examples of recurrent/feedback networks (see Figure 16).

The adaptive resonance theory (ART) model is an example for unsupervised learning method. ART networks are developed during Grosberg's studies on brain functions. This model has three characteristics: normalization, comparability and short term memory. Normalization means the adaptivity of the current situation; the comparability is the interpretation of occurred events on the environment. The sort-term memory stores the current events and assists by decision making. The main advantages of these networks are that they have capability to learn in on-line mode in unsupervised manner [Öztemel, 2006].

The SOM (self orginizing map) model was developed by Kohonen and generally, used for classification problems. The Kohonen's SOM networks have great ability to classify the input vectors and to learn the distributions of input vectors. The main feature of these networks is the unsupervised learning [Öztemel, 2006].

Hopfield is a single-layer and feedback model. All the process units are both input and output units, simultaneously. Hopfield uses an energy function to keep the link weights of the network. Hopfield can be considered in two groups: discrete and continuous Hopfield networks [Öztemel, 2006].

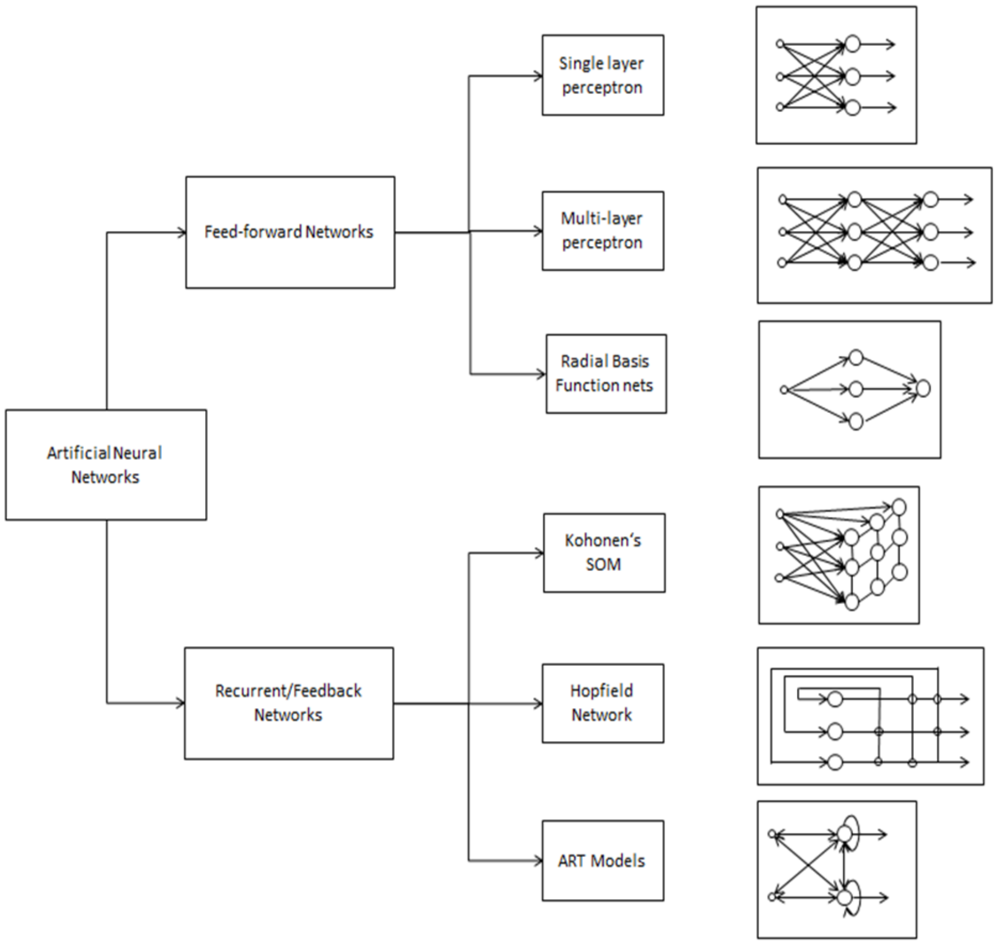


Figure 16: Types of artificial neural networks [source: based on Jain et al. 1996]

- ***Feed-Forward Networks:***

Feed-forward neural networks provide a general framework for representing non-linear functional mappings between a set of input variables and a set of output variables. This is achieved by representing a non-linear function of a single variable, called activation function [Bishop, 1995]. As shown in Figure 17, the feed forward neural networks are fully connected in the sense that each node in each layer of the network is connected to every other node in the adjacent forward layer [Haykin, 2005]. According to their architecture feed-forward networks are considered in two groups: single and multi-layer feed forward neural networks. As the name suggest single-layer feed-forward network has only one layer, while multi-layer network consists two or more layers. A simplified multi-layer feed forward network with one hidden layer is depicted in Figure 17.

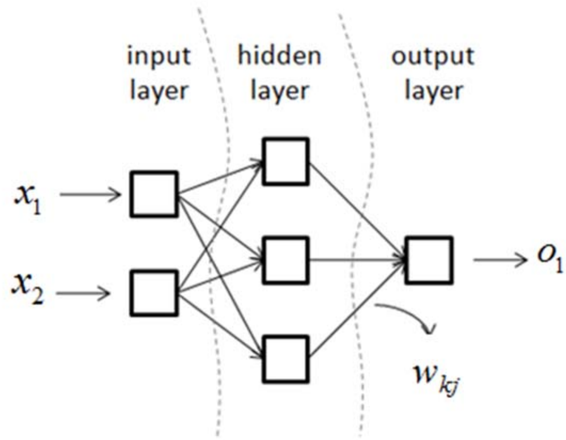


Figure 17: A simplified structure of multi-layer perceptrons

Multi-layer perceptrons consist of at least three layers; an input, an output layer and one or more hidden layers. Input layer receives the activity pattern (here χ_1 and χ_2) and send it to the adjacent hidden layer. Each processing unit contains one input and one output. This output is deployed to all processing units of the adjacent layer. Hidden layers receive the outputs of input layers and send them on after processing to the output or to next hidden layer if one exists. The output layer receives the outputs of hidden layer and processes and generates the actual response (y_1) of the network. In the network, each neuron or processing unit is fully connected to all the processing units of the previous layers [Öztemel, 2006].

3.3.3 Activation Function

For limiting the amplitude of the outputs an activation function is required in the artificial neural networks [Haykin, 2005]. Generally the activation functions squashes the values between 0 and 1. The threshold function was firstly proposed by McCulloch and Pitts in their first artificial neural networks. The threshold function separates the outputs either 0 or 1. This is generally used to classification into groups [Jain et al., 1996]. The threshold function is also used for output layer of multi-layer perceptrons for classification problems. The threshold function is formulated as:

$$\varphi(u) = \begin{bmatrix} 0 & if\ u < 0 \\ 1 & if\ u > 0 \end{bmatrix} \tag{8}$$

where u is the output of previous layer.

In the piecewise-linear function, the amplification factor inside the linear area of operation is assumed to be unity. Therefore, this function can also be viewed as an approximation to a nonlinear amplifier [Haykin, 2005]. The piecewise-linear function is applied using:

$$\varphi(u) = \begin{bmatrix} 1, & u \geq 0.5 \\ u & 0.5 > u > -0.5 \\ 0 & u \leq -0.5 \end{bmatrix} \tag{9}$$

It should be noted that the threshold in threshold and in piecewise-linear functions can be different determined.

The most used activation function is the sigmoidal function. It brings to the network the nonlinearity, which helps to the learning of the network due to its derivable characteristic.

$$\varphi(u + b_k) = \frac{1}{1 + e^{-u + b_k}} \tag{10}$$

where u represents the output of previous layer and b_k is the bias term. The bias has an effect to increase or decrease the net output of the activation functions [Haykin, 2005].

The feed-forward networks (multilayer perceptrons) employ mostly either the threshold function or the sigmoidal function [Jain et al., 1996]. However, the best suitable function for each problem should be chosen by the system designer [Öztemel, 2006].

3.3.4 Learning Methods

Neural networks have the ability to learn automatically from given samples. Instead of following pre-defined rules that are specified by human experts, neural networks appear to learn underlying rules from the given collection of representative samples. This feature makes them attractive and provides an advantage over traditional expert systems [Jain and Mao 1996]. Generally, three methods are used as learning method according to system and used learning algorithm. The term “teacher” is used by a description of these learning types.

- ***Supervised Learning:***

In supervised learning, the system learns under the supervision of a teacher. The teacher gives both the input and the target set to the system. The system should address the outputs of the system to the target data set that was given by the teacher [Öztemel, 2006]. The network weights are adapted under the combined influence of the error signal and the input vector. The adjustments of the weights are generally made by an iterative procedure until the output produced follows the target signal closely in some statistical terms [Khan, 1996]. Multilayer perceptrons, which is used in this study, learn using this learning strategy.

- ***Unsupervised Learning:***

In unsupervised learning, the system learns without a teacher. Only the input data set are given to the system. This learning type is generally used for classification problems. When the learning process is finished, a labeling is required to define the meaning of the outputs. Adaptive Resonance Theory (ART) Networks are typical examples that learn unsupervised manner [Öztemel, 2006].

- ***Reinforcement Learning:***

In reinforcement learning, a teacher also helps during learning. However, the target set for each input set is not addressed in this learning method. The system produces the output, after that, the teacher produces a signal, dependent on if the output is correct or not. With this information from teacher the system further learns up to an acceptable error rate. Linear vector quantization (LVQ) networks are common examples of for this learning method [Öztemel, 2006].

3.3.5 Learning in Multi-layer Perceptrons

Multi-layer perceptrons (MLP) have been applied successfully to solve some difficult and diverse problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm [Haykin, 2005]. This method is similar to error minimization which is an attempt to fit a closed-form solution to a set of data points such that the solution deviates from the exact value by a minimal amount [Abdennour and Ghamdi, 2002]. Though some discussion and research is conducted ([Huang and Huang 1991], [Yuan et al., 2003]) about the optimal number of hidden layers and the number of neurons there is not any fixed structure for the multilayer perceptrons. The number of hidden neurons as well as the number of neurons in a hidden layer varies according to problem types. The key issue in the feed forward neural network is the determination of the link weights between the neurons in each layer.

Multilayer perceptrons learn automatically from given samples. Instead of following pre-defined rule that are specified by human experts, neural networks appear to learn underlying rules from the given collection of representative samples. This feature makes them attractive and provides an advantage over traditional expert systems [Jain et. al., 1996].

Basically, training or learning of multilayer perceptrons consists of two passes: forward and backward pass. An activity pattern (input vector) is applied to the receiver of the network in the forward pass. Its effect propagates through the network layer by layer and a set of outputs is produced as the actual response of the network. In the forward pass the synaptic weights of the network are all fixed, while during the backwards pass are all adjusted according to the error-correction rule. The current output of the network is subtracted from target response to calculate the error. This error is propagated backwards through the network and the synaptic weights are adapted to bring the current output closer to the desired response in a statistical sense [Haykin, 2005]. An example for learning of MLP networks can be found in Appendix A.

Generally, the learning process can be preceded in two ways: sequential and batch mode. In sequential mode, weight adjustment is performed after the presentation of each training sample. In other words, the first sample pair in the epoch is presented to the network, forward and backward computations are performed, weights are adjusted and the second pair is sent to the network. This process is continued until the last sample pair. The sequential mode is exclusively relevant for online application. On the other hand, in batch mode of back-propagation learning, the adjustment of the weights is performed after the presentation of all the training samples [Haykin, 2005].

3.3.6 Stopping Criteria

Basically, the learning method of multilayer perceptrons is similar as that of the other networks. The main philosophy of the learning is to find the right values of each weights of the network. The desired weights are not known before. Therefore to understand and to explain the behaviors of neural networks is not possible [Öztemel, 2006]. The main question in the learning process is the stopping criteria. There are no any well-defined criteria for stopping the learning [Haykin, 2005]. The user cannot know if the error rate to global minima reached or only a local minima. To overcome this problem it is logical to define an error tolerance before. The problem to solve is to determine all weights to reach minimum or predefined error rate. As in Figure 18 depicted, suppose that w* represents the weight vector for minimum error signal. This weight is best suitable for this problem. To reach this value ΔW change in weights are realized in each iteration to provide ΔE change in the error signal. It is significant that the error space is not always one or two dimension. Figure 17 shows a more complex error space. In this case to find suitable weight w* could be not possible. Despite w* is the best suitable weight for this problem, there is no any information to reach this value. The neural network tries to find this weight itself, but it could not be possible every time. Therefore the users define an error tolerance for his network before learning according to importance and safety of his system. If the network reaches the error tolerance it is supposed that the system finished its learning. The other solution in Figure 19 (w_0, w_1,…) are the local solutions and if one of this peak lower the error tolerance is reached, the system stops the learning. Therefore determining of the error tolerance is significant [Öztemel, 2006].

After the learning process the network must be tested using new (another part of sample data) for its generalization performance. The learning process is completely stopped when the generalization performance is adequate [Haykin, 2005].

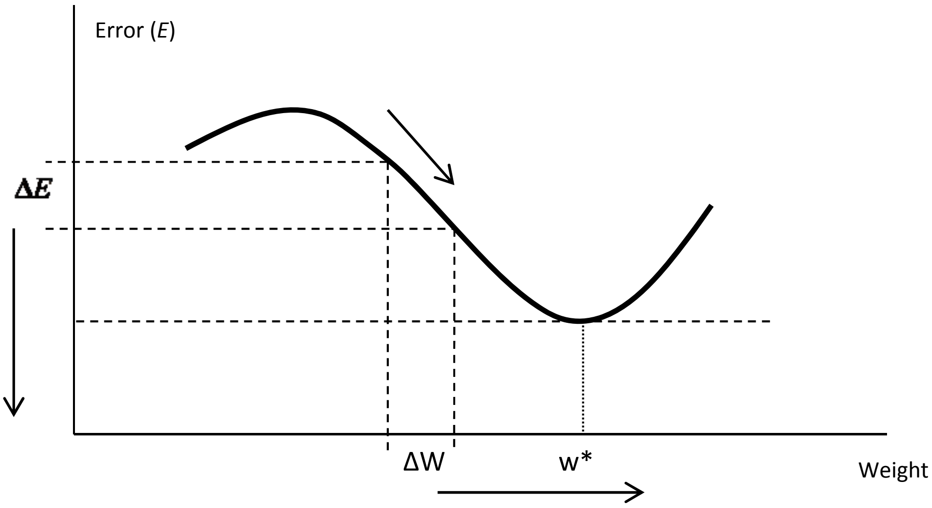


Figure 18: system learning and error space [source: Öztemel, 2006]

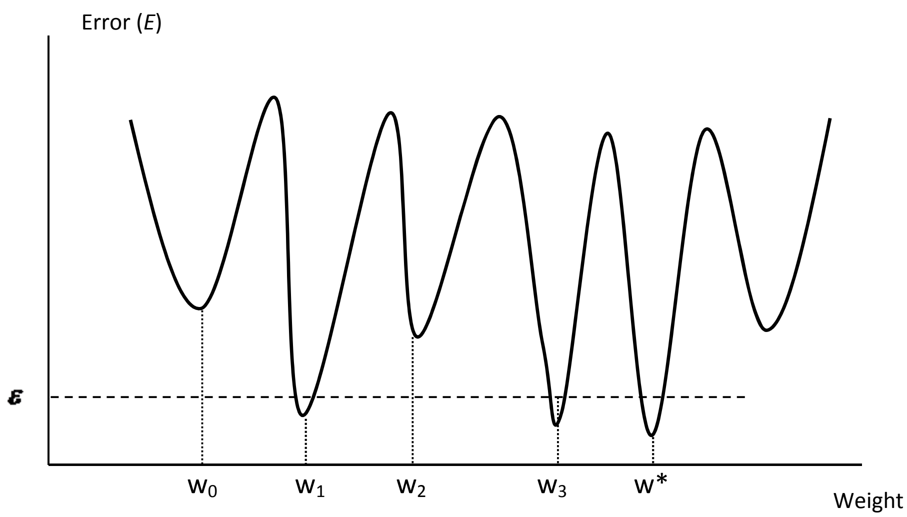


Figure 19: the effect of starting point in multilayer perceptrons [source: Öztemel, 2006]

4 Traffic signal control in developing large cities

The cities of developing countries face major problems in traffic control due to effect of urbanization rate, increasing population and car ownership with improving social and economic development. These changes have an effect also on traffic demand and this requires new infrastructure or effective use of existing infrastructure. Unfortunately, the cities in developing world spend their budget founds for construction of interchanges without considering improvement in traffic signal control infrastructure and methods.

To determine the current situation in signal control in developing large cities two different preliminary studies have been conducted during this research. The first preliminary study was to determine signal control strategies, architectures and equipment in developing large cities in comparison with developed one. The objective of the second preliminary study was to determine variations in saturation flow, which directly influence the capacity of signalized intersections and hence also replicate also traffic situation in those cities.

This chapter consists of three main parts. The first part summarizes the results from a web-based survey dealing with traffic signal control in large cities, while the second part presents the results of a study about capacity utilization at signalized intersections particularly based in one large city. The last part describes the traffic control related problems of developing large cities based on these studies.

4.1 Comparative study on control methods

4.1.1 Literature review on comparison of signal control method

To determine the use of signal control strategies and related equipment in developing large cities in comparison with developed cities a questionnaire-based survey was conducted with the transportation departments of the participant large cities.

Before the interview, a literature review was carried out in order to assess the status in cities and also to compare the current situation with the situation in previous studies. The following paragraph summarizes the comparative studies that investigated particularly traffic control in cities.

Gardner [1993] has described and evaluated deployments and costs of traffic signals in developing countries. He conducted a questionnaire involving 17 cities from developed as well as less developed countries. The main focus of this work was the number of controllers and the cost of traffic signal control in different cities. Fellendorf et al. [1999] investigated the system components of 9 European cities. They concentrated on the system architecture, data collection, processing and archiving structure, and the communication system of the some cities. Gordon and Braud [2009] conducted another comprehensive survey in the United States with 7 cities. The population of the responding cities varied from 40,000 to 900,000 inhabitants. They intended to provide a methodology for identifying operations and maintenance objectives and performance measures of signal control departments. Von Mörner et al. [2009] conducted a comprehensive survey with 22 German cities. Besides general questions about signal control in Germany, they concentrated on the detailed planning of signal controller, microscopic and macroscopic strategies, and energy costs etc. For most questions they also compared former West German cities with cities in the new states. Schmöcker and Bell, 2009] conducted a questionnaire-based survey with 28 cities around the world. They focused on traffic signal control and general traffic management operations in that survey. Stevanovic [2010] conducted a survey of 30 US cities and 15 cities from other countries that have implemented adaptive traffic control systems. The benefits of adaptive traffic signal control and its difficulties are investigated via interviews with experts of corresponding agencies. Stevanovic compared the workings principles of adaptive systems and their system requirements (hardware, software, communications etc.) as well as evaluated the cost/benefit of those systems. Kaparias, et al. [2010] investigated ITS implementations within 34 European cities including questions regarding Urban Traffic Control Systems (UTCS). Their study involved general statistics of the transport systems, organisational structures, monitoring and forecasting, provision of traffic information and urban traffic control. The special focus of their study was traffic demand management, traffic control centres, public transport infrastructure and parking facilities. They also gathered statistical figures on traffic signal density measured in signals per kilometre of road network in 24 participant cities. However, besides the study by Gartner [1993], these studies do not make a comparison between developed and developing cities and consider the general situation in all cities.

Except the study in [Gardner, 1993], generally previous comparative studies investigated either general transportation issues or signal control in developed cities. In this context, there was a need an investigation of the signal control strategies in developing large cities. Therefore, the objective of this comparative study was to investigate the current situation particularly about traffic signal control in large cities and to evaluate the situation in developing large cities in comparison with developed cities.

4.1.2 Survey conducted in 2010

In this study, the situation in developing cities is researched in comparison to developed ones. In Mai 2010, the invitation letters were sent out to traffic departments of large cities around the world. For this web-based expert's interview, 42 transportation departments were invited to participate. Overall, 19 large cities have completed the survey, thus the response quote lay at approximately 45%. The cities are selected according to size and level of development. It is also regarded to take as far as possible the same number of participants from developing and developed countries. There were two main problems in obtaining responses. The first being that finding the relevant authority and the contact person in some cities and the second being that transportation authorities could not respond to some of the detailed questions because of their administrative structure.

Each region or country has a different definition and tradition in traffic signal control. Therefore the participants had to be contacted again for further explanation and questions. Some cities outsource signal control operations to several companies; therefore they could not gather the answers. This also reduced the response rate. Some general statistics were compared with the World Bank, OECD and UN statistics for a crosscheck. Considering the population Zurich was the smallest city with 400,000 populations and Shanghai as the largest city with that of 16 million populations. Demographic statistics, the number of passenger cars per 1000 inhabitant and the number of traffic signals of participant cities are listed in Table 3. The following section reports the findings of interview with city governments.

Table 3: demographic statistics, car ownership and traffic signals of participant cities

Name	Country	Population 2010[1)]	Population increase (2005-2010)[1)]	Density (pop./km^2)[2)]	Car ownership[3)]	Number of Signalized intersections[3)]
Shanghai	China	15,789,000	1.70	6.0	101-200	7500
Beijing	China	11,741,000	1.82	5.0	201-300	3700
Istanbul	Turkey	10,530,000	1.62	9.7	101-200	1747
London	U.K.	8,607,000	0.24	5.3	301-400	6183
Berlin	Germany	3,423,000	0.19	3,5	301-400	2084
Cape Town	S. Africa	3,357,000	1.68	5.0	201-300	1370
Izmir	Turkey	2,724,000	1.82	10.7	101-200	430
Toronto	Canada	5,447,000	1.57	2.7	401-500	2200
Taipei	China	2,651,000	0.35	7.3	201-300	2338
Bursa	Turkey	1,589,000	2.24	13.4	101-200	145
Hamburg	Germany	1,777,000	0.43	2.5	501-600	1713
Vienna	Austria	1,710,000[2)]	1.04	3.8	401-500	1265
Munich	Germany	1,300,000	0.72	2.9	401-500	1130
Belgrade	Serbia	1,096,000	-0.18	4.7	401-500	507
Konya	Turkey	978,000	2.32	8.1	101-200	131
Amsterdam	Netherland	1,044,000	0.41	2.5	201-300	380
Athens	Greece	3,256,000	0.17	4.8	401-500	1600
Copenhagen	Denmark	1,087,000	0.04	2.6	301-400	308
Zurich	Switzerland	400,000[3)]	0.35	3.0	301-400	400

[1)] UN, United Nation Statistics
[2)] Demographia 2012
[3)] Participant responses

4.1.3 Results of survey 2010

- ***Number of traffic signals and its yearly increase***

Generally, the number of traffic signals can be determined either per 1000 inhabitant or per urban road network (km). In this study, the number of traffic signals per 1000 inhabitant is used, so the result could be compared with some previous research studies.

The average number of signal controllers in participant cities from western European countries is 0.9 controllers per 1000 people, whereas this value reduces to 0.6 controllers per 1000 people when considering all participant cities (see Figure 20). In developing countries, in comparison with previous comparative studies about signal control the number of signalized intersections has increased quite a lot in during last decade. However, the average number of signalized intersection per 1000 inhabitant lay at 0.3.

The total number of signal controllers varies not only according to the population but also the human development index (HDI) of United Nations. For example, Shanghai is the largest city with the most number of signal controllers (7500 SC), but per capita, Shanghai has 0.4 signal controllers per 1000 people.

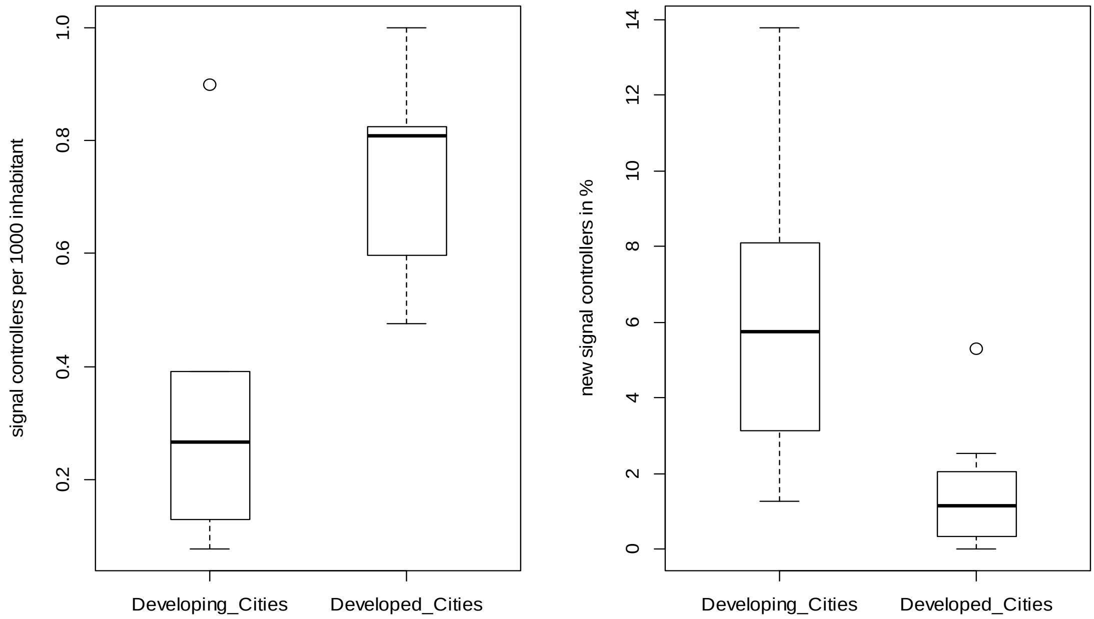


Figure 20: Number of traffic signals per 1000 inhabitant and its yearly increase in (%)

Reference quantity: n_1 (developing cities) = 10

n_2 (developed cities) = 9

Source: own survey [2010]

In London, the number of signal controllers has increased from 1800 to 6300 in the 20 years. The number of signal controllers per 1000 people was only 0.26 in 1991 [Gardner 1993], but today it has increased to 0.82. In 1991 the transportation authorities in Istanbul operated 400 signal controllers [Gardner, 1993]; they installed 1200 new controllers and increased the ratio from 0.04 to 0.12. However, it is still lower than the European average. Figure 10 compares developing and developed cities in the number of signal controller per 1000 inhabitants. The outlier in the category "developing countries" is the City of Taipei. Taipei is also above the European average; owning 2300 signalized intersections. As expected, the percentage of new installation of signalized intersections in developing cities is higher than in developed cities. The largest cities Istanbul, Beijing and Shanghai are also the cities that installed more new signalized intersections. Istanbul installed about 160 new signal controllers in 2010, while Beijing and Shanghai took 300 and 600 new signalized intersections into service, respectively.

The number of signalized intersections and its yearly increase represents not only the increase in population and car ownership, but also public interest in traffic control measures in urban areas.

- ***Age of controllers***

The questionnaire also asked how old the signal controllers are. The results about the ages of the signal controllers indicate that most of the participant cities replace old controllers with new ones. Only 9% of the total controllers are older than 20 years. This value increases in developed countries up to that of 20%. 61% of signal controllers are younger than 10 years old. In general, developing cities have younger controller than the developed cities. The developing cities have 35% of their controllers at less than 5 years old, while this ratio decreases to 21% in developed cities. The reason of this difference in proportion between developing and developed countries can be that due to an increase in population, in the number of passenger cars and as a result the increase in the number of signal controllers (and so newer controllers) in developing countries. They have younger signal controllers and fewer old controllers than the developed countries (see Figure 21). However, it is still quite interesting that only 4% of the total controllers are older than 20 years in developing cities.

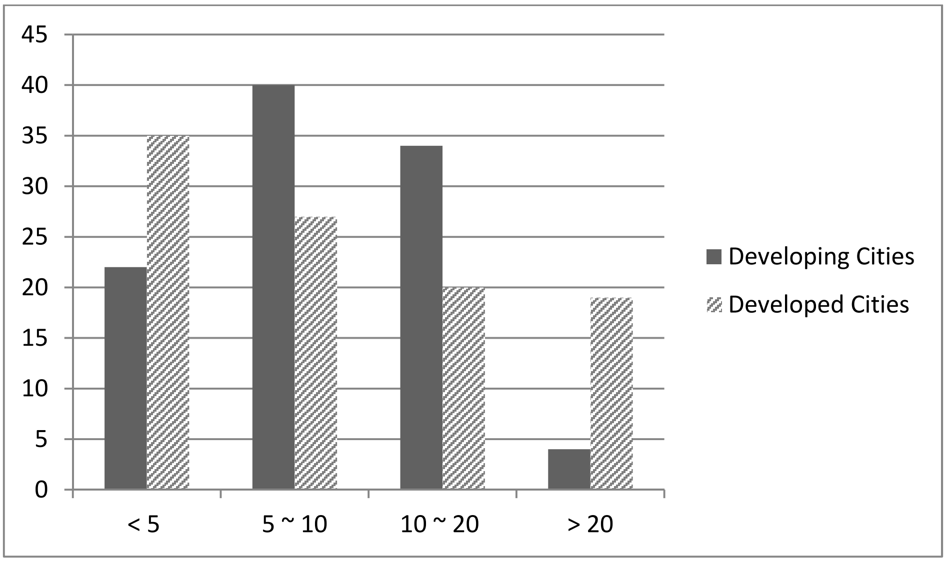


Figure 21: Age of signal controllers

Reference quantity: n_1 (developing cities) = 10

n_2 (developed cities) = 9

Source: own survey [2010]

- ***Signal Control Architecture***

Generally, transportation authorities operate their signal controllers using three types of signal control architecture: 1) central with only one traffic computer in the traffic control center, 2) central control using more than one regional computer with central controlling process and 3) distributed system architecture, where the intersections have more powerful signal controllers. In the distributed system each intersection calculates and operates itself. The central controllers provide the coordination between intersections. A well-known example of central control using more regional control is London and the SCOOT system. Munich, Hamburg and Shanghai operate their signals using more regional computers. Some cities use only one central computer in the control center. However, they do not communicate them at shorter intervals, since they do not use traffic actuated or traffic responsive control methods. Only warnings about defective detectors and bulbs are sent to the center. Aggregated responses to signal control architecture and the numbers of regional computers are demonstrated in Table 4 and 5. In this question, the total participant cities are considered without grouping them according to the development level.

Table 4: ratio of signal control architecture in participant cities

	Sample size (Number of Cities)	Central control (with regional computers)	Central control (without regional computers)	Distributed control
		(%)	(%)	(%)
Signal control architecture	19	37	37	26

Table 5: ratio in number of regional computers in participant cities

	Sample size (Number of Cities)	No regional computer	Regional computer (1 to 9)	Regional computer (>9)
		(%)	(%)	(%)
Number of regional computers	19	47	37	16

- ***Signal Control Strategies***

One of the most important parts of this interview was the traffic signal control strategies of the cities. According to their traffic engineering expertise, own experience, budget or socio-economic development of the country, cities have different signal control strategies. Looking at the types of signal control strategies at intersections or networks it can be found that most developed cities have real-time coordinated control system (traffic-actuated, TRPS or adaptive). This group reaches almost 65% of the total control types. Whereas, in developing countries the most preferred types is with 50% the isolated control.

Traffic adaptive control is considered separate from other real-time control types. In developing countries only 5% of the total intersections are controlled with adaptive strategies, this value increases to 10% in developed countries. London with its years of experience in SCOOT operates 41% of their intersections with adaptive control. Besides London some more cities have changed a part of their system to adaptive control. Cape Town (17%), Beijing (14%) and Toronto (15%) also operate some intersections with adaptive control strategies. Some cities also reported that they are currently testing out adaptive systems. Due to the trend and technological development in communication systems and adaptive control methods it is expected that the ratio may be changed in near future.

Isolated signalized intersections and pedestrian crossings were asked in the questionnaire separately. However, since they both are isolated and uncoordinated, these two types are combined. Traffic actuated and traffic responsive control types are generally confused due to the local and network level considerations. Therefore this was asked in a very detailed way to receive correct information from different regions, all of which have their own definitions. The ratio of fixed-time coordinated arterials is almost the same in both developing and developed cities. Control types and their ratio are illustrated in Figure 22.

Figure 22: signal control strategies in participant cities

Reference quantity: n_1 (developing cities) = 10

n_2 (developed cities) = 9

Source: own survey [2010]

- ***Public Transport Priority***

Priority at signalized intersections for public transport vehicles can reduce their travel time and increase the operating speed and of course increase the capacity of them. An intelligent signal controller and special components are required for public transport priority. We also asked the participant about the number of signalized intersections with PT priority.

The huge difference between developed and developing countries appears in public transport (PT) priority. Zurich with 80% PT priority and Vienna with 50% PT priority are the leaders in the participant cities in this category. In developing countries, Beijing reached the highest value with 8% PT priority of total signalized intersections. Beside engineering experience, the capacity or structure of signal controllers can be mentioned here as a main reason of low rates in developing countries. Some participant agencies also reported that the public transport priority is one of the most important reasons to change their system architecture. Figure 23 shows the difference in PT-Priority between developing and developed cities.

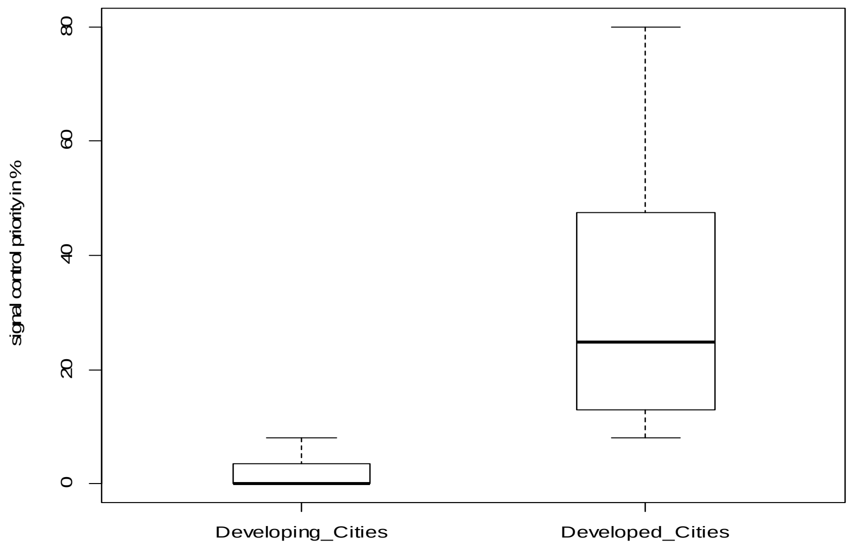


Figure 23: Percentage of public transport priority at intersections

Reference quantity: n_1 (developing cities) = 9

n_2 (developed cities) = 9

Source: own survey [2010]

- ***Timing plans and updates frequency***

Generally, independently if fixed time control method is implemented or traffic adaptive, traffic signals work with timing plans or frame plans. However, the timing plans or frame plans are adjusted in traffic adaptive control online, whereas in fixed time and traffic responsive plan selection methods the timing plans should be updated in scheduled or unscheduled periods. The participant agencies were also asked about the number of timing plans and update frequency in timing plans. Most of the participant agencies have more than 4 periodic and non-periodic timing plans. Shanghai and Bursa operate with the most timing plans. Both large cities have 8 to 10 weekday and weekend timing plans. Hamburg, the second largest city of Germany, uses normally 4 timing plans but they can increase them up to 16 with non-periodic timing plans. Hamburg also keeps evacuation timing plans for seaport areas. 40% of the participant cities do not have any scheduled timing plan update. Some of them are using full-traffic actuated signal control systems; therefore they do not need planned actualization. Nearly 50% of the participant cities actualize their timing plans in every 1-2 years. Responses of participant cities about timing plans and update frequencies are listed in Table 6.

Table 6: Number of timing plans and frequency of revision in 19 cities [source: own survey in 2010]

Agency	Number of weekday timing plans	Number of other periodic and non-periodic timing plans	Revise timing plans
1	6	6	every 1-2 years
2	4	6	every 1-2 years
3	10	10	every 1-2 years
4	5	1	every 1-2 years
5	5	1	every 1-2 years
6	3-5	2-3	no scheduled update
7	4	2	no scheduled update
8	4-16	4-16	1-10
9	5-7	individual	no scheduled update
10	no response	no response	no response
11	4	2	no scheduled update
12	3	5	every 3-5 years
13	4	0	no scheduled update
14	3	3	no scheduled update
15	no response	no response	no response
16	7	2	every 1-2 years
17	4	2	no scheduled update
18	6	1	no scheduled update
19	8 to 10	8 to 10	every 1-2 years

- ***Installation and maintenance cost of signalized intersections***

Some of the participants avoid responding the questions about installation and maintenance costs of the signal controllers due to privacy. The installation cost in European and developed countries starts from €80.000 and reaches up to €350.000 depending on control type and intersection size. By absolute value, it varies less in developing countries ranging from €5000 up to €50000 according to size and the type of intersections. Most of participants only responded the average maintenance cost. The average maintenance cost is about €2000 in developing cities, while the value increase to €5000 in developed one. The variations of cost of installation and maintenance are depicted in Figure 24 and 25.

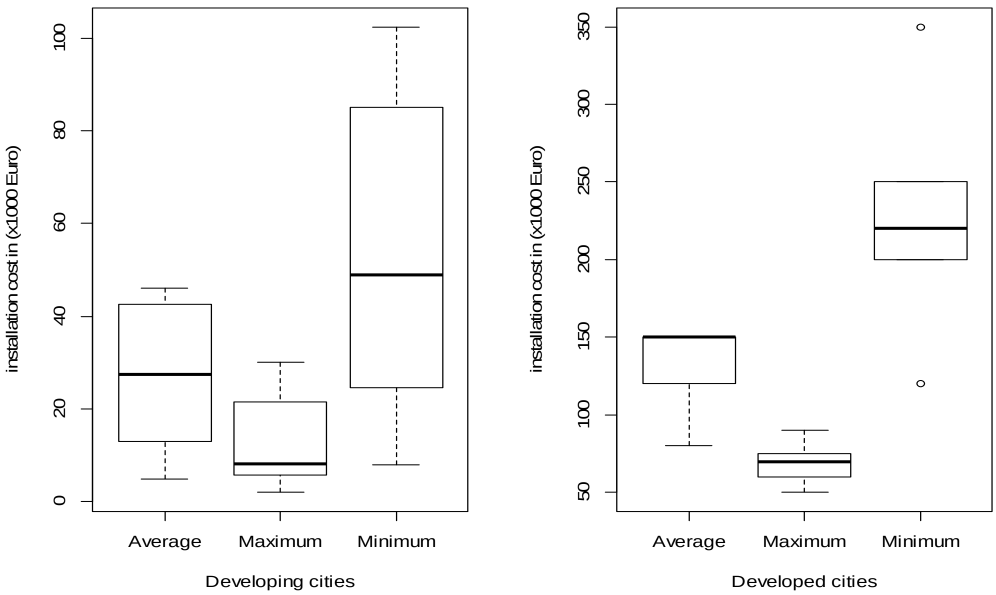


Figure 24: Installation cost of traffic signals

Reference quantity: n_1 (developing cities) = 8

n_2 (developed cities) = 7

Source: own survey [2010]

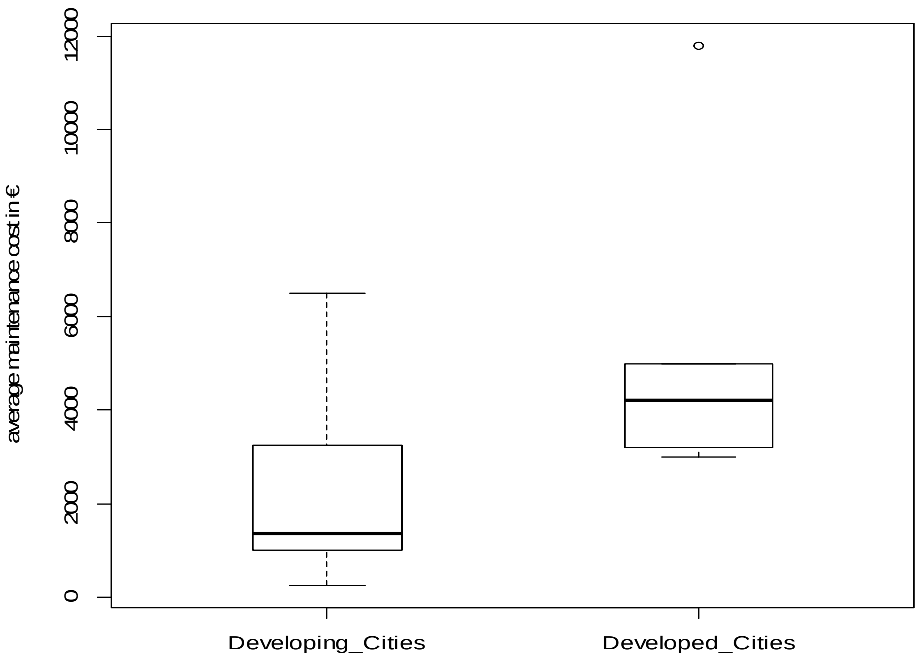


Figure 25: Average maintenance cost in participant cities

Reference quantity: n_1 (developing cities) = 8

n_2 (developed cities) = 7

Source: own survey [2010]

- ***Intent to change existing system***

At the end of the questionnaire the participant agencies were also asked about their intention in future technologies and control systems. In total 37% of the participants reported that they are planning to change their existing systems. Most of these responses are from developing countries. 52% of the all participants responded that their existing system fulfils their requirements. If we look only at developed cities, only one of 9 cities is thinking about changing the current system. 32% of replies stated that they have own expertise with existing system and therefore they do not intend on changing their systems. Some participants point out that they do not have any additional budget to change the system; these responses were about 15%. Some cities do not want to change the whole system, but they would like to improve their system with some parameter settings. Table 7 shows the responses all together in this category. Of the cities questioned, 50% of the cities responded in saying that they are in favour of changing the current system with the other 50% in favour of retaining their current system. Reducing the cost of retiming of signals, Handling with oversaturated conditions and fluctuating traffic conditions are aims of participant cities, which are going to change their system. They also want to realize priority for public transport vehicles with a new system.

Table 7: reason and ratio of changing the existing system (within 19 participants)

	%
Yes; to realize public transport priority	11
Yes; Handling fluctuating traffic conditions during weekdays	9
Yes; Handling oversaturated traffic conditions	14
Yes; Reducing costs of retiming signals	8
No; Our existing system fulfills our requirements	25
No; We have own expertise with our system	17
No; We have not any additional budget for a new system in the near future	8
No; We only adapt parameter settings within our system	8

4.1.4 Conclusions from signal control survey 2010

As an interim summary, what can be noted is that the number of controllers and control methods differ in developed and developing cities due to several reasons. However, the development in developing and emerging countries brings also interest and change to traffic control in urban areas. All participant cities prefer to prioritize public transport. Some of them can realize this and some are intending to change their system to provide priority to public transport. This requires more or less change in the existing system. The cities should develop their own system structure or install one of existing systems, which is suitable for their existing infrastructure and own experience.

In addition to this, in the questionnaire, the number of detectors is also asked about in detail (i.e. number of inductive loops, the number of video detectors etc.). However, we could not receive sufficient data for statistical analysis. But, in general, the cities use 1 to 5 detectors per intersection. Of course, the number of detectors depends on the major control method of the cities.

4.2 Capacity Utilization at Signalized intersections

4.2.1 Saturation Flow Measurements

The saturation flow rate is one of the most important input parameters in the calculation of cycle length as well as the determination of the intersection capacity. The existing research studies in Western countries, Europe and the United States indicate that the actual saturation flow rate is between 1700 pcu/h and 2080 pcu/h. However, in developing countries it is between 1200 veh/h and 2000 veh/h for the same conditions (3,5m wide gradient lane, without opposite flow). The difference of scattering lays in pedestrian activity and other external and regional factors [Turner and Harahap 1993]. Some further research studies [i.e. Webster and Cobbe, 1966 and Akcelik 1982] indicate that the geometrical and behavioral factors as well as the environment, influences the saturation flow rate. However, these local characteristics in new guidelines are not considered in detail. Particularly, in developing countries due to the behavior of traffic participants (drivers and pedestrian) and the deficit intersection area the field saturation flow can vary more than as expected. Hence, generally, saturation flow can be overestimated when using western guidelines. However, proper determination of its value is significant, since the saturation flow rate is one of the main parameters of traffic signal control.

Particularly, developing cities has some distinctive factors that impact the capacity utilization of signalized intersections. For example, lane disciplines, pedestrian behaviors or behavior of drivers of some special vehicles (rickshaw, minibus, motorcycle etc.) reduces the intersection capacity and should be considered particularly.

In order to determine the capacity utilization in a developing large city, a study was undertaken, in several districts of Istanbul, with the following objectives:

- To determine the base saturation flow rate in ideal or nearly ideal intersections
- To determine the existing saturation flow rate in several districts according to several significant factors.
- To evaluate the saturation flow rate and capacity utilization in comparison with the developed cities.

It is recommended to measure the saturation flow rate directly in the field [Akcelik, 1982]. Generally, there are two methods for direct measurement of saturation flow:

- Cross-section method
- Headway method

In the cross-section method, the green period is divided into three intervals and the saturation flow rate is measured at the middle interval. The first and last intervals are considered as lost time. In the headway method, time headway times between discharging vehicles are measured [Axhausen et al., 1989]. In both methods, first discharge or saturated headway is measured during the effective green time and then the saturation flow rate is calculated using:

$$q_s = \frac{3600}{t_B} \qquad (7)$$

where:

q_s = saturation flow rate (veh/h/ln)

t_B = discharge rate (sec/veh)

In this study, the headway method is used, so that vehicle position related factors could also be determined, since during the measurement vehicle types are also registered. Hence, beside the measurement of saturation flow the impact of vehicle types can also be determined.

4.2.2 Measurements on saturation flow in Istanbul

Istanbul is the largest city of Turkey with its population over 13 Million. The municipality operates 1747 signalized intersections mostly in time of day mode. Most of the intersections are on hilly terrains due to land structure. Pedestrian activities near intersections, parking maneuver of cars and loading/unloading of commercial vehicles are some general problems that reduce the capacity at intersections. Additionally, Istanbul has a special public transport mode: minibus. The minibuses do not have any regular stops and therefore stop also directly at intersections.

To determine the ideal value and its variation according to local factors such as geometrical standards and driving behavior, saturation flow measurements were conducted in the field. Therefore, 19 intersections from several districts of Istanbul are chosen (see Figure 26). Since Istanbul is a megacity with over 13 million inhabitants, Istanbul has several districts that reflect different intersection standards and also different behaviors. To determine these impacts the intersections are grouped into three categories according to Akcelik [1982] and, Webster and Cobbe [1966]:

- ***Group A:*** ideal or nearly ideal geometric and physical condition, good visibility, very few pedestrians, no interference due to loading and unloading of goods vehicles or parking turnover
- ***Group B:*** average condition, adequate intersection geometry, few numbers of pedestrians, some interference due to loading and unloading of goods vehicles or parking turnover
- ***Group C:*** poor conditions, large number of pedestrians, poo visibility, interference due to loading and unloading of goods vehicles or parking turnover

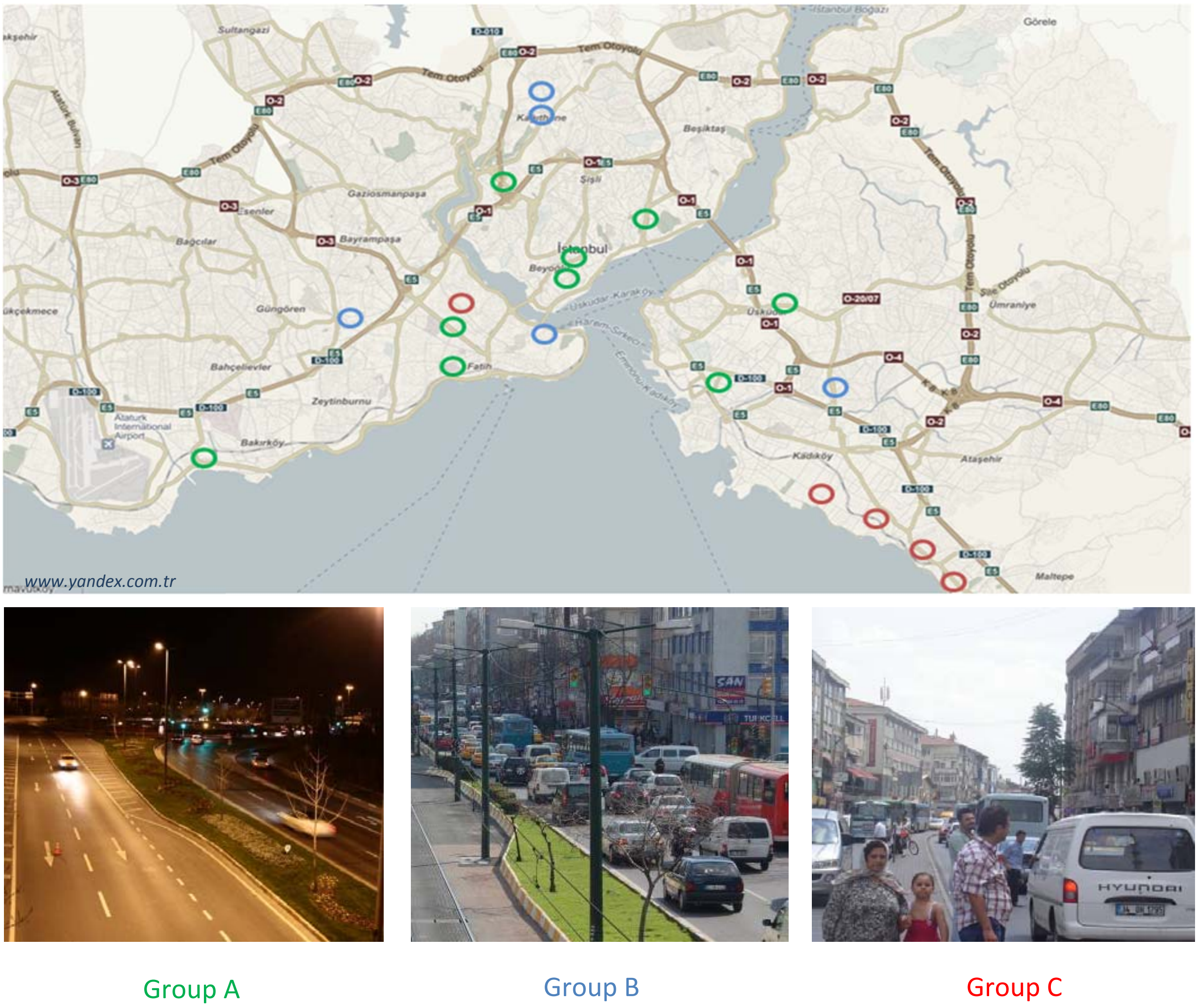


Figure 26: Measurement points and groups with different geometric standards (Group A, B, C)

4.2.3 Results of saturation flow measurements

- ***Ideal Saturation Flow Rate***

The Group A represents ideal or nearly ideal geometric and traffic conditions. As listed in Table 8 and 9 as an example from Groups A, the discharge time of vehicles are close to the base saturation flow defined in guidelines. At this intersection the saturation flow was about 1900 veh/h, which is still lower than in recent study in Germany, where the value was about 2000 [Schnabel et al., 2005]. However, the required discharge time for the first vehicle (1.27 sec) is smaller than in the study in German study (1.62 sec) [Schnabel et al., 2005]. This can derive from aggressive driving. Figure 27 represents the boxplot chart of intersection 15 (Ayrilik Cesme) that represents the ideal or nearly ideal conditions.

Table 8: discharge time considering single vehicles at intersection nr.15

Vehicle Position	1	2	3	4	5	6	7	8	9	10	11	12	13
Av. discharge time (sec)	1,27	2,51	2,41	2,22	2,08	2,07	1,92	2,14	1,82	1,83	1,76	1,74	1,52
Min. discharge time (sec)	0,40	0,50	1,40	1,30	1,30	1,40	1,20	1,00	1,00	1,10	0,90	1,10	1,00
Max. discharge time (sec)	6,20	3,50	4,80	5,50	4,00	3,40	4,20	6,40	3,90	3,50	2,90	2,60	2,10
Standard Deviation (sec)	1,05	0,63	0,76	0,76	0,61	0,47	0,58	1,02	0,58	0,54	0,50	0,41	0,43

Table 9: totalized discharge time at intersection nr.15

Vehicle Position	1	2	3	4	5	6	7	8	9	10	11	12	13
Av. discharge time (sec)	1,27	3,78	6,19	8,42	10,49	12,56	14,48	16,62	17,84	19,84	21,15	22,31	22,46
Min. discharge time (sec)	0,40	1,40	2,80	4,70	6,80	8,50	10,60	13,30	14,90	16,80	18,70	20,20	21,20
Max. discharge time (sec)	6,20	9,00	11,70	13,10	14,80	16,80	18,30	20,20	21,60	23,00	23,80	24,80	23,60
Standard Deviation (sec)	1,05	1,35	1,57	1,60	1,71	1,75	1,70	1,76	1,84	1,72	1,50	1,42	0,85

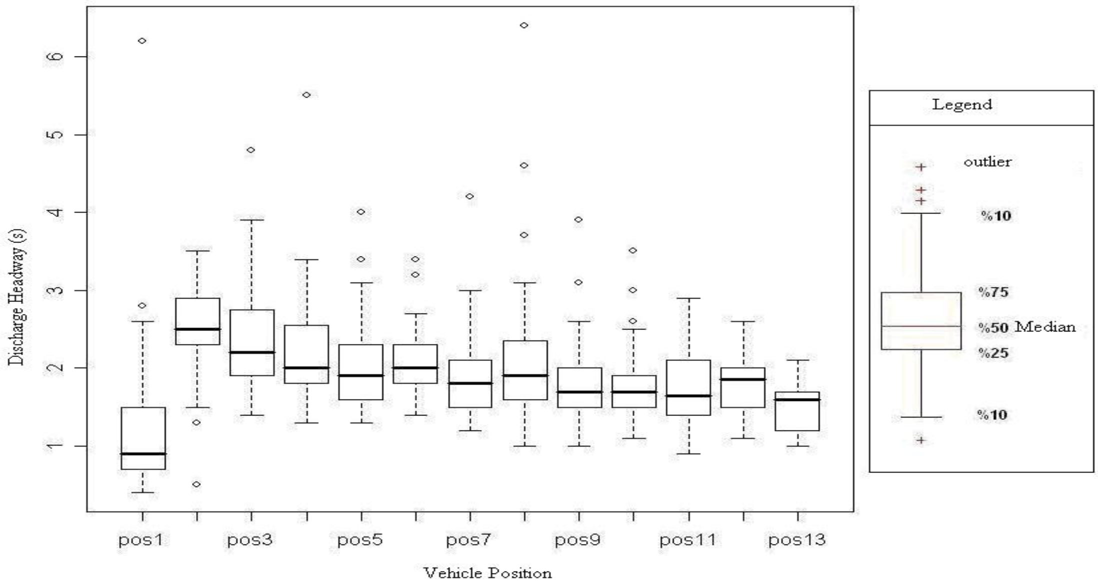


Figure 27: Discharge time at intersection 15 of group A

- **Saturation flow rate in the three groups**

The average discharge time at each intersection and the average saturation flow of each group are listed in Table 10. As the results show, intersections in Istanbul have different traffic characteristics and so different capacities. As in the group definition mentioned, the location of the intersection and the pedestrian activity affects the results. A specific factor for this metropolitan city was the “minibus”, in which nonobservant minibus drivers also stop during green intervals and reduce the capacity of that approach. Therefore during the categorization, intersections are also grouped according to the number of minibus lanes. Figure 28 represents required green time for 13 vehicles in different area types. For example, at intersection 15 (Ayrilik Cesme) 23 seconds are required for 13 vehicles, while at intersection 17 (Maltepe Camii) 43 seconds are required for the same number of vehicles.

Table 10: Discharge time and saturation flow in three groups

Group A	Intersection	2	15	7	3	14	9	12	Average Saturation Flow
	discharge time (s)	1.89	1.91	1.97	2.27	1.94	2.03	2.08	1827 veh/h
Group B	Intersection	13	1	8	4	5	6	11	
	discharge time (s)	2.11	2.02	1.98	2.09	2.01	2.2	2.44	1698 veh/h
Group C	Intersection	18	10	17	20	19			
	discharge time (s)	2.64	2.75	2.61	2.76	2.54			1353 veh/h

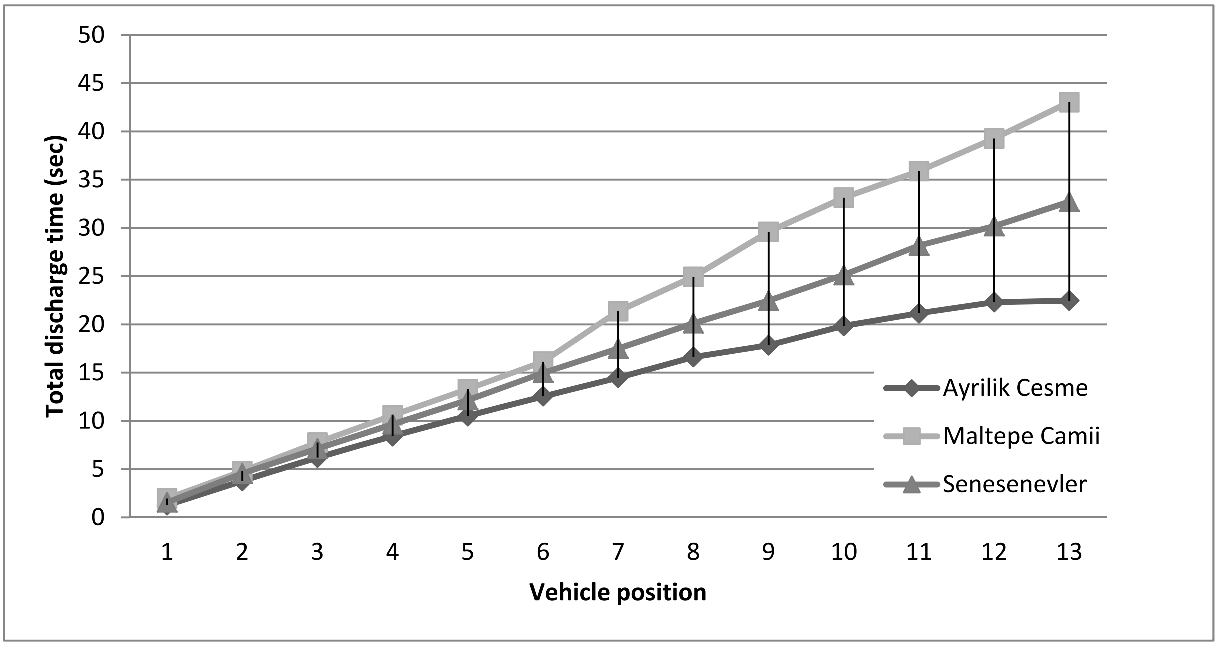


Figure 28: Required green time in different groups with different number of minibus lanes

4.2.4 Conclusions on capacity reduction

The following conclusions can be drawn from this study:

- In comparison with other studies in developed countries, the ideal saturation flow rate is not low. However, this value decreases swiftly with other factors.
- As expected, the results indicate that the saturation flow rate varies according to environmental conditions and is generally overestimated.
- The predetermined group categories represents relative homogenous saturation flow rate
- The average difference in capacity between the Group A and Group C is about 25%
- The specific factor “minibus” reduces the capacity about 50% particularly at intersections that are attractive for the pedestrian activities.

4.3 Traffic control related problems in developing large cities

Traffic congestion, emissions and traffic safety are some general traffic related problems of all large cities. However, in developing cities, some problems are greater than that of developed cities and some other problems emerge, such as the behavior of traffic participants, the lack of intersection layout or poor signal control. Meanwhile some other issues like environmental problems (traffic noise, exhaust gas emissions etc.) are not on the agenda of most cities of developing countries.

Considering traffic control related problems and the related solutions, the developing and emerging countries cannot be considered in the same category according to their national income as well as their traffic characteristics. For example in some developing cities, motorcyclists come at the forefront whereas in some other it is the cyclist. But in some developing cities motorcyclists and cyclists create a minority. Therefore local traffic conditions in each country should be considered separately.

Traffic accidents occur frequently not only in the developing world, but also in the developed world. Generally, population is used to compare traffic accidents. If the number of traffic accidents per population is compared developed cities can be more risky than developing cities. However, the number of passenger cars or registered vehicles could be the correct parameter for those comparisons.

This study focuses largely on general developing large cities, but does not concentrate on special problems of motorcycle (or bicycle) dependent cities, which are already investigated in depth by Khuat [2006] and Cuong [2009]. In this study, only signal control related problems are investigated and summarized in the following sections, not the whole transportation problems of those countries. Traffic signal control related problems of developing large cities and interactions are illustrated in Figure 29.

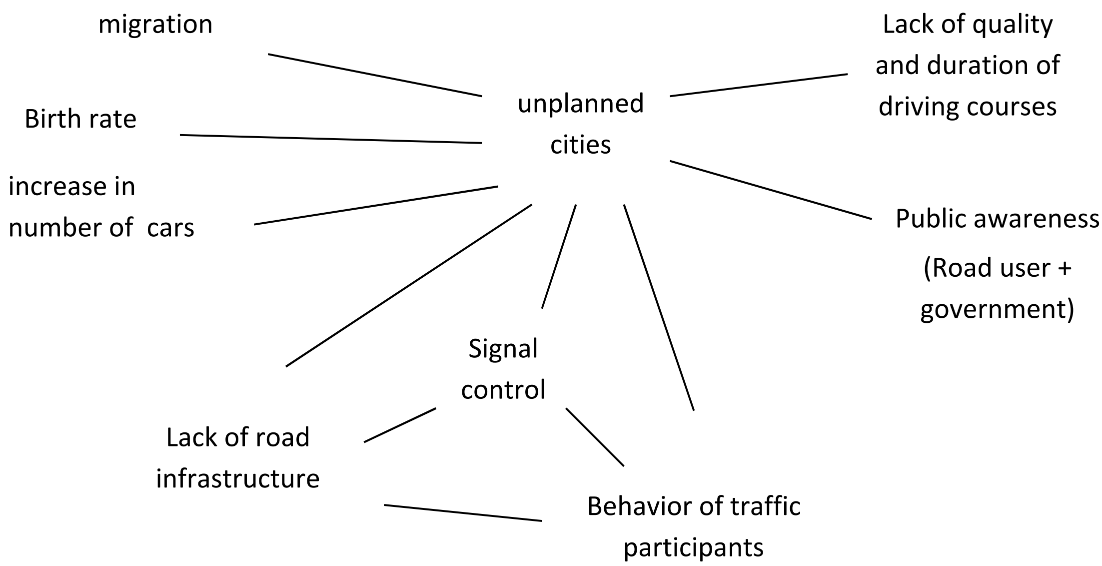


Figure 29: Traffic related problems

The increasing population due to birth rates and migration, as well as an unplanned city structure are problems inherited from previous politics. Lack of traffic education, public awareness and the employment of traffic engineers are partially inherited problems of those countries. As sub-problems the lack of intersection infrastructure, behavior of road users and poor signal control can be listed. As mentioned above, these problems are not completely independent. A significant mutual interaction exists between each other. The problems according to their reasons and scope of required measurement are listed in Table 11.

Table 11: Problems according to their roots and scope of required measurements

	Root of Problem		**Scope of required measurement**	
Problem Type	**inherited**	**generated**	**local**	**Regional**
Unplanned city	xxx	x	x	
Increasing population (birth, migration)	xx	x	x	xxx
Increasing number of cars	x	xxx	x	x
Education	xxx	xx	x	xxx
Public awareness	xxx	xx	x	xxx
Lack of road infrastructure	x	xxx	xxx	x
Behavior of road users	xx	xx	xx	xx
Signal control related problems	x	xxx	xxx	x
x minor impact	xx medium impact		xxx major impact	

4.3.1 Unplanned cities

The cities can be considered in three groups: car oriented cities, motorcycle dependent cities, and transit oriented cities (low-cost bus oriented, modern rail oriented). Due to rapid motorization the developing cities should consider several scenarios; either a modern car oriented city with a weak city center or a rail oriented city with a strong city center [Khuat, 2006].

Some researchers consider that in unplanned cities, narrow roads and high densities are also a cause of traffic problems. This argument is then correct, if only the car dependent cities are considered. However, on the other hand, high-density cities are considered as more sustainable than the other cities due to length of the daily trips. The type of city and the major transportation mode are significant to correlate the city structure with traffic problems. It means that the correct transportation mode for the existing city structure should be chosen.

Inadequate road networks can be considered as one of the reasons for traffic problems but not all. As seen in Figure 29, developing larger cities have only 6 to 12 percent road density, whereas developed larger cities have between 20 and 25 percent [Vasconcellos, 2001]. However, if the road density is discussed as a problem of developing larger cities (i.e. by Morichi [2005], the ratio of car ownership have to be considered for comparisons, since the ratio of car ownership in developing countries is far lower than the ratio of road densities.

As a conclusion, an 'unplanned' city will only be problem if the city structure is not considered during the planning of transport infrastructure. But, this is a common situation in developing cities.

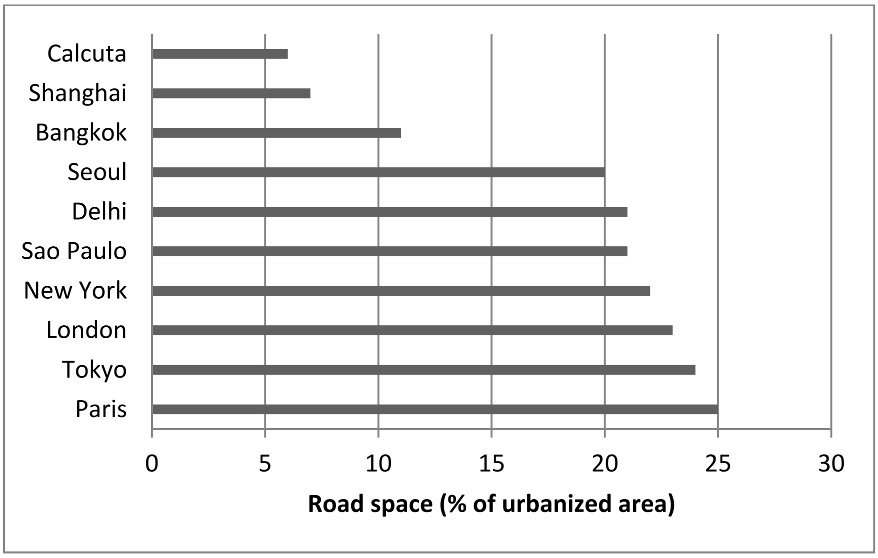


Figure 30: Road space in some larger space [source: Vasconcellos, 2001]

4.3.2 Education and Public Awareness about transportation

In developing countries due to less number of car ownership the public discussions about traffic problems and solution seeking are started after a population threshold, more accurately, after a ratio of car ownership or traffic demand.

Traffic engineering education and traffic engineer employment in the municipalities, and the duration and quality of driving courses are also considered under this section. The issue of public awareness then has two sides; that of the road user and the government. Of course, besides financial deficit, the lack of public awareness about transportation plays the main role in the traffic control related problems. If awareness about transportation increases, some traffic problems can also be solved with small budgets. Unfortunately, the lack of awareness of road users and transportation authorities triggers some other transportation problems. Generally, obtaining a driving license is easier in developing countries than in developed one. The duration of courses and examinations are shorter and compulsory attendance in the lectures is generally not required or tolerated. Of course, this deteriorates the driving behavior in those countries.

4.3.3 Lack of road and Intersection Infrastructure

The problem is not only a lack of guidelines or directly translation problem from developed countries as Cuong [2009] mentioned, but also implementation of regulations defined in guidelines. Beside the differences between recommendations in guidelines (or standards) and intersection layout, there are also some differences between intersection layout and the intersection itself in the field; for example a lack of stop bar marking, lane marking or marking for pedestrian etc. This occurs either because of a lack of maintenance operations or due to taking it incompletely into operation.

4.3.4 Behaviors of road users

Generally, in developing countries, car users are not trained sufficiently to follow lane discipline, and as a consequence of poor lane discipline, especially at intersections, the already overcrowded intersection situation deteriorates. [Jain et al., 2012] Furthermore, red light running in free traffic causes accidents and thus traffic congestion. Also, in peak hours causes further traffic congestion due to intersection blocking [Karadag et al., 2012].

Several researches proved that enforcement systems could reduce penalties significantly and change traffic behavior. By this means, at least the intersections or emergency lanes can be kept free. [Karadag et al., 2012; Hess and Polak, 2003; Christie and Lyons 2003]

Intersection geometry and lane markers also have an effect on the behavior of traffic participants. For example, the geometry itself of signalized roundabout in some cities deteriorates the existing poor lane discipline. Figure 31 and 32 represents an intersection type with a regular left turn lane and a roundabout with left turn using the roundabout. Signalized roundabouts have some advantages (i.e. reducing lost time) and are applicable by low left turn volumes. However, generally in larger cities left turn demand is also high at most intersections. In this example, the vehicles expand to three or more lanes, after the intersection is reduced again to one lane. First, this intersection type create bottleneck itself, second, it reduces the lane discipline of the drivers.

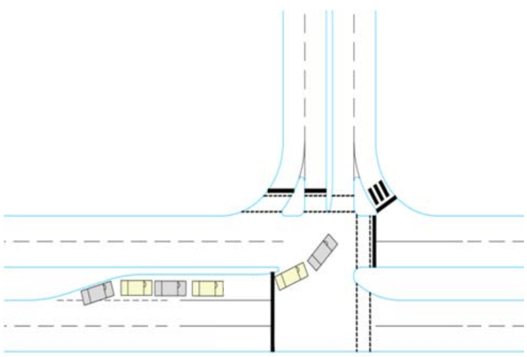


Figure 31: left turn movement at signalized intersection

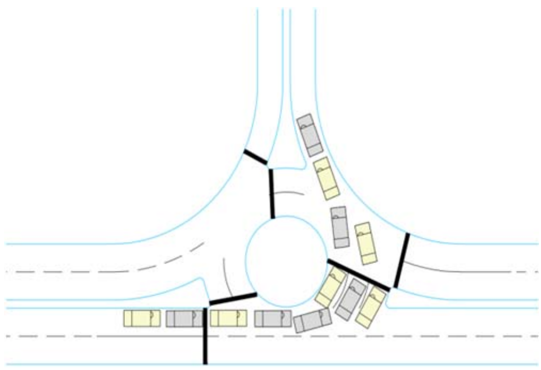


Figure 32: left turn movement at signalized roundabout

Nonobservant road users are not only the car drivers. Pedestrians as jaywalker also affect the capacities of intersections and risk their own life. On the other hand, car drivers do not pay attention at the priorities of pedestrians. If the right turn vehicle movement has the same phase with pedestrians, mostly the car drivers do not pay attention to the pedestrian. Therefore, unfortunately, transportation authorities in those cities avoid creating the same phase for the right turning vehicles and crossing pedestrians. Thus, an effective operational measure that reduces the lost time must be avoided. Zebra crossings for pedestrians are also ineffective measures to bring the pedestrians over to other side.

4.3.5 Problems of signal control

In most developing cities, traffic signals are placed only to regulate the traffic streams. In fact, traffic signals can do more than traffic regulation and traffic safety. Since the capability or actual capacity of traffic signals are not known, interchanges (multi-level junctions) are constructed to reduce traffic congestion with the increasing traffic demand.

As mentioned earlier, traffic signal control in most developing cities comes to discussion only if their population and/or car ownership ratio surpasses a threshold. One can say that only larger cities or megacities discuss the current signal control infrastructure and seek new alternatives to improve.

Generally, signal coordination and the maximum green band for main approaches are privileged on arterial streets. However, in a peak traffic situation left turn movements block the main approaches and the green wave coordination is interrupted. Even, in some intersections, a separate left turn phase is prepared with a very narrow left turn lane, to provide more green time for main approaches and to maximize the green band. However, since the left turn traffic blocks the main road, both the intersection capacity reduces and the coordination is interrupted.

Furthermore, in some motorcycle dependent cities by design and operation of signalized junctions only car traffic is considered. Even, some countries do not have any guidelines or standards for intersection design [Cuong, 2009]. In some countries, only one cycle time with two phases is used for the whole day and in some intersections pedestrian movements or clearing time for pedestrians is not considered [Cuong 2009].

Some signal regulations like right turn with pedestrian or permitted left turn phases that reduces lost time cannot be used in some countries due to the acceptance of drivers and traffic safety. In some cities (i.e. some districts of Istanbul) nonobservant minibus drivers wait also during the green interval to pick up passengers, and so reduce the capacity of that lane and interrupts the coordination of the arterial.

Furthermore, the usage of detectors that allow more flexibility in signal programming is too low in developing cities. Due to uncoordinated frequent road works of different agencies like phone, electricity or gas companies, the detectors cannot be used properly, which sends the traffic agencies away from the traffic-actuated control.

4.4 Conclusion

Some traffic problems must be solved with more comprehensive political regulations, like internal migration, urbanization policy, city planning and to choose the main transport mode of the cities. Education and Public awareness about transportation issues can also be listed in comprehensive midterm political decisions. Especially, the training of the engineers and the employment of traffic engineers will help overcoming some sub-problems, like intersection design or signal programming.

The problems directly related to intersection and traffic signals are interactive. For example poor intersection designs (with poor marking) can worsen the driving behavior or due to the driving behavior some signal regulations cannot be implemented.

The traffic control strategies for developing countries should also consist of measures for resolving these problems and improving traffic condition.

5 A signal control framework for developing large cities

The previous chapter presented two preliminary studies that represented the current situation in developing large cities, and described the traffic control related problems of those cities. This Chapter proposes a framework for developing large cities. Firstly, the objectives and the required actions are described with consideration of the strategic goals of an urban transport system. Secondly, multi-level signal control architecture is proposed for the signal control system of developing large cities.

5.1 Traffic management for developing cities

A signal control system is not an isolated system that works without an external impact. There are lots of factors, which influences (generally negative) the traffic flow at intersections and consequently the quality or level of service (LOS). Therefore first some general traffic control related problems should be resolved. For this purpose, both short- and long-term strategies must be developed, and actions must be taken to achieve these objectives.

More or less some strategic goals of urban transport system are defined in several guidelines. These goals generally consist of sustainability, environmental aspects, traffic safety, the economy, the accessibility of regions and to save natural resources [Retzko and Boltze, 1993; Khuat, 2006]. Based on these strategic goals of urban development, traffic management includes the influence of traffic supply and demand, and consists of political and operational strategies to avoid traffic (reducing traffic volume, reducing distances etc.), to shift not avoidable traffic (temporal, local and modal) and to control the rest of the traffic efficiently and making it environmentally compatible (see Figure 33) [Boltze, 2005].

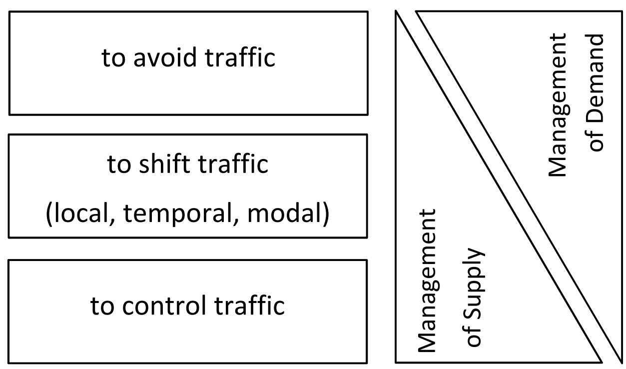


Figure 33: impacts of traffic management [source: based on Boltze 2005]

As described in Chapter 4, the cities in the developing countries have some unique problems: such as driving licensing procedure, general education and/or public awareness about transportation (i.e. car dependency), the behaviors of road users (not only drivers but also pedestrians and cyclist), the lack of road infrastructure design and equipment. A general strategic plan for a signal control framework for developing cities should also contain strategies as a prerequisite to resolve aforementioned problems. Therefore, the following short- and long-term strategies should be implemented:

- ***To provide well-balanced regional development***

The first and inherited traffic control related problem of large cities is uncontrolled population increase and urban planning. A part of population comes from internal migration, because of unbalanced regional development (job, education and infrastructure opportunities centered on a few cities). Additionally, the cities themselves are not planned in a galaxy system structure that considers centers, sub-centers etc., but unplanned like an oil stain. This strategy focuses on a balance in urban and regional development and increasing the attractiveness of other cities and regions. Thus, maybe the population will not reduce but its increase will decelerate. The cities themselves should also be planned according to theorem of a "city of short distances" to reduce the total distance per daily activity in the city.

- ***To reduce vehicle use and to decelerate car ownership***

As presented in the comparative study in the previous chapter, the current car ownership in developing countries is very low in comparison with developed cities. However, the situation is dynamic; it means that car ownership increases rapidly with an increasing economic growth. The governments invest always in more roads for automobiles to resolve or to satisfy the traffic demand. However, this strategy creates more automobile dependent traffic demand. Therefore a new strategy to slow down car ownership and to reduce car usage for daily trips is required particularly for large cities.

- ***To improve quality and duration of education, and training for driving licenses***

Generally, the lack of driving courses (time, quality, exam difficulty) affects on the one hand traffic safety (untrained driver) and on the other hand it has an impact on number of violators. Therefore, this examination procedure should be upgraded and the duration of the courses should be extended with improvement in their content and quality.

- ***To deploy enforcement systems to reduce negative impact of nonobservant road users***

This strategy focuses on an improvement in traffic safety and traffic flow. Several studies proved that the use of electronic enforcement systems reduces the red light running at intersections or speed violations. Therefore this strategy can also help to improve traffic flow (i.e. prevent blocking the intersection due to vehicles that cannot be served during green interval) and reduces external impacts on the designed signal control system.

- ***To support public transport (and priority) project***

This strategy focuses on providing "mobility of people" but not the vehicle in urban areas. Therefore public transport should be supported both with priority for public transport construction projects and with public transport priority at intersections. Particularly, in developing cities, improvement quality of public transport vehicles (by increasing the number of vehicles and so to provide comfortable travelling, and by increasing the operating speed) will change the modal-split towards public transport due to their economic condition. However, this may require a change in the traffic signal control system.

- ***To improve geometric standards of roads and intersections***

The following step is a proper design and construction of road and intersections according to guidelines. This strategy includes improvements on geometric standards, lane marking, and the determination of a correct stop point at intersections. Intersection geometries that may affect driving behavior should also be reconstructed (described in Chapter 4). This strategy also includes proper segregation of traffic modes on the roads particularly in those cities, where non-motorized vehicles comes at the forefront.

- ***To increase the employment of transportation engineers in municipalities***

This strategy is also a prerequisite of road projects, intersection design and construction in view of demand oriented traffic management.

- ***To establish a strategic management plan for maintenance and control of traffic signals***

The advantage of each form of real-time control method against to fixed time control method is proven in several comparative studies. A real-time control can require a change in controller type and proper maintenance of infrastructure (i.e. detectors, communication system). Therefore a concept for scheduled change of an existing system and a strategic plan for maintenance coordinated with other departments is required.

- ***To prepare guidelines for signal control included real-time control strategies***

Generally, developing cities either do not consider any guidelines for intersection design and signal control or directly translate the guidelines from developed countries. However, used parameters are generally unsuitable for special condition in these countries (i.e. mixed traffic conditions are not considered). Therefore these cities require their own guidelines or current guidelines should be well adapted to current conditions. For example Cuong [2009] adapted the German Guidelines for signal control to Vietnamese traffic conditions.

5.2 Multi-Level Architecture

The signal control management system proposed in this study has a two-level architecture and has bi-directional communication with the transportation management center. The Transportation management center should also be connected with other control parts (i.e. freeway management, parking management etc.). As depicted in Figure 34, a simplified signal control management unit consists of two levels; the upper level act as the planning level and the lower level as the operational level. This would control the signalized intersections of each sub-region at a pre-determined time interval (5 min. to 15 min.).

The main idea is a relative simple structure with reduced detection and communication. One can say that the upper level is relatively static, while the lower level more dynamic. Traffic signal control related data could be directly sent to signal control management unit or sub-regions. The upper level also contains manual counts and general transportation modeling data.

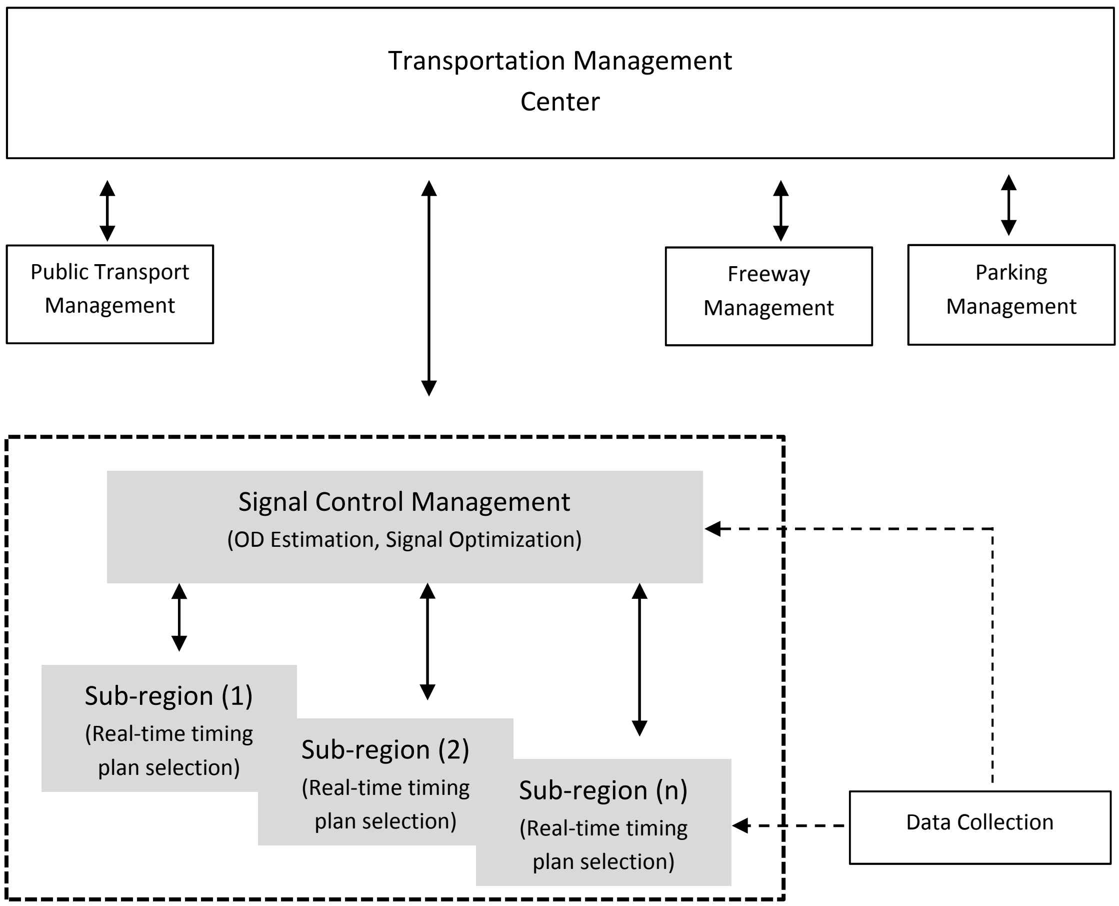


Figure 34: Multi-layer traffic management concept for developing large cities

5.2.1 Planning Level (Macroscopic Level)

On the upper level, strategic and political criterions are determined for the strategic control of traffic signals. The objective function for optimization and related parameters are defined and signal control related data is stored in this level. As Schlaich and Haupt [2012] mentioned, traffic counts are the most obvious and most used data for signal control and optimization. However, in particular, developing large cities have some limitations using traffic counts. Also, origin-destination matrices for transportation planning are either not available or aged. This is not surprising for those cities, where the population increases rapidly. Therefore, an origin-destination matrix estimation technique (i.e. Willumsen [1978], van Zuylen [1979] or Ploss [1993]) could help firstly; for the estimation of the OD demands, secondly; for the estimation of the traffic demand on links or intersections, where actual traffic counts are not available.

The tool that will be used at the upper level of a signal control management unit should provide the origin-destination matrix estimation, traffic signal optimization and scenario management. The OD matrix estimation is needed for the calculation of the demand on each link or intersection and particularly for scenario management. The scenario management can be used to prepare several timing plans in consideration of a change in the traffic demand and in the network. Generally, the traffic demand varies daily and seasonally. The primary timing plans are based on daily variations on traffic demand and the secondary timing plans consider an event-based change in traffic demand (i.e. weekend, leisure traffic or big events etc.). The changes in the supply part are at the partial or entirely closing of road sections (generally short-term) or extension in the road network (generally long-term). Or in some cases both the supply and demand site changes for a short-time period (closing road section for demonstration). During the scenario management origin-destination matrices are estimated and so the link capacities are determined again and the traffic signals are optimized.

The last task that should be accomplished on this level is the traffic signal optimization. During the optimization, the interaction between signal optimization and traffic assignment should be considered, since the signal optimization influences the route choice, which then changes the flows and turning movements and thus the basis for the signal optimization (see Figure 35). Therefore the traffic signal optimization framework can be extended with an outer iteration that considers traffic assignment with a junction model and uses a node impedance calculation method [Schlaich and Haupt 2012].

If required (particularly for large cities) the city network can be divided into (n) sub-regions according to the interrelation of intersections or arterials. Traffic signals in each generated sub-region can be optimized under consideration of interactions between sub-regions.

The prepared timing plans included eventual scenarios that are sent to the operation level. The communication interval between levels can be fixed in advance or new plans can be sent only depending on a change in demand or in network.

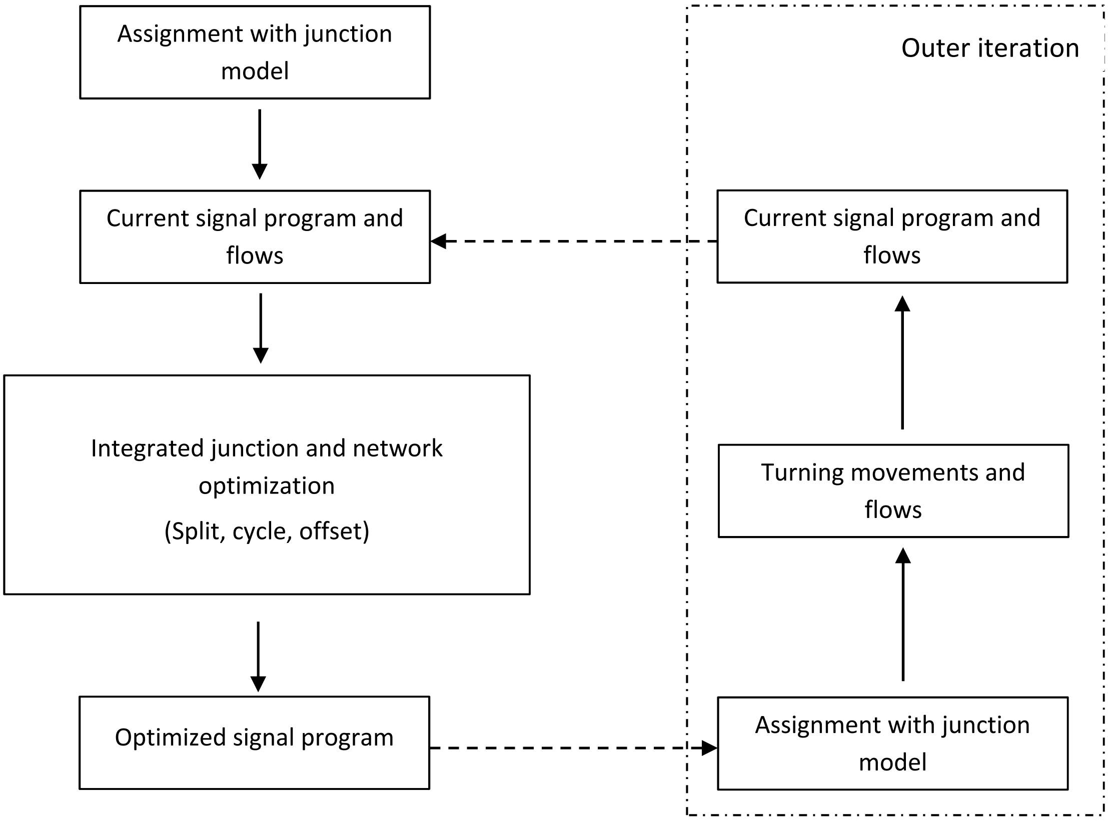


Figure 35: Optimization of fixed timing plans with outer iteration loop [source: adapted from Schlaich and Haupt 2012]

5.2.2 Operation Level (Microscopic Level)

The basic function of the lower level is the implementation of a traffic responsive plan selection algorithm, which will be described in the next chapter in depth. The optimized timing plans are received from the upper level and selected using a pattern recognition algorithm according to the information collected from detectors. Pattern recognition provides a competitive advantage by using a fewer number of detectors. In brief, traffic counts and occupancy data on a predetermined sub-network is collected and sent to the neural network controller. The controller recognizes the predefined traffic state and selects a suitable timing plan for the current situation. The collected information is sent to the upper level. Additionally, the system will also send a signal if an undefined situation occurs. In this case, the sub-region is optimized at the upper level and new timing plans are prepared. The following chapter describes the pattern recognition method using the artificial neural networks in detail.

6 Pattern recognition approach for simplified traffic responsive signal control

6.1 Introduction

As described in the previous chapter, the lower level of the proposed system is responsible for the selection of suitable timing plans for the actual situation. This can effectively be described as a pattern recognition problem. Considering daily, weekly or seasonal traffic in a predetermined network, this in turn represents either primarily pattern or some secondary “event” patterns. The simple task of this algorithm is first to select the presented feature and second to recognize the traffic state. Additionally, if a new pattern comes up, a signal will be sent to the user, thereby the user can start the optimization and training again. Eventual primary and secondary predictable and unpredictable traffic pattern on an arterial or network are listed in Table 12. Daily traffic patterns (or clusters) can be increased while bearing in mind the capacity of the signal control equipment. To determine these traffic patterns a cluster analysis can be performed. To simplify and to bring the engineering judgment to the forefront, instead of a cluster analysis the known clusters and timing plans of agencies can be used. In the next chapter, a cluster analysis is performed to examine the current traffic state clusters and to compare it with existing classification.

Table 12: primary and secondary traffic pattern on arterial or networks

Primary traffic patterns	Secondary traffic patterns
morning peak hours	big events (sports, cultural, etc.)
afternoon peak hours	road incidents, constructions
off-peak hours	fairs, street markets
nights	weekend and leisure traffic

The proposed system works similarly to a rule-based traffic responsive system. However, both systems differ by identifying current traffic conditions. The pattern recognition method is used for estimating the prevailing traffic state instead of logical queries as used by rule based traffic control. In a pattern recognition method each pattern has some features and is recognized by these features. The local traffic state that is determined directly from detectors is used to estimate the regional traffic state on a selected arterial or network using the pattern recognition process (see Figure 36). This can also be considered as the ‘local’ and ‘central’ level as described in [von der Ruhren 2006]. Von der Ruhren [2006] classifies the detected traffic data at the local level and the general traffic state in that region is determined at the central level. The classification at the local level is based on actual detected traffic data (i.e. traffic counts, occupancy or speed). The measured traffic data generates the daily flow profile with the data collected at other intervals. The results of the local level are analyzed at the central level and the overall traffic state is estimated [von der Ruhren, 2006]. The advantage of this multi-level classification method is that on-site classified traffic states can also be used in traffic information center for local user information. In a similar way, in the coordinated traffic signal control, each detector represents a demand state for a predetermined time interval. As depicted in Figure 35, these detectors can represent different demand states for the same time interval. The problem is to determine (or to estimate) the overall traffic state on the network.

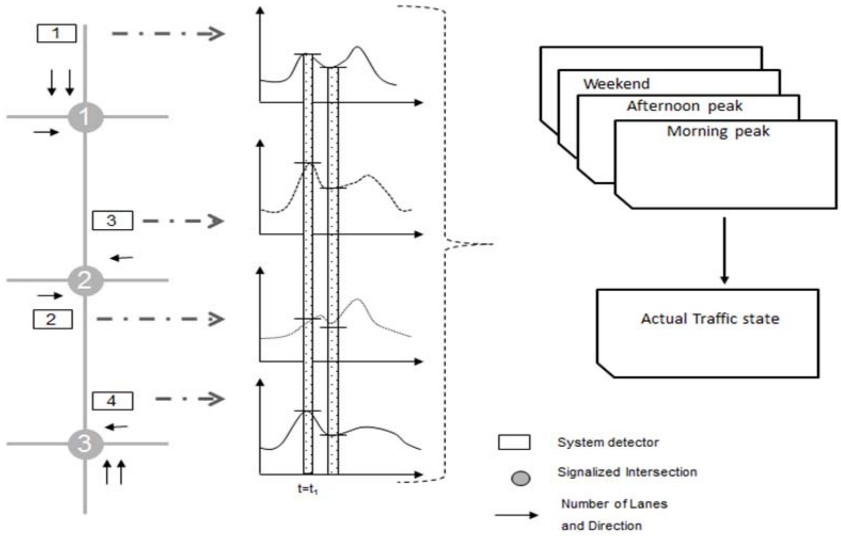


Figure 36: Traffic state estimation from system detectors

In the traffic signal control, traffic volume (*q*) and occupancy (*k*) can be used as the features of the traffic patterns. The detected traffic data at each detection point $f(q,k)_j^t$ constitutes a scenario after a classification process. The following equations present the relationship between detected traffic situation, network traffic state (or scenario) and the timing plans.

$$S^t = \begin{bmatrix} S_1^t \\ \dots \\ S_i^t \\ S_n^t \end{bmatrix} = \begin{bmatrix} f(q,k)_1^t & f(q,k)_j^t & \dots & f(q,k)_k^t \\ \dots & \dots & \dots & \\ f(q,k)_1^t & f(q,k)_j^t & \dots & f(q,k)_k^t \\ f(q,k)_1^t & f(q,k)_n^t & \dots & f(q,k)_k^t \end{bmatrix} \quad (12)$$

$$S_i^t \rightarrow TP_{m,n}^t (TP_{m,1}^t, TP_{m,2}^t, \dots TP_{m,n}^t) \quad (13)$$

where

$f(q,k)_j^t$ = traffic count (*q*) and occupancy data (*k*) of detector *j* at time *t*

S_i^t = scenario (*i*) at time *t*

$TP_{m,n}^t$ = corresponding timing plan (m) for intersection (n) at time (*t*)

The planning level keeps all collected data together and prepares scenarios and corresponding optimized timing plans for each scenario, and is sent to the lower level, where the training of the system and the operation occurs. The neural network control applies pattern recognition method to recognize the traffic state. After that the corresponding timing plan is selected. Since, one of the main advantages of neural network is its ability to learn, the system will continuously learn with the new data that is collected during the operation. If a new traffic state appears, the user will be informed and new timing plans are prepared for the new traffic state. This occurs, either if the traffic state is not completely recognized or the error rate increases and exceeds a (second) threshold value. Since the system will continuously learn, the scenarios or traffic states can increase by the time. Thus the required timing plan will also increase. The Figure 37 shows the general flow diagram of this algorithm. The details of the learning and implementation process of the proposed methodology are explained in the following sections.

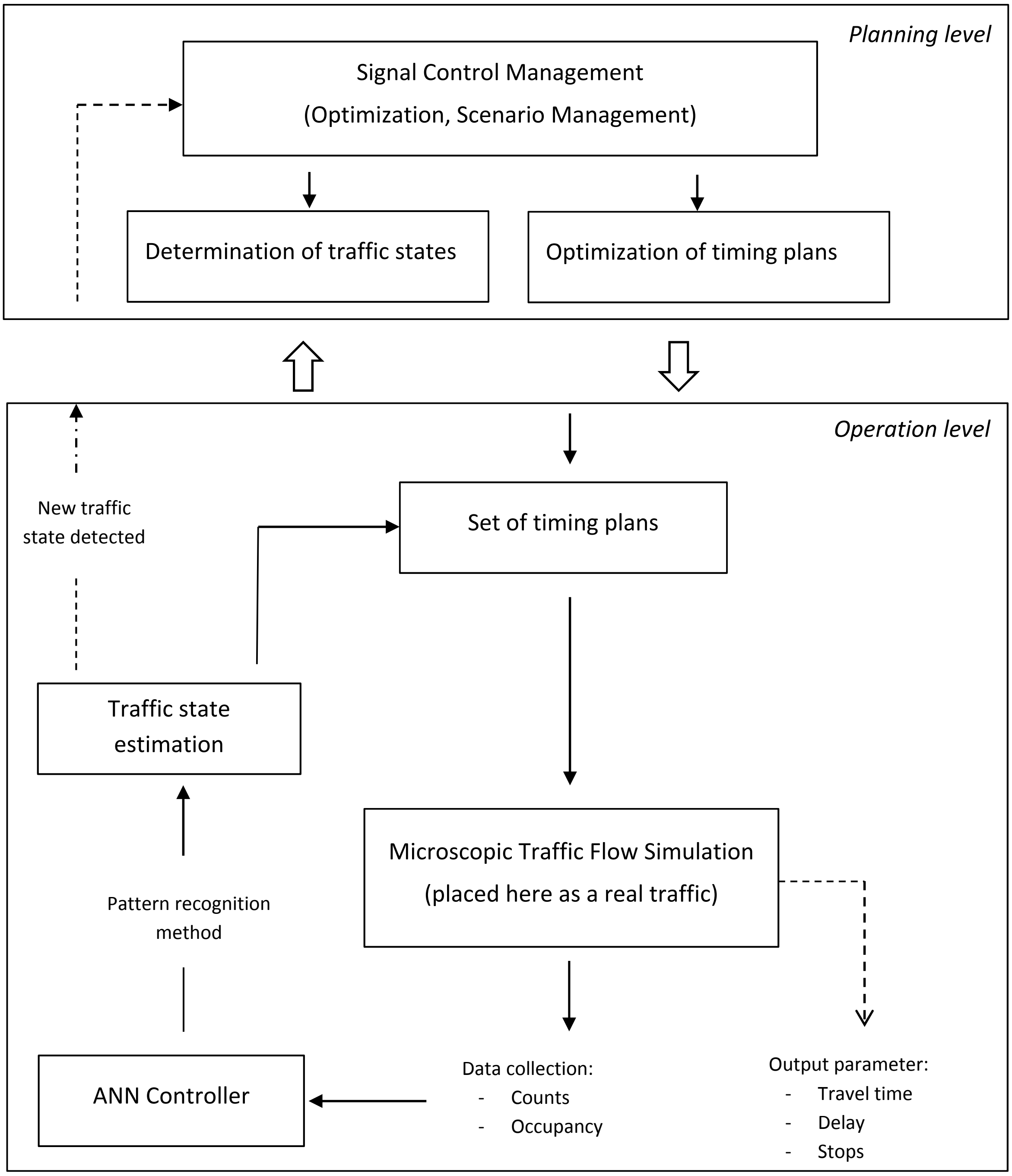


Figure 37: Structure of proposed traffic responsive plan selection model using pattern recognition

6.2 Implementation process of pattern recognition using neural networks

The implementation of pattern recognition based timing plans selection systems using neural networks is applied in eight steps, which represents a combination of a TRPS implementation process in [Balke et al., 1997] and a pattern recognition process in [Jain et al.,2000]. Here, the general processes for setting up a closed-loop traffic responsive control system and for pattern recognition are almost overlapping. For example both processes have row data processing, weighting, scaling etc. Figure 38 shows this process in general, while the remainder of the chapter describes it in detail.

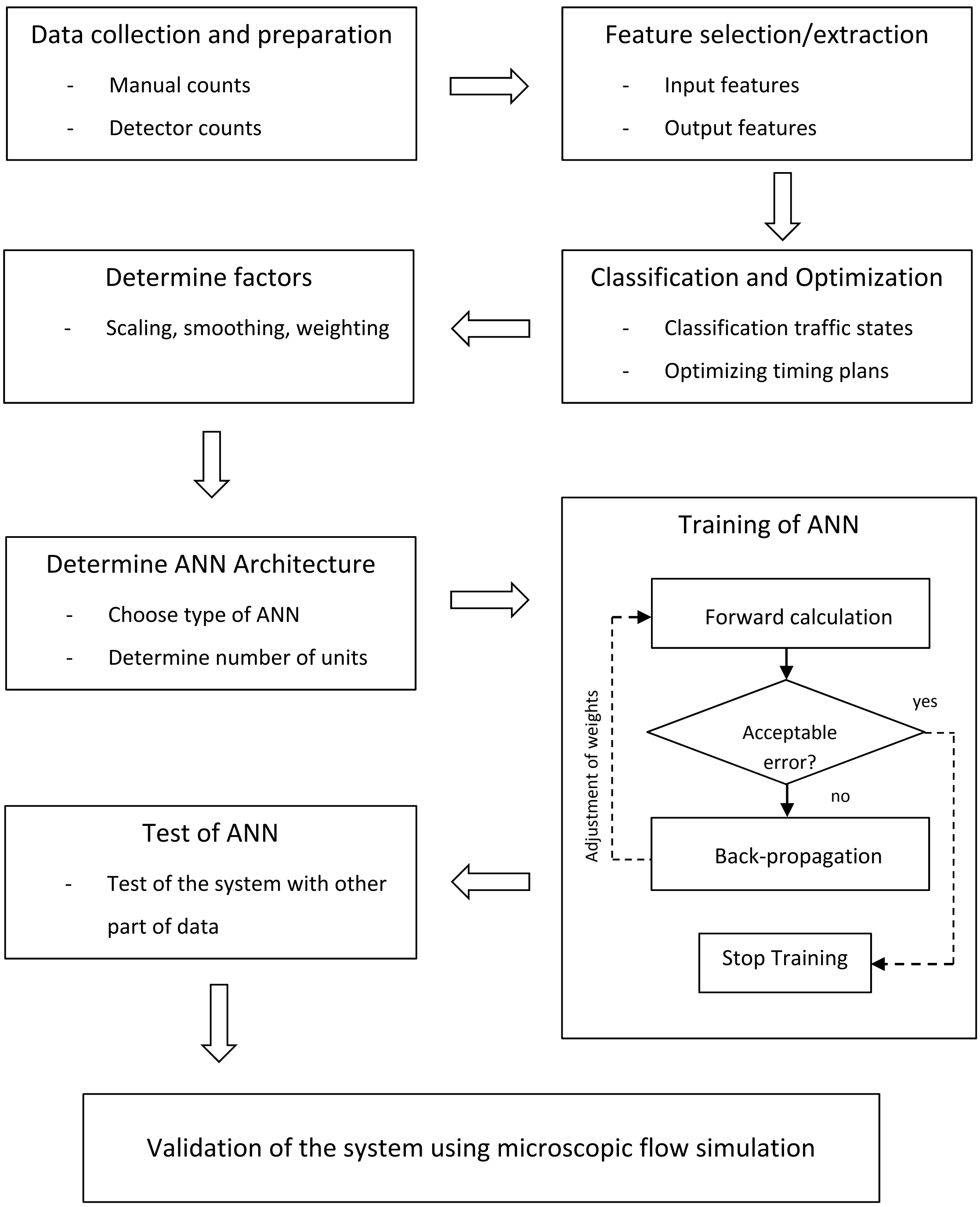


Figure 38: General process implementation of pattern recognition using ANN

6.2.1 Data collection and pre-processing

Principally, the first step of pattern recognition is the data collection and the pre-processing. Generally, the data is collected using sensors. This data requires a pre-processing that varies according to problem type. The pre-processing module segments the pattern of interest from the background, removes the noise, normalizes the pattern, and any does any other operations, which will contribute in defining a compact representation of the pattern [Jain et al., 2000]. In traffic flow, the role of the pre-processing module is the preparation of the raw data. This includes plausibility tests to identify the failure or gap in raw data. The failure or gap can come up during the data collection due to detector malfunctions or interruption of communication systems. This failure or gaps should be eliminated, corrected or completed using relevant methods [von der Ruhren, 2006].

In signal control, this step includes manual traffic counts, installation of system detectors and collection of data for the training of pattern recognition systems. Manual traffic counts are used for long-term planning and updating of the system in the control center. The proposed real-time system is so designed that the traffic state is recognized using system detectors. However, generally, closed-systems allow limited detectors to be assigned (i.e. Naztec up to 10 and Econolite only 4 system detectors) [Balke et al., 1997]. Therefore the placement of this limited number of detectors is significant, since the detectors are needed to estimate the traffic state at a predefined time interval.

Especially, in developing countries maintenance of detectors and keeping them in operation is difficult, due to uncoordinated road excavations by several agencies (post, electricity etc.) and direct road constructions. The collected data is also sent to the signal control management unit to be kept and to update the system.

Although in some literature (i.e. [Webb, 2002]), the feature selection/extraction step is considered under pre-processing, this step is separately described in this study.

6.2.2 Feature selection / extraction

The role of the module feature selection is the selection of correct feature. It is significant for pattern recognition problems and generally, complex part of pattern recognition. An incorrect selection of feature can cause totally incorrect classification [Fletling, 2010]. For example, the “color” or “brand” of a car are not correct feature in the emission modeling, but “horsepower” or “engine size” could be the correct classes.

Generally, two methods are used for feature extraction module [Niemann, 2003]:

- The heuristic method, based on intuition, imagination or experience
- The analytic method, based on systematically search for optimum feature

For the simple pattern problems, generally, experience of experts is used to find the suitable features, while for complex problem an analytic method should be used [Niemann, 2003].

Another role of this module is the transformation of pre-processed raw data to a feature vector (x) that enables a simple and robust classification. The dimension of the input vector has a significant role in pattern recognition. If it is too less, then the pattern contains less information to classify correctly. Contrary to this, if the input vectors have high dimensions, and so contain redundant and irrelevant information, a simple and quick classification can be difficult. The objective of the use of feature vectors is that the reduction of dimension of input features by elimination redundant and irrelevant information [von der Ruhren, 2006].

Fundamentally, the inputs of a real-time signal control system are the traffic counts and the occupancy data, with the output being the traffic state. For example, very common traffic states can be the morning peak and the afternoon peak hours.

The output feature can be increased according to the timing plan capacity of the controllers and regional computer. Today, more powerful signal controllers are available, but it is not the case for all cities in developing countries. Therefore, the number of output features can be selected according to regional circumstances. In this study only traffic counts and occupancy data are used as an input feature. However, increasing the number of input feature is also possible. For example, speed, queue length or delay (if a real-time measurement is possible) can also be used as input features.

6.2.3 Classification and Optimization

After the data collection and the determination of features, the control conditions should be determined. Pattern recognition in this study is applied in a supervised manner. It means that the system outputs are previously labeled and therefore, the traffic states should be previously classified. The determination of the traffic states is conducted using a classification method.

The classification describes a process or a method for the separation of objects in classes or categories. A class or category forms a set of objects which show as similar as possible or the same distinctness of their features. Object of different classes should have distinctness of their features as different as possible [Fletling, 2010].

The classification can be conducted in two ways [Fletling, 2010]:

- Expert methods
- Automatically classification methods (cluster analysis)

The expert knowledge can be used, generally, for simple problems (with fewer components) that are relative easy to cluster or the cluster alternatives are relative less. For complex problems, generally, several classification methods (also called cluster analysis) are used.

Many of the classification methods require similarity (or dissimilarity) measure or distance between two pattern vectors. Thus, in some cases, data can arise directly in the form of a dissimilarity matrix [Webb, 2002]. The more similar are the objects, the smaller the distances of the features [Fletling, 2010]. Many methods are available for the similarity (or dissimilarity) measures (i.e. Euclidean distance, city-block distance, Canberra distance etc.).

For the classification in arterial signal control the expert knowledge is also important. Then, the control conditions are determined not only based on collected data at one single point, but other parameters and conditions and transport political decisions should be considered. Additionally, the timing plan capacities of controllers can also be a factor as the constraint in classification.

Classification is done by using the manual counts on intersections and cross-sections, and the collected data from system detectors. The number of classes varies depending on the fluctuation in traffic volume and the capacity of the controller for corresponding timing plans. As listed in Table 8 (on page 70), generally, three to five clusters for weekday and several special clusters for leisure and event traffic could be adequate. This is also verified by the results of the preliminary comparative study, where most of cities responded that they generally use 3 to 6 weekday timing plans. Traditional TRPS systems calculate the total demand, directional demand and cross street demand, and select the corresponding cycle, split and offset times from a timing plan matrix [Balke et al. 1997].

After classification of traffic states the corresponding timing plans are generated and optimized. Since the optimization of traffic signals consider all links or approaches of intersections or networks, the total delay and number of stops can be reduced in comparison traditional coordination of traffic signals. To achieve the coordination a common cycle length is required along the entire (or selected) arterial or network. The optimization of traffic signals can be done either analytic methods or commercial signal optimization software can be used (i.e. TRANSYT, SYNCHRO, VISUM etc.)

In this study, cycle, split and offset times represented on the timing plan and could only be used together instead of a timing plan matrix, since timing plans (with all parameters) have been previously optimized. It could be an alternative to a timing plan matrix that consists of required green splits, cycle and offset time for several control conditions according to predetermined threshold values as in traditional traffic responsive systems.

It should also be noted that the older intersection controllers have a limited capacity (i.e. PLC controllers) for timing plans, while newer ones have a very large capacity (i.e. microprocessors). Timing plans are optimized at the planning level, where the signal control related data is kept. Several macroscopic tools enable to keep the traffic count data and to optimize timing plans for several scenarios (i.e. VISUM or TRANSYT).

6.2.4 Determination of scaling, smoothing and weighting factors

After the generation and optimization timing plans, the factors for scaling, smoothing and weighting should be determined. Generally, each signal company uses two different scaling factors: one for volume and another for occupancy. As a general rule, the volume scaling factor should be set to the saturation flow rate of the approach, where the system detector is located [Balke et al.,1997]. The significance of saturation flow rate and its determination methods with a case study in Istanbul is described in chapter 2 and 3, respectively. The scaling factor for occupancy depends upon whether or not queues build on the system detectors. For foreknown-congested approaches, the scaling factor should be set to equal the highest occupancy rate [Balke et al., 1997]. Beside these rules about scaling the collected data considering traffic responsive control system, the collected data should also be scaled for the neural network controller. Then, most of the neural controllers can only learn with the input values less than 1.

Another factor to determine in this step is the smoothing factor. Smoothing factor is used to eliminate the negative impact of short-term fluctuations in traffic volume. Generally, exponential smoothing is used:

$$\overline{m}_i = \alpha m_i + (1-\alpha)\overline{m}_{i-1} \quad (14)$$

Where, $\overline{m}_i$ refers to a new smoothed value, m_i refers to the actual measured value and $\overline{m}_{i-1}$ is the old smoothed value in the last period. α refers to the smoothing factor and generally is set between 0.15 and 0.5. Smoothing is significant to avoid frequent change of timing plans due to stochastic change in traffic volume. However, if the plan selection interval is large enough, smoothing is not necessary.

Since the significance of streets is not equal, the detectors can be previously weighted according to their importance. For example, the user can assign a higher weight for the system detectors near the critical intersection. However, it is recommended that the detectors should be weighted equally for initial implementation. If required, the weighting factors can be adjusted later [Balke et al., 1997].

The weighting is also used for neural network (see Figure 39). The weights are used to calculate output of corresponding layer. At the beginning of training, these weights can be assigned randomly and are adjusted during the training to reduce the error and to close to the best solution. Smaller values are also recommended for starting weights.

6.2.5 Determination of neural network architecture

One of the useful methods for the implementation of a pattern recognition algorithm is the use of artificial neural networks (see chapter 3 for other methods). In this step, the architecture of the artificial neural networks should be determined. Since offline optimization is required and timing plans should be previously calculated, pattern recognition method with the supervised manner is suitable for the proposed system. Feed-forward neural networks (also called multi-layer perceptrons) with a back propagation algorithm are suitable tool for pattern recognition. Multi-layer perceptrons (MLP) provide a general framework for representing non-linear functional mappings between input and output sets. This is achieved by representing a non-linear function of a single variable called the activation function [Bishop, 1995]. As shown in Figure 39, multi-layer perceptrons are in each node, and each layer is connected to the adjacent layer. The suitable numbers for the number of hidden layers and the number of neurons (units) in each hidden layer varies according to the problem type. There is not any optimum solution, which could be applied for each problem type. In this study, the number of neurons in the input and the output layer varies according to the number of system detectors and the number of traffic state or corresponding timing plans, respectively. The written code for this algorithm allows this flexibility to change the input and output vectors according to the system detectors and the required timing plans, which also enables finding of the optimum number of units (layer and neuron) for the hidden layer.

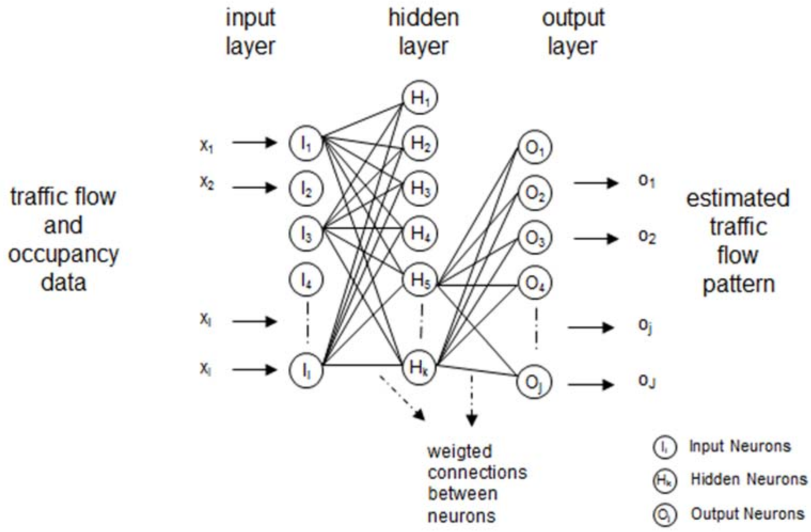


Figure 39: Architecture of multi-layer perceptrons

6.2.6 Training of Neural Networks for Recognition

As aforementioned, multi-layer perceptrons have a supervised learning strategy. That means that each input vector has to be labeled with a target vector, which represents previously clustered traffic states. During the training it is intended that output vectors will be close to the target vectors. This will be achieved after numerous iterations. The stopping criteria should be previously determined and it varies according to the importance of failure for particular problem. Essentially, the learning of multi-layer perceptrons is accomplished with the following four steps:

1. *Preparation:*

Before starting, the data should be scaled and divided into two parts: learning data and test data. With the learning data the network weights are modified in order to close up the system outputs to the desired responses. After training with a back-propagation algorithm, the system should be validated with other elements of the sample before implementation. The forward calculation starts with initial random weights.

2. *Forward calculation:*

In this step, first the sum of the weights is calculated using (15) and the output of corresponding layer is determined using an activation function. Generally, MLP networks use the sigmoidal function. This action is performed for each layer of the network.

$$u_k = \sum_{i=1}^{I} w_{ki} * x_i \tag{15}$$

$$\varphi(u_k + b_k) = \frac{1}{1 + e^{-u+b_k}} \tag{16}$$

where:

u_k = input of the processing unit for hidden layer

x_i = output of the previous layer

w_{ki} = weights between layers

b_k = bias

The previous equations are used for each layer of the network.

In the output layer, the error of the system is calculated using (17):

$$E_j = \frac{1}{2}\sum_{n=1}^{J}\left(t_j - o_j\right)^2 \tag{17}$$

where:

E_j = system error as cost function

t_j = target value

o_j = output of the system

J = number of output units

The root mean squared error for total samples is calculated using (18)

$$E_m = \frac{1}{M}\sqrt{\sum_{m=1}^{M} E_{j,m}} \tag{18}$$

where:

M = number of sample in the output layer

The aim of the training in neural network is to minimize this error during the learning process.

3. *Back-propagation:*

After feed forward calculation the first results are obtained to see the difference between the output and the target values. The aim of training in the neural networks is to minimize the error with adjusting the weights. The adjustment of the weights and the minimization of error can be achieved with back-propagation algorithm using the generalized delta rule. As shown in Figure 40 and 41, the back-propagation starts the adjustment of weights from the output layer and ends at the input layer. The calculations that are presented in both figures must be applied for each neuron in each layer. The figures represent a network with only one hidden layer.

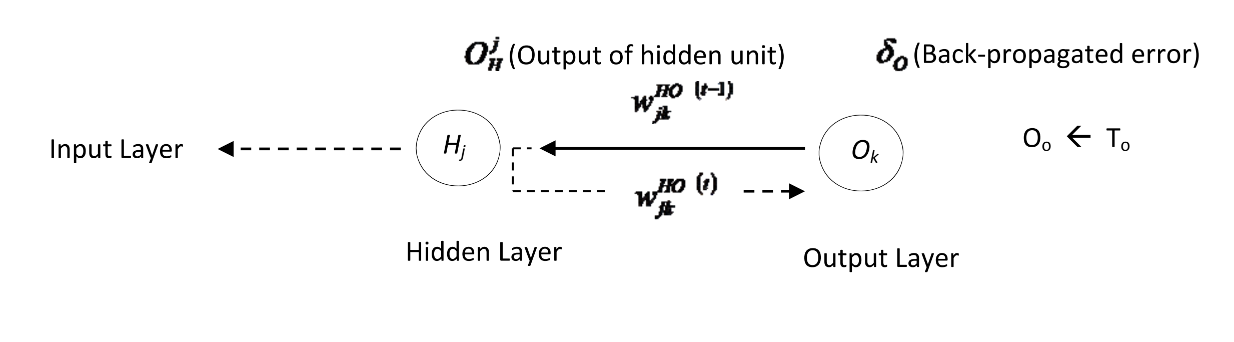


Figure 40: Back-propagation and weight adjustment between hidden and output unit

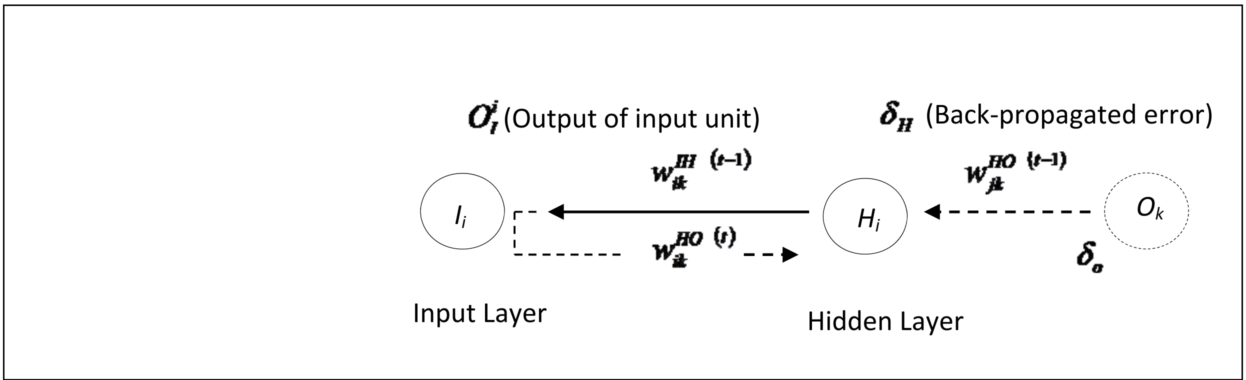


Figure 41: back-propagation and weight adjustment between input and hidden units

The general formula for the adjustment of weights is:

$$w_{jk}^{new} = \alpha w_{jk}^{old} + \Delta w_{jk} \tag{19}$$

The term Δw_{ij} is the first derivative of error (*E*) with respect to the weights

$$\Delta w_{jk} = -\eta * \frac{\partial E}{\partial w_{jk}} \tag{20}$$

The derivation can be formulated as the chain rule

$$\frac{\partial E}{\partial w_{jk}} = \frac{\partial E}{\partial O_o} * \frac{\partial O_o}{\partial u_k} * \frac{\partial u_k}{\partial w_{ij}} \tag{21}$$

$$\frac{\partial E}{\partial w_{jk}} = -(T_o - O_o) * O_o * (1 - O_o) * O_H \tag{22}$$

The first part of this formula is named as back-propagation error of output layer. The back-propagated error can be separated from weight adjustment and the formulas can be rewritten as follows:

$$\delta_O = O_O(1-O_O)(T_O - O_O) \tag{23}$$

The adjustment of weights between hidden and output layer can be rewritten as the following:

$$w_{jk}^{HO\ (t)} = \alpha w_{jk}^{HO\ (t-1)} + (\eta \delta_O O_H) \tag{24}$$

With the same way the back-propagated error for the hidden unit is calculated, and the input weights are adapted using (25) and (26):

$$\delta_H = O_H(1-O_H)\sum_O (\delta_O w_{jk}^{HO(t-1)}) \tag{25}$$

$$w_{ik}^{IH(t)} = \alpha w_{ik}^{IH\ (t-1)} + (\eta \delta_H O_I) \tag{26}$$

where:

O = outputs corresponding layers

δ = error of units in each layer

η = learning rate

α = momentum value

4. *Iterative process*

Neural networks learn after numerous iterations. The iteration will continue either until an acceptable error or predetermined number of iteration. The error rate or the number of iteration is defined by user and varies according to the problem type. Details about the stopping criteria are described in Chapter 3. After the training phase the system is tested and validated.

6.2.7 Test of the system and validation using microscopic traffic flow simulation

After the training, the neural networks should be tested to prove the performance. Sometimes, the system can memorize, but not generalize, due to several reasons like the type of the problem, structure of ANN or input samples. It means that the system memorizes and therefore recognizes the patterns during the training, but cannot give adequate results with new samples. Therefore, the system is tested using the data, which is not used during the training. The performance of the system is presented as a ratio of correct outputs to the total responses.

$$P = \frac{C}{T} * 100 \tag{27}$$

where:

P = performance of the system

C = correct responses

T = total responses

To increase the system performance and to reduce the iteration number the system architecture (number of neurons and other parameters) can be adjusted using an experimental design. This process will be described in the following chapter.

If the system training and the test performance are good enough, the system can be validated using a microscopic traffic flow simulation. The case studies to evaluate the performance of the proposed system are described in the next chapter. During these simulation studies, first a hypothetical network is created to test the ANN architecture, to adjust learning parameters and to examine system performance with the place and number of system detectors. Secondly, a real traffic network is taken from Istanbul with current traffic data and examined with the adjusted parameters and compared the performance of the system with current timing plans.

6.3 Conclusion

In this chapter a simplified traffic responsive control method is described. The system is based on pattern recognition method and artificial neural networks. The main advantages of this system are its relatively simple structure and the use of a less number of detectors. The main reason for that is to reduce the installation and maintenance costs, and to reduce the problems of developing cities malfunctioning detectors.

Since the neural network is used for pattern recognition, the system can continuously learn during the operation. If a new traffic state appears, the used scenario and corresponding timing plans will increase.

The next chapter will evaluate the proposed system with two arterial networks using microscopic traffic flow simulation VISSIM. The next chapter will also examine the further reduction in the number of detectors and a ranking of detectors to reduce redundancy.

7 Simulation environment

In the previous chapter a framework for developing large cities was presented and a simplified traffic responsive control method was proposed. This chapter will evaluate this proposed method using two case studies. Each step that was presented in Figure 38 (on page 70) is implemented for both case studies. After the training of the proposed system a simulation link is conducted to test the system using a microscopic traffic flow simulation environment (see Figure 42). The next section will describe the details of implementation of simulation study.

7.1 Simulation Model

The core of the microscopic traffic flow simulation is VISSIM [PTV, 2011]. VISSIM offers two external control modules: dynamic link library (DLL) and component object model (COM) that gains an advantage over other models. COM interface enables the automation of certain tasks in VISSIM by executing COM commands from an external program. The COM interface works with several programming languages (i.e. Python, Java, C++ or VBA) [PTV, 2011]. In this research, COM interface is used to access the model data and simulations and so to control the timing plans, which was the specific aim of the study. Microsoft Excel and Visual Basic for Applications (VBA) is used to load the network, to set different random seeds and to control the traffic flow. Another reason for choosing VISSIM and COM interface is that the artificial neural network controller is coded using VBA (a sample code for training of ANN is attached in Appendix B). Hence, one combined tool is used for the training and the testing of neural network controller, and for validating the system using microscopic traffic flow simulator. For the simulation studies a PC with Intel® Core™ i5-2400 CPU (3.1 GHZ and 8 GB RAM) was used. As illustrated in Figure 42, the trained and tested ANN-controller sends actual timing plans to VISSIM and receives the volume and occupancy data from VISSIM. The communication interval is determined as 5 min, but the interval can be adjusted according to specific conditions. The Figure 42 is the represents only the step 8 of Figure 38 (on page 70).

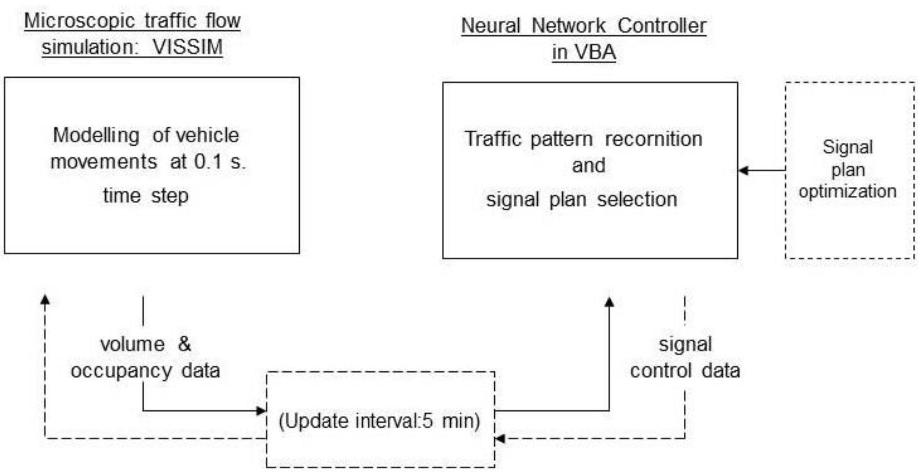


Figure 42: Simulation environment VISSIM with COM interface

The proposed controller is first tested using a hypothetical small arterial network to examine the number of detectors, the detectors positions etc. Secondly, a real arterial network is chosen in the historical district of Istanbul to validate the system. For the optimization of timing plans TRANSYT [TRL, 2011] and VISUM [PTV, 2011] were used. TRANSYT is well-known optimization software developed in UK and VISUM is a macroscopic transportation modeler developed in Germany. VISUM recently also introduced an optimization tool.

7.2 Sample arterial network and sensitivity analysis of parameters

The evaluation of the pattern recognition based system was conducted at first using a small hypothetical arterial network. The arterial consists of three signalized intersections and eight system detectors (see Figure 42). 12 different scenarios (see Appendix C) were developed, that represented different traffic states like morning and afternoon peaks, weekend etc. Volume and occupancy data is collected from the detectors that are placed on VISSIM network and used for the training and testing of neural networks.

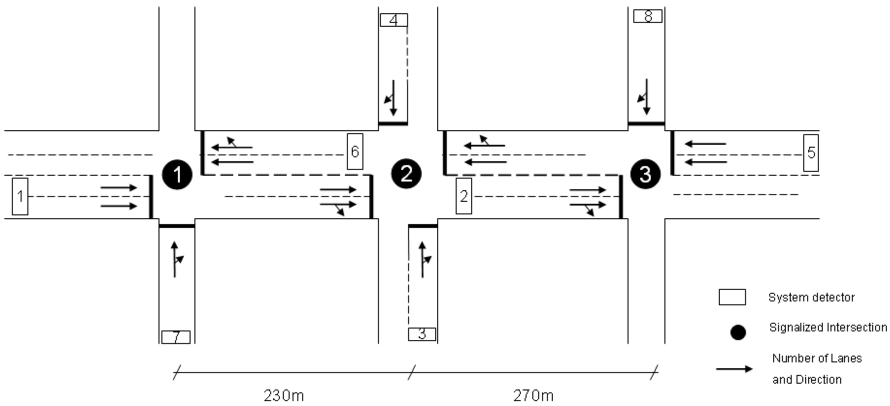


Figure 43: hypothetical arterial network and system detector locations

Feed forward neural networks (also known multi-layer perceptrons) are chosen for pattern recognition. In multi-layer perceptrons (MLP), input and output units are determined according to the number of detectors and the number of traffic states, respectively. However the hidden unit and some learning parameters can be selected according to problem types. Therefore an experimental design is performed to determine the optimal structure of MLP. During this experiment the number of hidden neurons, momentum value and learning rate was adjusted. As shown in Figure 43, the system learns with some parameters outstanding, whereas with some other parameter settings the system cannot learn but only memorizes the training data, and hence the test performance reduces. After 2000 iteration steps the best performance is achieved with the design group 26 (25 hidden neurons, momentum term = 1 and learning rate = 0.05). The Figure 44 shows the comparison of the learning performance of four data sets. The data set group 26 also learns faster than other groups.

The 'root mean squared errors' (RMSE) for the training and test sets of 30 groups are listed in Table 13. These results were reached after 5000 iterations for each group. Especially, increasing the learning rate causes memorizing and increases the test error. The results show that the momentum term and the number of hidden neurons have minimal effect on both the training and test performances. However, these analysis and results assist the determination of the optimal structure only for this problem. The number of input and output units and sample data may change the optimal settings and also effect of hidden neuron and momentum term.

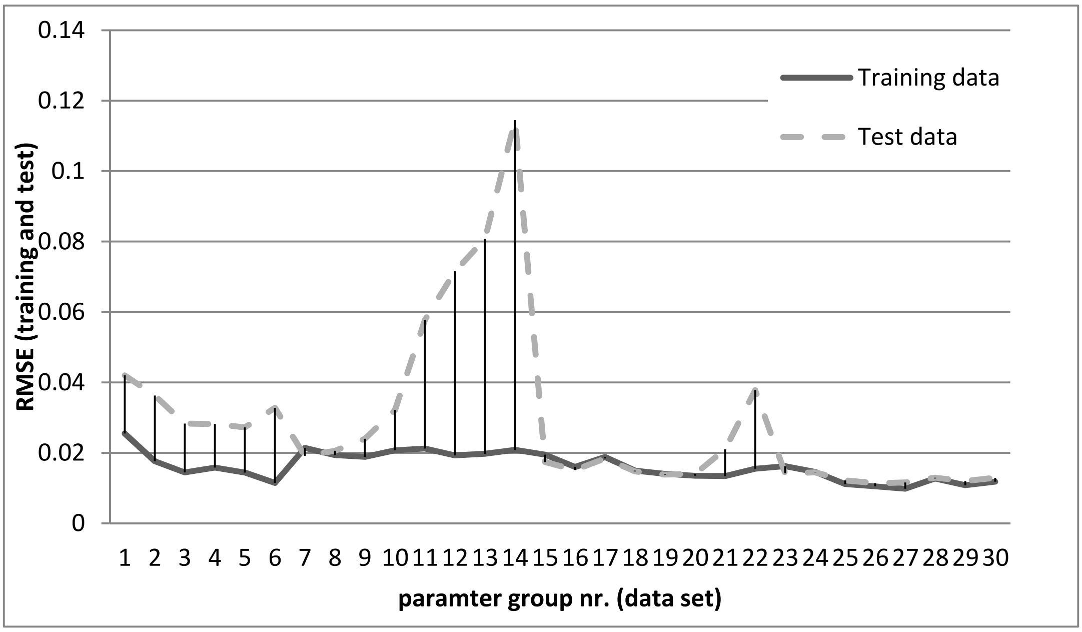


Figure 44: training and test errors with several parameter groups

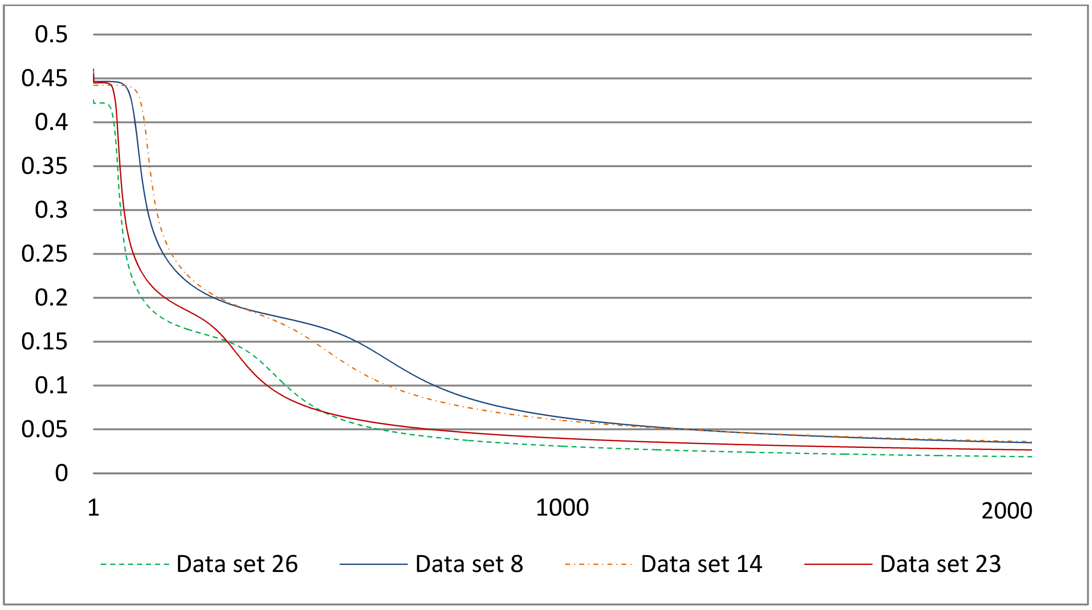


Figure 45: training process according to four parameter groups

Table 13: System performance according to several parameter settings

Data set no.	number of hidden neuron	momentum value	learning rate	RMSE Training	Test
1	12	0,5	0,3	0,025497	0,042014
2	12	0,6	0,3	0,017625	0,036273
3	12	0,7	0,3	0,014459	0,028317
4	12	0,8	0,3	0,015827	0,028191
5	12	0,9	0,3	0,014403	0,027222
6	12	1	0,3	0,011419	0,032776
7	12	0,5	0,05	0,021406	0,019112
8	12	0,5	0,1	0,019331	0,020572
9	12	0,5	0,2	0,018847	0,023994
10	12	0,5	0,3	0,020717	0,032141
11	12	0,5	0,4	0,021239	0,057735
12	12	0,5	0,5	0,019257	0,071533
13	12	0,5	0,6	0,019733	0,080745
14	12	0,5	0,7	0,020866	0,114473
15	12	0,6	0,05	0,019468	0,017398
16	12	0,7	0,05	0,015997	0,01512
17	12	0,8	0,05	0,018816	0,018403
18	12	0,9	0,05	0,014931	0,014735
19	12	1	0,05	0,014052	0,013814
20	12	0,9	0,1	0,013465	0,014026
21	12	0,9	0,2	0,013397	0,020987
22	12	0,9	0,3	0,015473	0,037821
23	8	1	0,05	0,016208	0,014217
24	15	1	0,05	0,014566	0,014428
25	20	1	0,05	0,011141	0,012147
26	25	1	0,05	0,01048	0,011395
27	30	1	0,05	0,009808	0,011577
28	20	0,9	0,05	0,012713	0,012942
29	25	0,9	0,05	0,010813	0,011996
30	30	0,9	0,05	0,011796	0,012898

- **Simulation results**

After the sensitivity analysis the system was evaluated using a microscopic traffic flow simulation. Since this study was on a hypothetical network, the traffic volume was determined using several scenarios. Therefore 12 different scenarios were developed and optimized using macroscopic optimization tool TRANSYT. The optimized TRANSYT timing plans and a sample transit output file is attached in Appendix D.

This experiment is performed using three cases. Since, the traffic state is at its peak in the morning and afternoon hours and the unpredictable traffic demand usually appears during off-peak periods, the first experiment with a three-intersection arterial network is performed with fluctuating traffic during off-peak periods. The three cases are:

Case 1: optimized pre-timed (signal plans are optimized using TRANSYT 13 for each time period)

Case 2: optimized pre-timed (change the traffic situation to am-peak for a short-time period)

Case 3: optimized real-time (real-time signal plan selection using proposed controller)

The results verify that a well-designed and optimized time-of-day signal plan can work well enough as long as the traffic situation does not change. But the traffic patterns change during the day as well as seasons. Figure 45 shows the results of three-intersection arterial network based on the midday traffic volume and signal data. Case 1 contains the base scenario that corresponds to the midday traffic volume and related signal plans. It should be noted, that the traffic volume was increased for Case 2 for 30 minutes without any change in signal plans. Case 3 proposed the use of MLP controller, which chooses the optimal timing plan for the prevailing traffic conditions. With the introduction of a MLP controller, the average speed increases by 27%, the average delay and average number of stops decreases by 15% and 55%, respectively. The results indicate that Case 1 is better than Case 3, because the total traffic volume in Case 3 is higher than in the base case. It should be considered that all three cases use optimized timing plans.

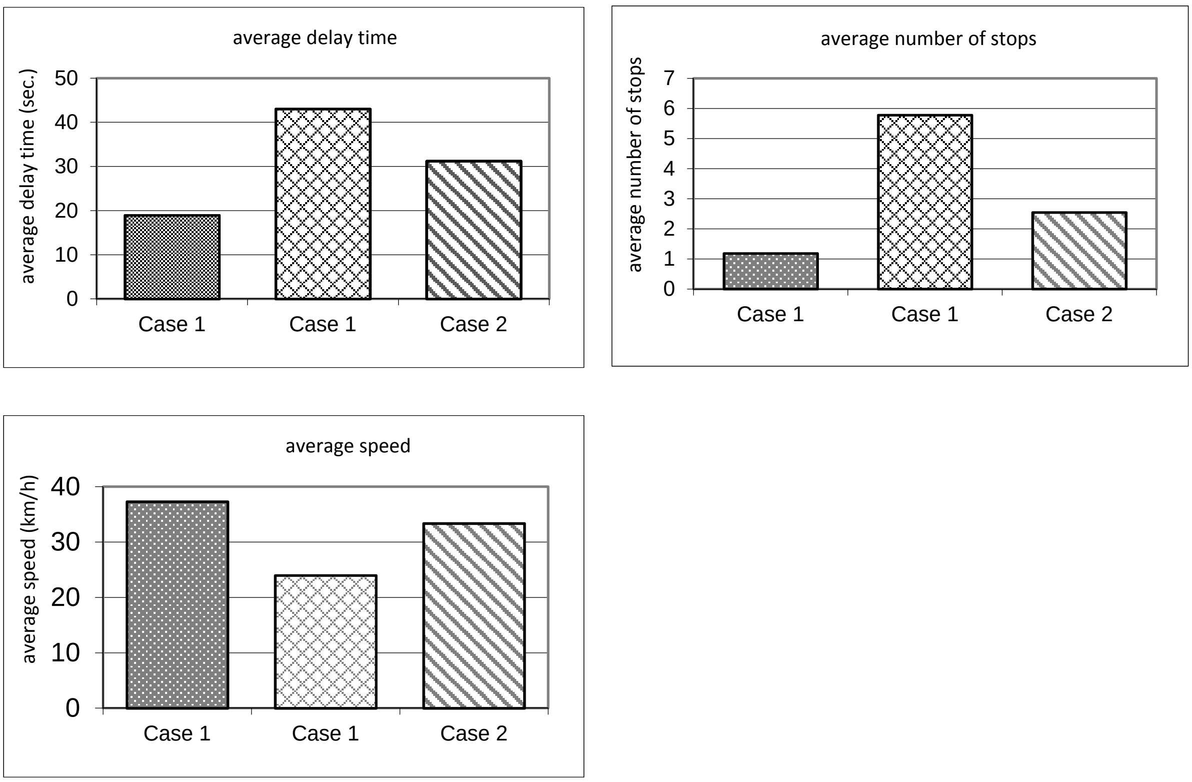


Figure 46: Simulation results of three-intersection hypothetical arterial network

7.3 System performance by reduction and malfunction of detectors

This section analyses the system performance by a reduction in the number of detectors in the planning phase and the influence on the system performance of potential detector malfunctions or communication failures during the operation phase. The use of a less number of detectors was one of the main objectives of this study, since it will reduce the maintenance and also the communication cost of the system. Another aspect is that the detector malfunctions or communication failures during the operation. For example, Sunkari et al. [2010] report that about 50% of the inductive loops are malfunctioning at any given time in the United States. This point is especially important for traffic actuated control and adaptive control systems. Generally, adaptive control systems switch to time of day mode or use historical data if detector malfunctions are perceived. Some adaptive systems also consist of a traffic-actuated system inside. These systems can switch to an actuated mode [Stevanovic, 2010]. Considering the situation in developing countries, it is not just technical problems with communication systems or on detector cards that cause data loss, but also some cities suffer under detection problems (especially with inductive loops) because of uncoordinated road works (pipeline excavation etc.). This section examines both detector reduction; to reduce redundant information (input features) and the system performance under detector malfunctions.

Through reduction methods the number of features can be decreased and thus the number of quantities, which need to be observed, measured and transmitted, respectively. Additionally, fewer input features reduce the training time of neural networks. [Teeuwsen et al. 2002] Since the proposed system in this study is based on a less number of detectors, a feature reduction method can be a useful tool in planning phase.

Neural network researchers introduced several methods. For example, Castellano and Fanelli [1999] proposed an approach to select the features in the neural networks. They developed a simple criterion to select the input features to be removed. Tsai [1999] proposed a reduction algorithm for probabilistic neural networks (PNN). Tsai developed an iterative approach that utilizes a weighted PNN with one smoothing factor for each variable in the feature reduction stage. Another feature reduction method, which was proposed and evaluated in [Wang et al., 1998] and [Wang et al., 2000] will described in the following paragraphs more in detail.

Wang et al [2000] used the clamping technique, which is based on a hypothesis that was previously described and demonstrated in [Wang et al., 1998]. The main idea of the clamping technique is that the more effective a feature is, the more severe the adverse effect on the performance of system will be, when the input for this feature is clamped to its mean value [Wang et al. 2000].

$$\varepsilon(x_i) = 1 - \frac{g(X(x_i = \overline{x_i}))}{g(X)} \tag{28}$$

where g(X) represents the generalization performance of the neural network, and $X(x_i = \overline{x_i})$ is the clamped generalization performance of the network when the input x_i is clamped to its mean $\overline{x_i}$.

The clamping technique and the factor ranking are carried out by the following procedures (Wang et al.2000):

- Train the neural network through a commonly used learning algorithm.

- Test the network with the test data that contains P patterns and calculate the generalization performance of the network, g(x).
- Calculate the mean value of each input in the data set;

$$\overline{x_i} = \frac{1}{P}\sum_{p=1}^{P} x_{ip} \quad \forall\, i = 1,2, \dots n \tag{29}$$

- Clamp each input by replacing x_{ip} with x_i for all test patterns. Put the clamped test data set through the net and calculate the clamped generalization, $g(X(x_i = \overline{x_i}))$.
- Calculate the impact ratio $\varepsilon(x_i)$, $\forall\, i = 1,2, \dots n$.
- Rank the impact ratio $\varepsilon(x_i)$ in descending order. The higher $\varepsilon(x_i)$ indicates that input x_i has a stronger effect on the output, and hence more important.

As expected, detectors (Nr. 2 and 6) on the major intersection of these example networks (see Figure 43 on page 84) are significant for pattern recognition. Interestingly, further accuracy could be achieved if the Detector Nr.8 is removed from the system, rather than being included. This can be caused from the collected data or from the detector itself, which is placed on Minor Street. The number of required input features and its redundancy is problem specific and depends on the size of the signal control network and the required data. Therefore, it should be noted that feature reduction techniques should be re-applied for different traffic network and detectors. The ranking of detectors are listed in Table 15. Considering detector malfunctions the system performance starkly reduces. Figure 47 shows the root means squared errors with the failure of corresponding detectors and Figure 48 indicates the difference in system performance in removing the detector at the planning and operation level using two sample detectors.

Table 14: The clamped generalization and the ranking of detectors on sample network

Rank order	Factor	$\varepsilon(x_i)$
1	Detector 2	0,704
2	Detector 6	0,593
3	Detector 3	0,296
4	Detector 4	0,259
5	Detector 5	0,111
6	Detector 1	0,037
7	Detector 7	0
8	Detector 8	-0,074

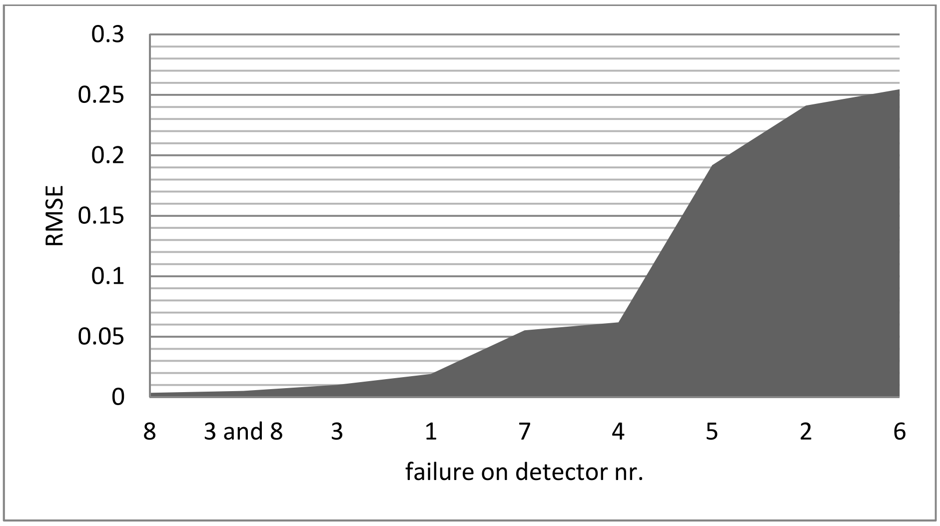


Figure 47: Change in RMSE with malfunction of corresponding detector

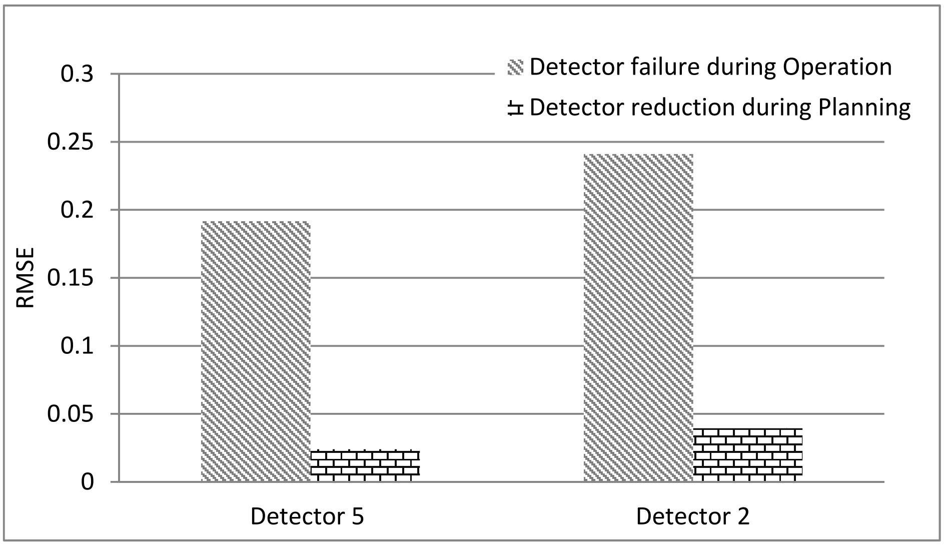


Figure 48: Comparison RMSE during malfunction and reduction of detectors

Additionally, the number of outputs (in this case the number of traffic states and corresponding timing plans) has also an effect on the performance of the system. The more timing plans that are created, the less performance should be expected. The overall results indicate that the learning and generalization feature of neural networks further enable the reduction of the number of detectors. If the optimal numbers of input and output features for respective problems are determined, the system could perform better.

7.4 Evaluation of the proposed system within a realistic arterial network

7.4.1 Study area and data collection

In order to validate the performance of the system on a realistic arterial network of relevant size, the system is examined within a simulation study. For the test network Fevzipasa Street in the historical district of Istanbul was chosen. This arterial network consists of nein signalized intersections and three signalized pedestrian crossings (pelican crossing) as shown in Figure 48. Except the pedestrian crossings, all signals are operated in time-of-day mode. Since the detectors are not available on this arterial street, traffic counts are obtained manually.

Detailed study was also conducted at intersection Nr.1136 “Akdeniz Street-Fevzipasa Street” for calibration of simulation parameters point-speed and saturation flow measurements. Traffic volumes on three approaches of this intersection are illustrated in Figure 50. Highly recurrent morning and afternoon peaks are not common on this arterial, because of the traffic behavior, pedestrian activities, and congested traffic on freeways, which are close to this arterial. Since this arterial is located in a busy shopping area and therefore traffic volume differs only marginally, the transportation authority uses the same timing plan with small adjustments for weekdays in the current traffic signal controllers. The main arterial consists of two lanes for both directions. However, most of the intersections do not have any separate turn lanes. High mixed public transport, pedestrian and parking activities reduces the physical speed and causes congestion on this arterial. The traffic counts at these nine intersections are listed in Table 15 and the current timing plans and detailed layouts of these intersections are presented in Appendix D.

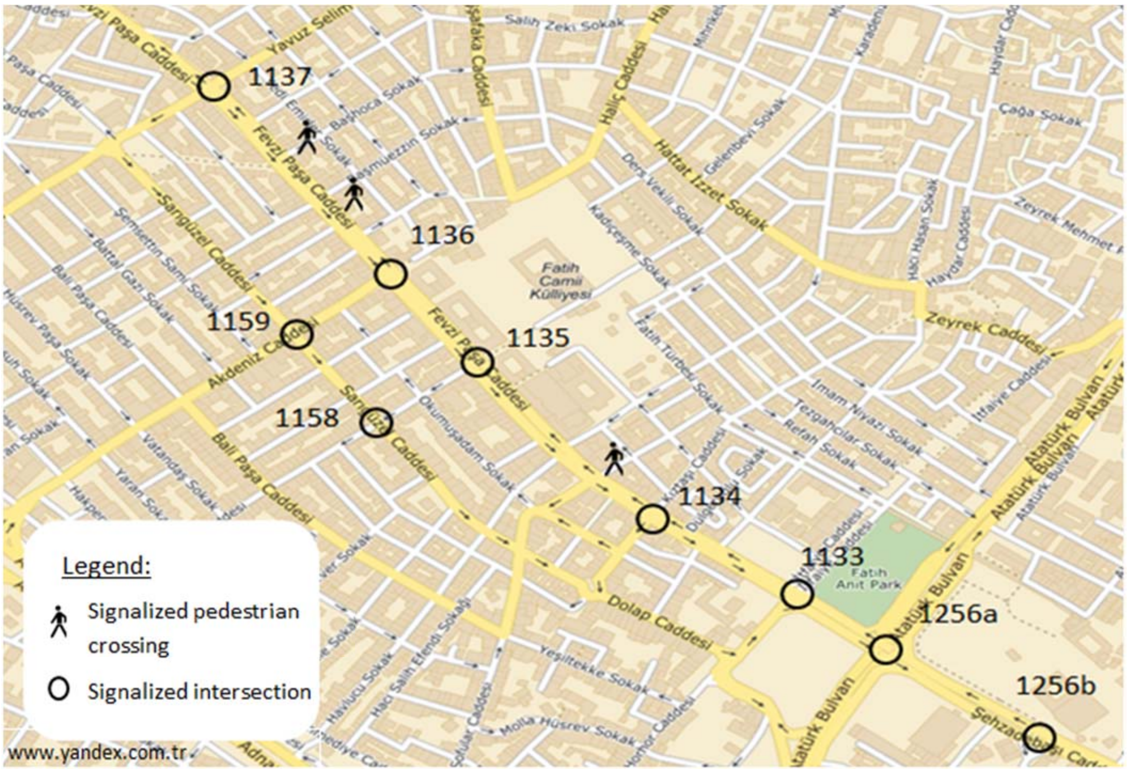


Figure 49: Study arterial and signalized intersections in historical district of Istanbul

Using the traffic data at intersection Nr.1136 a cluster analysis is performed to examine the current traffic state clusters and to compare it with existing classification. The k-means clustering is implemented and the best silhouette value (mean value: 0.7079) is found with five clusters. However, the transportation department of the city operates these signals using four timing plans. To simplify the comparison with existing situation, four clusters were used in this study. Clustered traffic volume and occupancy data are illustrated in Figure 51.

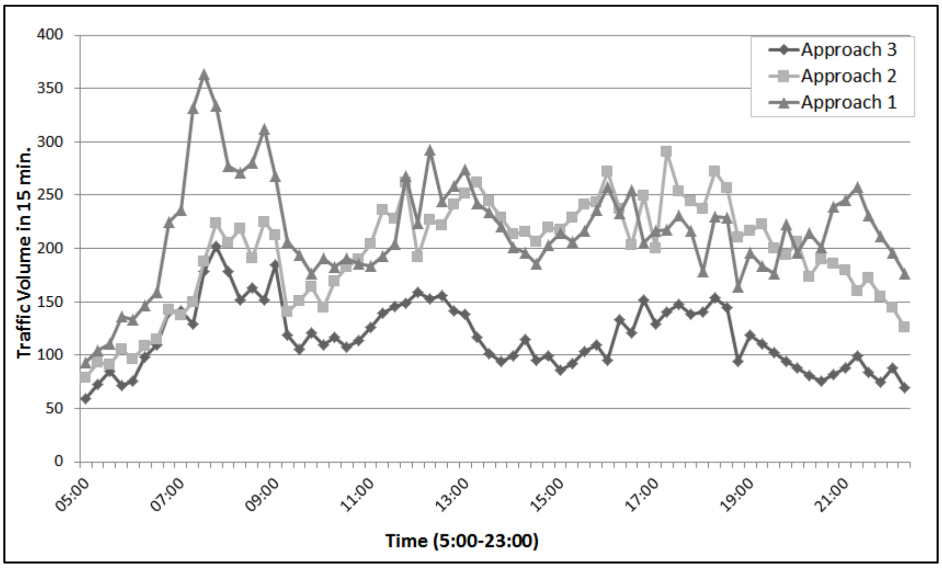


Figure 50: Traffic volume on Akdeniz-Fevzipasa intersection (Nr.1136)

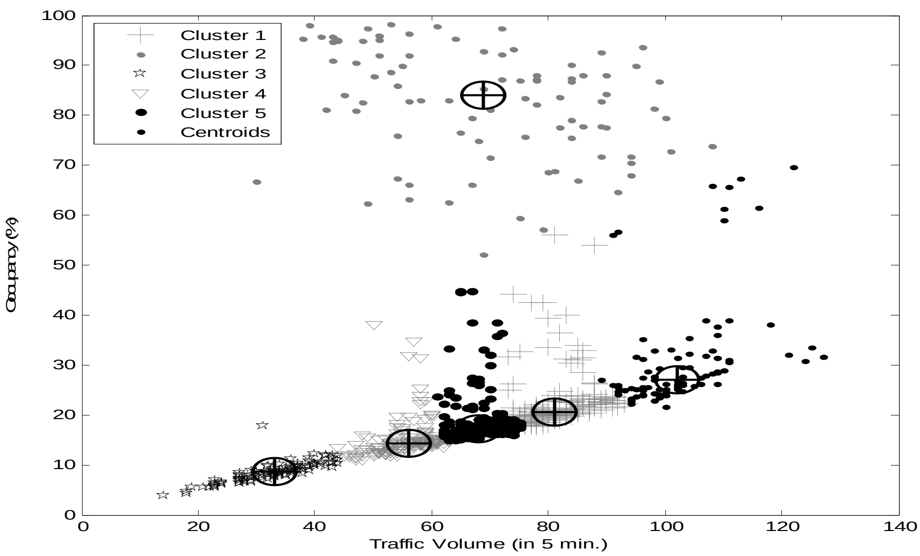


Figure 51: Data clustering using traffic volume and occupancy

Table 15: Traffic counts on the network

Intersect. Nr.	Count Period	Northwest Bound			Southeast Bound			Northeast Bound			Southwest Bound		
		L	T	R	L	T	R	L	T	R	L	T	R
1133	Morning (3h)	0	2360	436	22	2899	0	301	367	2449	0	0	34
	Midday (2h)	0	2285	378	35	1483	0	272	347	1549	0	0	49
	Afternoon (3h)	0	4373	1133	41	1791	0	237	549	1894	0	0	45
1134	Morning (3h)	380	2523	0	0	2723	420	235	159	63	0	0	0
	Midday (2h)	484	2372	0	0	1366	348	121	137	73	0	0	0
	Afternoon (3h)	1043	4009	0	0	1650	487	186	270	96	0	0	0
1135	Morning (3h)	808	2198	0	0	3223	133	0	0	0	0	0	0
	Midday (2h)	745	1926	0	0	1813	60	0	0	0	0	0	0
	Afternoon (3h)	1609	2883	0	0	2225	94	0	0	0	0	0	0
1136	Morning (3h)	523	2741	0	0	2135	13	665	0	1171	0	0	0
	Midday (2h)	477	1531	0	0	1875	10	401	0	757	0	0	0
	Afternoon (3h)	703	1889	0	0	2884	41	495	0	1093	0	0	0
1137	Morning (3h)	857	1898	409	487	2623	731	0	0	0	413	621	12
	Midday (2h)	643	1471	333	382	1527	579	0	0	0	317	319	3
	Afternoon (3h)	984	2351	320	464	1941	955	0	0	0	349	458	218
1158	Morning (3h)	0	0	0	0	934	87	0	0	0	128	879	0
	Midday (2h)	0	0	0	0	610	152	0	0	0	106	830	0
	Afternoon (3h)	0	0	0	0	796	211	0	0	0	174	1776	0
1159	Morning (3h)	0	0	0	227	486	235	0	1607	578	11	496	0
	Midday (2h)	0	0	0	171	349	210	0	1059	369	33	504	0
	Afternoon (3h)	0	0	0	178	379	325	0	1486	581	41	861	0
1256a	Morning (3h)	0	1885	725	2416	2945	0	0	7833	2388	0	6536	1185
	Midday (2h)	0	1864	790	1269	1748	0	0	4844	1586	0	4901	1163
	Afternoon (3h)	0	4110	1031	1559	2000	0	0	6226	1807	0	8437	1708
1256b	Morning (3h)	0	1304	0	0	3577	444	295	0	159	0	0	0
	Midday (2h)	0	1516	0	0	2259	433	325	0	128	0	0	0
	Afternoon (3h)	0	3829	0	0	2286	459	555	0	230	0	0	0

7.4.2 Software-in-the-loop simulation

The timing plans of the first arterial network were optimized using TRANSYT 13. For this second simulation study, VISUM is used for the optimization of traffic signals, since the new version of VISUM also provides an optimization module. The VISUM network and the timing plans are provided by Istanbul Metropolitan Municipality. Optimized timing plans can be found in Appendix D and the existing timing plans with intersection layouts are attached in Appendix E. The following steps were conducted during this simulation study.

At first, the current signal plans of the city and the manually obtained turn counts are inputted into the network. Secondly, T-Flow Fuzzy module of VISUM was used to estimate the origin-destination matrices and path flow from intersection counts. OD estimation and path flows are not directly required for traffic signal optimization, but can be used to examine several scenarios. After the optimization of timing plans the network was exported to VISSIM with its network, demand and timing plans. 8 system detectors were placed on the VISSIM network to generate cross-sectional counts. The algorithm was trained with a part of data and tested with the other part. The timing plans are prepared using Vehicle Actuated Programming (VAP) tool of VISSIM, to prevent the problems by switching the timing plans.

In simulation phase, the VISSIM generates and sends volume and occupancy data to the external controller. The ANN-controller runs the algorithm and recognizes the traffic state and selects the corresponding timing plans. The selected timing plan is recalled from VAP controller.

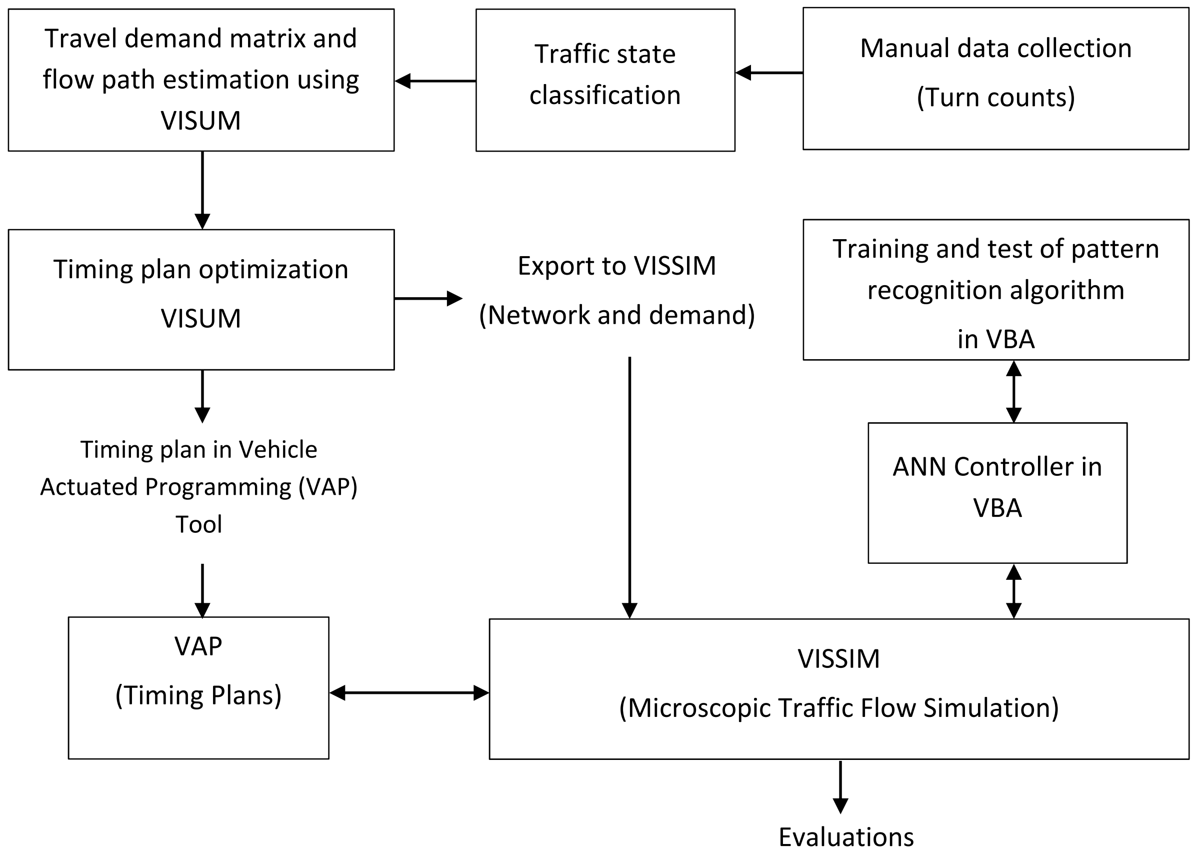


Figure 52: Flow diagram of simulation case study

The simulation results indicate that despite relative homogenous traffic demand near or above capacity, proposed pattern recognition based traffic responsive controller reduces key traffic indicators such as delay and the number of stops. Considering the average number of stops, speed and average delay the proposed controller improve the traffic quality on this arterial networks.

The average number of stops decreases during morning peak hours by about 15%, midday hours 43% and afternoon peak hours by about 23%. The average speed increases during morning, midday, and afternoon peak hours at, 2%, 28% and 12%, respectively. As illustrated in Figure 50, the average delay time is also reduced in the time period of the three experiments by 6%, 43% and 22%, respectively. The highest profit in this experiment could be achieved in midday (or off- peak) time period. One reason could be that the traffic technicians pay more attention on peak hours and prepare or recalculate the timing plans at peak hours more frequently. Another reason could be the fluctuating traffic on the arterial network that cannot be served with fixed-time traffic signals.

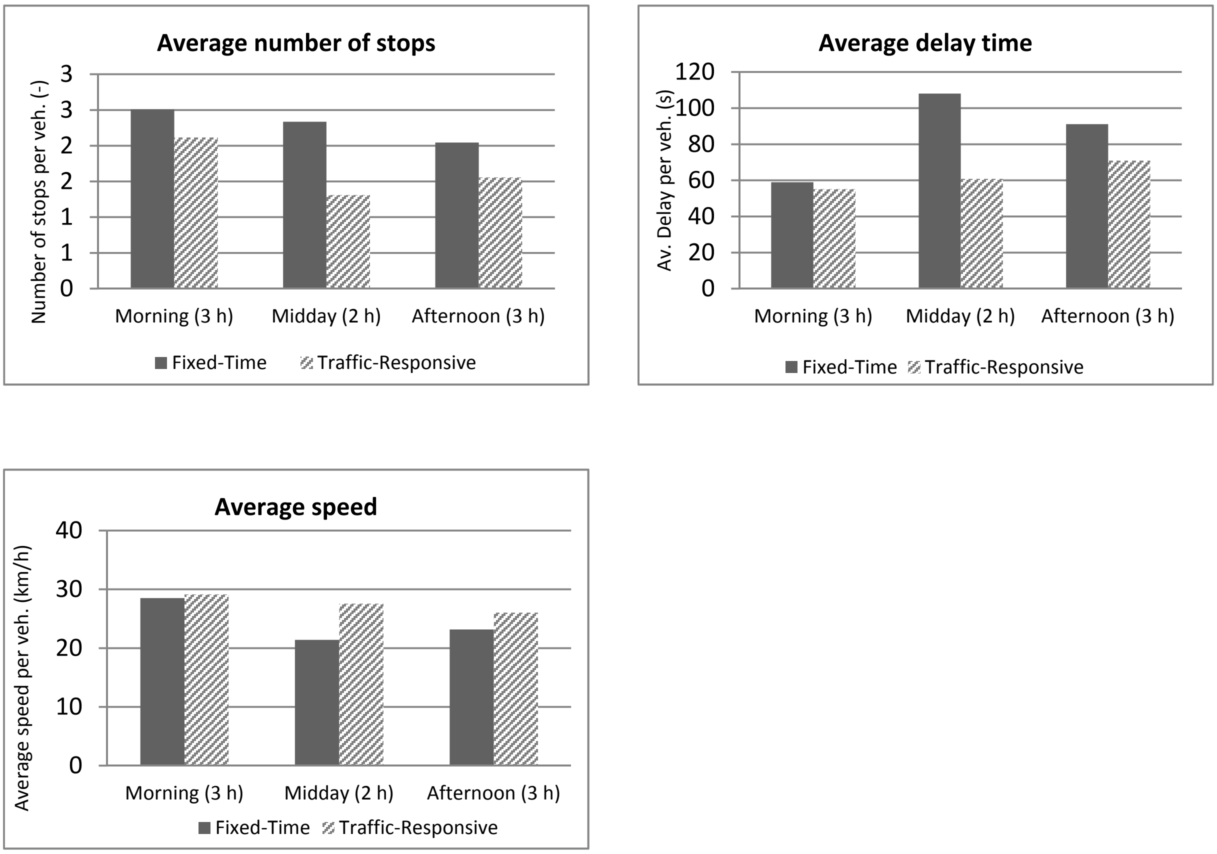


Figure 53: comparison of performance with and without proposed controller

7.4.3 Scenarios

Since the proper data collection is costly and onerous particularly for developing and emerging countries, scenario management at the macroscopic level can provide opportunities to examine several scenarios and to optimize corresponding timing plans. The scenario management is a useful tool not only for a fixed time controlled system but also for the real-time system as proposed in this dissertation study. As mentioned before, the system is based on a fewer number of detectors. Therefore, 4 to 8 detectors for a 10-intersection network could not be sufficient to optimize the traffic signals. But, if the traffic signals are optimized and several scenarios are prepared based on manual counts and travel behavior data, the system can then recognize the traffic state and select a suitable timing plan for corresponding traffic situation using few detectors.

As described in [von der Ruhren 2006] three cases occur generally in urban traffic networks:

- Change in traffic demand (i.e. peak hours, leisure traffic)
- Change in network capacity (i.e. incident, road works)
- Change in both demand and capacity (i.e. demonstrations)

In this section, changes in traffic demand and network capacity are evaluated within two scenarios. The first scenario reflects the demand changes in various directions of the network, while the second scenario involves change in the network capacity due to lane closure. For these scenarios previously optimized timing plans were used. Special timing plans for the scenarios were not prepared.

Scenario 1a: For this scenario the traffic demand in Northwest-Southeast direction is increased about 40% between 12.00 and 14.00. The base traffic demand and the variation in scenarios are illustrated in Figure 52. The average demand in southeast direction is about 800 vehicles per hour for midday situation.

The current signal timing plans that were used in Chapter 7.4.2 are also used in this scenario without any change. Since in this arterial network the off-peak period is also high as peak hours and almost reaches to the capacity, the increasing in traffic demand for this scenario affects the system performance drastically. As shown in Table 17, the use of proposed controller reduces the average number of stops in 32%. The average delay decreases 9%, while the average speed increases about 5%.

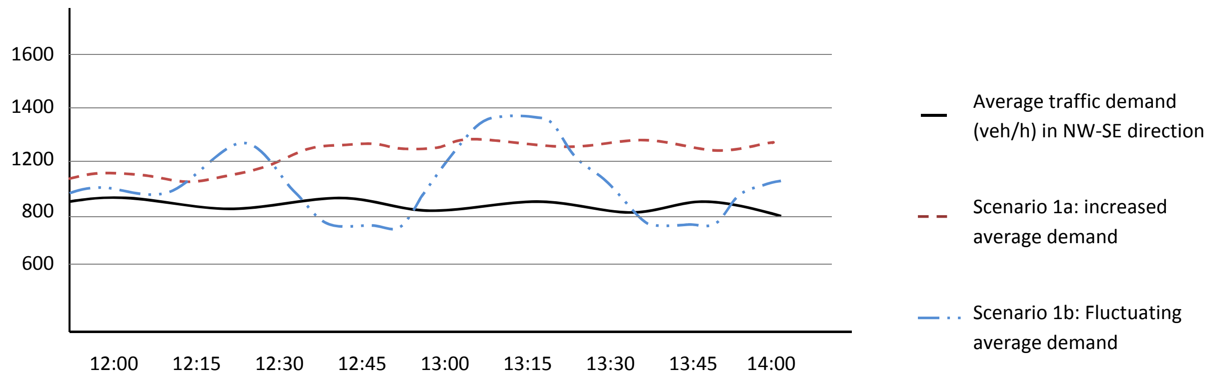


Figure 54: Traffic demand in Scenario 1a and 1b

Table 16: Performance comparison results of Scenario 1a (12:00-14:00)

	Fixed-time control	Traffic responsive control
Average number of stops (-)	12.6	8.6
Average delay time per veh. (s)	114.9	104.9
Average speed per veh. (km/h)	20.9	21.9

Scenario 1b: A significant factor to use traffic dependent control method is to cope with fluctuating traffic demand. This scenario tests the system under fluctuating traffic condition. As presented in Figure 52, the traffic demand in southeast direction was changed in each 15 minutes. In comparison with optimized fixed time control the proposed traffic responsive model reduces the average number

of stops and the average delay about 38% and 17%, respectively. The average speed increases %12. The performance results of Scenario 1b are listed in Table 18.

Table 17: Performance comparison results of Scenario 1b (12:00-14:00)

	Fixed-time control	Traffic responsive control
Average number of stops (-)	15.6	9.7
Average delay time per veh. (s)	134.2	111.8
Average speed per veh. (km/h)	19.0	21.3

Scenario 1c: This scenario examines the performance of the system with increasing traffic demand on the major cross street. The traffic volume on Akdeniz Street was increased about 50%. Instead lack of special timing plans for demand in main Cross Street as commonly used in TRPS control methods; the proposed controller outperforms existing system with increasing demand. The comparative results are listed in Table 19. By using the proposed controller, the average number of stops decreased about 35%. The average delay per vehicle decreased 23%, whereas the average speed increased about 18%.

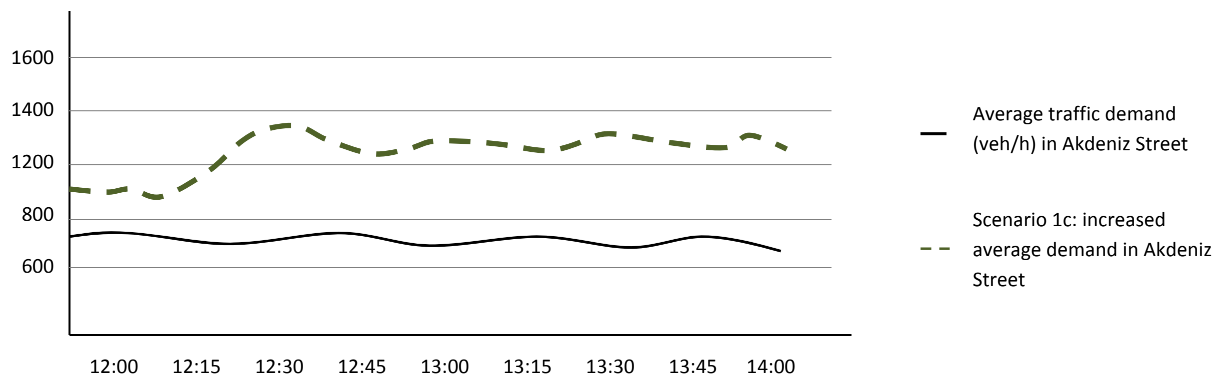


Figure 55: Traffic demand in Scenario 1c (Akdeniz Street)

Table 18: Performance comparison results of Scenario 1c (12:00-14:00)

	Fixed-time control	Traffic responsive control
Average number of stops (-)	18.7	12.3
Average delay time per veh. (s)	144.9	111.5
Average speed per veh. (km/h)	18.1	21.3

Scenario 2: Frequently occurring incidents traffic accidents or road works reduce the lane capacities by partly or entirely lane closure. Therefore to see the system performance during capacity reduction a lane is closed and evaluated with both fixed time and traffic responsive controllers. As shown in Figure 56, one traffic lane (for about 50 meter) on northwest bound is closed from 12:00 to 14:00 and the performance of the system is examined without any change in traffic demand and timing plans.

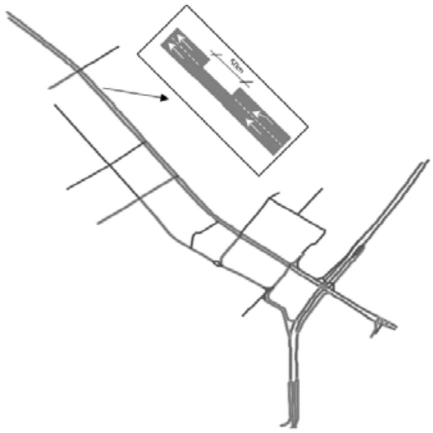


Figure 56: Lane closure in scenario 2

The proposed traffic responsive controller can reduce the average number of stops and average delay per vehicle about 25% and 5%, respectively. The average speed increases about 3% (see Table 20). The profit with the proposed system is not enough, since for this arterial only the existing timing plans are used with optimization. Additionally, because the traffic demand is almost same throughout the day, the optimized timing plans are also almost the same. Therefore, the real profit of proposed system can be seen with timing plans quite different from each other. The scenario management tool enables to study and to optimize several scenarios for road incidents as in this situation.

Table 19: Performance comparison results of Scenario 2 (12:00-14:00)

	Fixed-time control	Traffic responsive control
Average number of stops (-)	11.5	8.5
Average delay time per veh. (s)	94.3	89.8
Average speed per veh. (km/h)	22.9	23.6

7.4.4 Conclusion

This chapter presented the evaluation of proposed system in two different networks and using several scenarios. The first hypothetical arterial network was used to examine several scenarios (see Appendix C). The importance of detectors is also tested with the clamping technique and the ranking of detectors were presented. This method can be useful for the praxis to reduce the number of detectors.

On the real arterial network, beside the comparison of the system with existing demand, two scenarios for change in demand and supply were also examined. The proposed system outperforms the existing time of day method both in increasing and fluctuating traffic demand. In the practice, using a scenario management tool, several scenarios can be examined and more timing plans can be optimized for different traffic conditions. In that case the performance of the system can increase further. However, this depends on capacity of signal controllers. A scenario management tool may also serve for long-term change in traffic demand or long-term road construction works.

In this Chapter the system was tested only using microscopic traffic flow simulation. However, the algorithm itself is ready to use in real applications.

8 Conclusions and Recommendations

8.1 Conclusions

Developing cities must struggle with traffic problems under an increasing population due to high birth rates and migration to the large cities and an increasing number of car ownership due to economic development. It is not a case of developed cities. One can say that the developed cities uses dynamic traffic control method under a static population and static number of cars, while, generally, developing cities use static traffic control method under dynamic population and number of cars.

However, developing large cities also cannot be considered in the same category. For example cities with high cyclist or motorcyclist rates cannot have same signal control related problems as in car-oriented cities. It is also difficult to categorize the cities according to national income, since the large cities; particularly in developing countries are more prosperous than other cities and are privileged by the governments.

Traffic signals are one of the most important control methods in urban areas. They can control conflicts and hence improve safety, and can also increase the capacity of intersections. However, if they are not properly operated, they themselves could cause traffic congestion. The main focus of this dissertation study was to develop a simplified and low-cost traffic signal control strategy that uses a reduced number of detectors to recognize the traffic state and to select the best suitable timing plan.

Within the scope of this dissertation study, two preliminary studies were conducted to determine the current situation in developing large cities. Briefly stated, the first preliminary study investigated the technical situation, while the second one researched the physical or infrastructural situation for developing large cities.

To investigate state-of-the-art traffic signal control in developing large cities a web-based interview was conducted with traffic agencies of both developing and developed large cities. Responses were obtained from 19 of 42 cities in both developing and developed countries. The response rate was about 45%. The results indicated that there is a perceivable difference both in the number of signalized intersections and in control methods. The main control method in developed cities is real-time coordinated control, while isolated intersection control is the major control method in developing cities. Exclusively, in developing countries, the increase is not only in the number of signalized intersections due to increasing population, but also the ratio of signalized intersections per 1000 inhabitants. This also decreases the average age of controllers in developing large cities.

Besides the technological and institutional infrastructure in developing countries the physical and geometrical situations of intersections were also investigated within saturation flow measurements. To observe the saturation flow variations in different districts and traffic conditions the intersections were grouped into three categories. Each category represented different physical and behavioral conditions (i.e. geometric standard of intersection or pedestrian activities). Small variations between districts could be expected. However, this preliminary study indicated that there is a significant difference between the districts in Istanbul. This difference increases also according to the number of minibuses, which is a special transport mode without regular stops. Exclusively, the coming together of pedestrian activities and the number of minibuses reduces the intersection capacity by about 50%.

In this dissertation study, a simplified traffic responsive control method is presented and evaluated using microscopic traffic flow simulation. The algorithm is based on pattern recognition method. The application of this method in this study is that the recognition of the actual situation and the selection of suitable timing plans. The timing plans are, therefore, previously optimized using the macroscopic tools TRANSYT and VISUM.

The pattern recognition method is applied using the artificial neural networks. The learning and generalization ability makes the neural networks very useful tool in estimation problems. Thus, traffic state could be determined using only a few detection points, while modern sophisticated traffic adaptive systems require a high-density sensor mesh. Since it is relatively difficult to keep the inductive loop detectors in operation in developing cities due to a lack of coordination in agencies that are continuously doing road works and excavations, hence disturbing the detectors or communication, the use of a fewer number of detectors becomes apparent.

The pattern recognition based model is evaluated using the two different arterial networks. The first hypothetical network is used to test several scenarios and timing plans. The learning parameters and the impacts of the number of detectors are also examined in this case study. Here, the clamping technique is used for the ranking of the detectors. The clamping technique indicated that some detectors could be redundant for recognizing the traffic state.

For the second case study, a real arterial network in the historical district of Istanbul was chosen to evaluate the proposed pattern recognition based model. Since the detectors are not available on this network, the data is collected manually. The macroscopic planning tool VISUM was used to complete the data to calculate the flow paths and to optimize the timing plans. The model reduced the number of stops during morning peak hours by about 15%, midday hours 43% and afternoon peak hours by about 23%. The average speed increases during morning, midday, and afternoon peak hours at, 2%, 28% and 12%, respectively. The average delay time is reduced in the time period of the three experiments by 6%, 43% and 22%, respectively.

Since long-term data was not available on this network several traffic conditions were also examined using scenarios. In the demand scenarios, the traffic demand was fluctuated in Northwest, Southeast bounds and a cross street. The system shows better performance than existing fixed time control for all three scenarios. Additionally, one supply scenario was examined. In this scenario, one lane in Northwest bound was closed and the system performance was analyzed. The proposed controller improved the traffic flow instead of capacity reduction.

During this study the new optimization model of VISUM could also be examined using microscopic traffic flow simulations. Since the detector data was not available, counts and occupancy data could be generated using the microscopic traffic flow simulation. If long time data directly from detectors could be available the training of system can achieve more accuracy. The algorithm needs further investigation with detectors data particularly for the volume-occupancy relationship for oversaturated conditions.

8.2 Recommendations and Future Works

The European or American guidelines about traffic signal control and traffic management cannot be directly used in developing cities. They should prepare either new guidelines or calibrate the parameters. This requires particular research in detail.

One of the main problems in developing large cities is that the collection of data and the management of existing data. Particularly, to keep the inductive loop detectors in operation, continuous coordination between different agencies is required, since the road works and excavation works does not always disturb the detectors or communication.

Most of the developing large cities have their own special transport mode that is dominantly used. For example in Hanoi, motorcycles are dominant, while in some district of Istanbul the minibuses. These are generally not considered in modern guidelines, despite their significant effect on traffic flow. Special research is required to determine the parameters of these special transport modes for signal timing calculations.

In this study, a relative simplified model is proposed that can be an interim solution for some cities or a low-cost and long-term solution for some others. Pattern recognition method is used to select a complete timing plan with cycle, split and offset lengths. However, pattern recognition can also be used in signal control with several variations (i.e. estimation only cycle in actuated control). Some further investigation or improvements in the model should be done:

- In this study, 5 minutes intervals are used to query traffic situations, however larger intervals could allow the system to remain more stable and the change frequency in cycle time and offset in praxis more practical. The query interval of the model should be examined in further studies.
- Another issue related previous point is the negative impact of transition period from one timing plan to another. The transition period in general problem in signal control in praxis. Incremental change in timing plan or long intervals to change the timing plans are generally applied by several agencies. This point should also be considered and investigated in detail for real implementations.
- The model can also be examined with larger intervals for cycle queries and shorter intervals for green adjustment and offsets particularly for congested and high pedestrian activity arterials.
- The model can be improved to also provide traffic actuated signalization at intersections. In that case the model provides the minimum and the maximum durations of green time in each approach and estimate the suitable cycle length for each situation.
- Queue length/link capacity ratio, degree of saturation can be used as further parameters for estimation problem. For example, instead of continuous querying of the actual traffic state, the same timing plan can be used between the minimum and maximum thresholds of the degree of saturation. The model can be improved with additional parameters such as queue length and queue capacity and should be examined at least at isolated intersections.

- In this study, supervised learning method was applied due to a labeling problem with the traffic state and timing plan. However, an unsupervised learning method can also be used. In that case, instead of using pre-optimized timing plans the system will estimate the cycle, split and offset times.

9 References

Abbas, M.M.; Sharma, A. (2006): *Multiobjective Plan Selection Optimization for Traffic Responsive Control,* Journal of Transportation Engineering, pp. 376-384

Abbas, M.M.; Pesti G.; Chaudhary N.A.; Li P. (2009): *Illustrative Field Configuration and Evaluation of Traffic-Responsive Control*, In: Journal of Transportation Engineering, pp.591-599

Abdelaziz, S.L. (2008): *Study and Evaluation of Traffic Responsive Control on a Large Arterial Network*, Virginia Polytechnic Institute and State University, Master Thesis

Abdennour, A.; Al-Ghamdi A. S. (2002): *Artificial Neural Network Application to Estimation of vehicle Headways in freeway Sections*, In: Proceedings of 81st annual meeting of the Transportation Research Board Annual Meeting (2002), Washington DC

Al-Mudhaffar, A. (2008): *Impacts of Traffic Signal Control Strategies*, VDM Verlag

Akcelik, R. (1981): *Traffic Signals: Capacity and Timing Analysis*, Research report Nr.123, Australian Road Research Board

Asar, A.; Ullah, M.S.; Ahmed, J.; Hasnain, R. (2008): *Traffic Responsive Signal Timing Plan Generation based on Neural Network*, In: Proceedings of 4th IEEE Conference on Automation Science and Engineering (2008), pp.833-838

Axhausen, K.W.; Fellendorf, M.; Hook, D. (1989): Zur Abhängigkeit der Zeitbedarfswerte von der Knotenpunktsgeometrie, In: Straßenverkehrstechnik, Heft 5 (1989), pp.170-175

Balaji, P.G.; German, X.; Srinivasan, D. (2010): *Urban traffic signal control using reinforcement learning agents,* In: IET Intelligent Transport Systems (2010), Vol.4, pp.177-188

Balke, K.N.; Keithireddipalli, R.S.; Brehmer, C.L. (1997): *Guidelines for Implementing Traffic Responsive Mode in TxDOT Closed-Loop Traffic Signal Systems*, Research Report 2929-3F, Texas Transportation Institute

Barter, P.A. (1999): *An international Comparative perspective on Urban Transport and Urban form in pacific Asia: The challenge of rapid motorization in dense cities*, Murdoch University, Dissertation

Bishop, C. M. (1995): *Neural Network for Pattern Recognition*, Oxford University Press

Basu, J.K.; Bhattacharyya, D.; Kim, T.H. (2010): *Use of artificial neural network in pattern recognition*, In: International Journal of software Engineering and its application (2010), Volume 4, pp. 23-34

BMVBS, Bundesministeriums für Verkehr, Bau und Stadtentwicklung (2010): *Anwendung und Analyse modellbasierter Netzsteuerungsverfahren in städtischen Straßennetzen- Endbericht AMONES*,

Boltze, M. (1988): *Optimierung von Umlaufzeiten in der Lichtsignalsteuerung für Straßennetze*, Darmstadt University of Technology, Dissertation

Boltze, M. (2005): *Verkehrsmanagement und Telematikeinsatz systematisch planen – Vorstellung des „Leitfadens Verkehrstelematik des BMVBW“,* In: Berich zum FIV-Symposium „Neue Planungsinstrumente für die Verbesserung der Verkehrs und Umweltqualität (2005), pp.19-27

Boltze, M.; Busch F.; Friedrich B.; Friedrich, M.; Kohoutek S.; Löhner H.; Lüßmann J.; Otterstätter T. (2011): *AMONES: Anwendung und Analyse modellbasierter Netzsteuerungsverfahren in städtischen Straßennetzen- Teil 2: Analyse und Ergebnisse*, In Straßenverkehrstechnik (2011), Heft 6, pp.374-381

Braun, R. (2008): *Ein echtzeitfähiger Evolutionärer Algorithmus zur netzweiten Optimierung der Lichtsignalsteuerung*, Technische Universität München, Dissertation

Brilon, W.; Wietholt, T.; Wu, N. (2007): *Kriterien für die Einsatzbereiche von Grünen Wellen und verkehrsabhängigen Steuerungen,* Berichte der Bundesanstalt für Straßenwesen, Volume 163

Castellano, G.; Fanelli, A.M. (1999): *Variable selection using neural-network models,* In: Neuro-computing, pp.1-13

Chang, E.C.P. (1999): *Estimate Traffic Control Patterns Using a Hybrid Neural Network*, In: Proceedings of International Joint Conference on Neural Networks, Washington DC, 1999, pp.2763-2767

Chin, D.C; Spall, J.C.; Smith, R.H. (1999): *Evaluation of System-Wide Traffic Signal Control Using Stochastic Optimization and Neural Networks*, Proceedings of the American Control Conference, San Diego, 1999, pp.2188-2194

Chiu, S.; Chand, S. (1993): *Adaptive traffic signal control using fuzzy logic*, In: Proceedings of the first IEEE Regional conference on aerospace control systems (1993), pp.122-126

Christie, S.M.; Lyons, R.A. (2003): *Are mobile speed cameras effective? A controlled before and after study*, In: injury prevention, Volume 9, pp. 302-306

Cuong, D. Q. (2009): *Traffic Signals in Motorcycle Dependent Cities*, Darmstadt University of Technology, Dissertation

Di Taranto, C. (2007): *Absolute Priority*, In Traffic Technology International (2007), pp.100-101

Duda, R.O.; Hart, P.E., Stork, D.G. (2001): *Pattern Classification*, Wiley-Interscience

Favilla, J.; Machion, A.; Gomide, F. (1993): *Fuzzy Traffic Control: Adaptive Strategies,* In: Second IEEE International Conference on Fuzzy Systems (1993) pp.506-511

Fellendorf, M. (1991): *Graphentheoretisches Verfahren zur koordinierten Signalisierung von Knotenpunkten*, Karlsruhe Institute of Technology, Dissertation

Fellendorf, M.; Platzer, S.; Reich, B. (1999): *Innovative Steuerungskonzeptionen in europäischen Städten*, Research Report, PTV, Karlsruhe

FGSV, Forschungsgesellschaft für Straßen- und Verkehrswesen (2001a): *Hinweise zu Verkehrsrechnern als Bestandteil der innerörtlichen Lichtsignalsteuerung*, FGSV Verlag

FGSV, Forschungsgesellschaft für Straßen- und Verkehrswesen (2001b): *Handbuch für die Bemessung von Straßenverkehrsanlagen – HBS (Manual for the dimensioning of road infrastructures),* FGSV Verlag

FGSV, Forschungsgesellschaft für Straßen- und Verkehrswesen (2010): *Richtlinien für Lichtsignalanlagen – RILSA (Guidelines for Traffic Signals), FGSV Verlag*

Fletling, R. (2010): *Methodische Ansätze zur unscharfen Mustererkennung bei Deformationsmessergebnissen*, Technischen Universität Braunschweig, Dissertation

FSV, Österreichische Forschungsgemeinschaft Straße-Schiene-Verkehr (2009): *Richtlinien und Vorschriften für das Straßenwesen – RVS, Verkehrslichtsignalanlagen*, FSV Verlag

Friedrich, B. (1999): *Ein verkehrsadaptives Verfahren zur Steuerung von Lichtsignalanlagen*, Munich Technical University, Dissertation

Friedrich, B. (2000): *Steuerung von Lichtsignalanlagen, BALANCE- ein neuer Ansatz*, In Straßenverkehrstechnik (2000), Vol. 7, pp.321-329

Friedrich, B.; Schütte, C. (2000): *Strategic Control in Metropolitan Areas*, In: Proceedings of the 3rd IEEE Conference on Intelligent Transportation Systems (2000), Dear-borne, USA

Gardner, G. (1993): *The choice of traffic signal control for LDC's*, In: Conference of the Development and Planning of Urban Transport (1993), Tunis

Gardner, N.H. (1983): OPAC: *A demand responsive strategy for traffic signal control,* In: Transportation Research Record 906 (1983), pp.75-81

Gartner, N.H.; Pooran F.J.; Andrews C.M.(2002): *Implementation and Field testing of the OPAC adaptive control strategy in RT-TRACS*, In: Annual Meeting of the Transportation Research Board (2002), Washington, DC

Ge, Q.; Poschinger, A. (2011): *Forecast based cycle time calculation in Sitraffic Motion*, In: Proceedings ITS World Congress (2011), Orlando, USA

Gordon, R.; Braud, C. (2009): *Traffic Signal Operations and Maintenance Staffing Guidelines*, Report FHWA-HOP-09-006, Federal Highway Administration, Washington

Gordon, R.L.; Tighe, W. (2005): *Traffic Control Systems Handbook*, Report FHWA-HOP-06-006, Office of Transportation Management, Federal Highway Administration, Washington

Hamilton, R.J.; Pringle, R.D.; Grant, P.M. (1992): *Syntactic techniques for pattern recognition on sampled data signals*, In: IEE proceedings-E (1992), Volume 139, pp.156-164

Haykin, S. (2005): *Neural Networks: A Comprehensive Foundation*, Prentice-Hall Press

Hess, S.; Polak, J.W. (2003), *An Analysis of the Effects of Speed Limit Enforcement Cameras on Accident Rates,* Transportation Research Record, 180, pp.5-33

Hirschmann, K.; Fellendorf, M. (2010): *A toolbox to quantify emission reductions due to signal control*, in: TRB 2010 – 89th Annual Meeting (2010), Washington, USA

Huang, S.C; Huang, Y.F (1991): *Bounds on the Number of Hidden Neurons in Multilayer Perceptrons,* In*:* IEEE Transactions on Neural Networks (1991), pp.47-55

Hunt, P.B.; Robertson, D.I.; Bretherton, R.D.; Winton, R.I. (1981): *SCOOT – A Traffic Responsive Method of Coordinating Signals,* TRRL Laboratory Report 1014, Transport and Road Research Laboratory, Crowthorne, UK

Jain, A.K.; Man, J.; Mohiuddin, K.M. (1996): *Artificial Neural Networks: A Tutorial,* In: *Computer (1996)*, pp.36-44

Jain, A.K.; Duin, R.P.W.; Mao, J. (2000): *Statistical Pattern Recognition: A Review,* In: IEEE Transactions on pattern analysis and machine intelligence, Volume 22, pp. 4-37

Jain, A.K.; Duin, R.P.W. (2004): *Introduction to pattern recognition*, In: The oxford companion to the Mind, Second edition, Oxford University Press

Jain, V.; Sharma, A.; Subramanian, L. (2012): *Road Traffic Congestion in the Developing World*, In: Proceedings of ACM Symposium on Computing for Development (DEV), March 11-12, 2012, Atlanta, GA

Karaboga, D. (2004): *Yapay Zeka Optimizasyon Algoritmalari (artificial intelligence optimization algorithms),* Atlas Press, Istanbul

Karadag, A.; Alyuruk, M.; Gundogan, F. (2012): *Evaluation of Parking Violation System in Istanbul*, accepted for World Congress of Intelligent Transportation System World Congress, Vienna, 2012

Kaur, D.; Konga, E. (1994): *Fuzzy traffic lights controller,* In: Proceedings of the 37th Midwest symposium on circuits and systems (1994), pp.1507-15010

Khan, S.I. (1996): *Modular neural network architecture for detection of operational problems on urban arterials*, University of California, Irvine, Dissertation

Khuat, V. H. (2006): *Traffic Management in Motorcycle Dependent Cities*, Darmstadt University of Technology, Dissertation

Lapierre, R.; Obermaier A. (1988): *Entwicklung der Lichtsignalsteuerung in städtischen Straßennetzen (Development of signal control in urbarn road network),* In: 88 Jahre Strassenverkehrstechnik in Deutschland (1988), Kirchbaum Verlag

Li, Y.; Mück, J.; Matthias, P. (2010): *A recurrent neural network approach to network-wide traffic signal control*, In: Transportation Research Board Annual Meeting (2010), Washington DC

Lin, S.M. (1999): *Formulation and Evaluation of a methodology for network-wide signal optimization,* University of Florida, Dissertation

Liu, H.H.; Hsu, P.L. (2006): *Design and Simulation of adaptive fuzzy control on the traffic network*, In: International Joint Conference SICE-ICASE (2006), pp.4961-4966

Lu, W.; Zhang, Y.; Xie, Y. (2011): *A multi-agent adaptive traffic signal control system using swarm intelligence and neuro-fuzzy reinforcement learning,* In: IEEE forum on integrated and sustainable transportation systems (2011), Vienna, pp.233-238

Martin, P.T.; Perrin, J.; Chilukuri, B.R.; Jhaveri, C.; Feng, Y. (2003): *Adaptive signal control II*, Research Report, University of Utah

MathWorks, 2012: *Neural Network Toolbox,* Available under: http://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf (last access on 20.07.2012)

Mauro, V.; Di Taranto, C. (1990): *Utopia*, In: Control, computers, communication in transportation (1990), pp. 245-252

Mirchandani, P.; Head, L. (2001): *A real-time traffic signal control system: architecture, algorithms, and analysis*, In Transportation Research Part C (2001), pp.415-432

Möhring, R.H; Nökel, K.; Wünsch, G. (2006): *A model and fast optimization method for signal coordination in a network,* In: Proceedings of the 11th Symposium on Control in Transportation Systems (2006)

Morichi, S. (2005): *Long-Term Strategy for Transprt system in Asian megacities*, In: journal of the eastern Asia Society for Transportation Studies (2005), Vol. 6, pp.1-22

Mück, J. (2008): *Neue Schätz- und Optimierungsverfahren für Adaptive Netzsteuerungen*, In Straßenverkehrstechnik (2008), Vol. 12, pp.761-763

Mück, J. (2009): *The German approach to Adaptive Network Control*, In: Proceedings of mobil.TUM (2009), Munich

Murat, Y.S. (2001): *A Fuzzy logic traffic responsive signal control model for isolated intersections*, Istanbul Technical University, Dissertation

Murat, Y.S. (2006): *Comparison of fuzzy logic and artificial neural networks approaches in vehicle delay modeling*, In: Transportation Research Part C (2006), pp.316-334

Niemann, H. (2003): *Klassifikation von Mustern* (2. Edition in Internet), Available under: http://www5.informatik.uni-erlangen.de/fileadmin/Persons/NiemannHeinrich/klassifikation-von-mustern/m00-www.pdf (last access on 20.07.2012)

Niittymaki, J.; Kosonen, T.; Nevala R. (2001): *Fuzzy traffic signal control in major arterials*, In: Proceedings IEEE Intelligent Transportation Systems Conference (2001), pp. 346-351

Öztemel, E. (2006): *Yapay sinir aglari (Artificial neural networks),* Papatya Press

Ploss, G. (1993): *Ein dynamisches Verfahren zur Schätzung von Verkehrsbeziehungen aus Querschnittszählungen*, Technische Universität München, Dissertation

Priemer, C. (2011): *Kommunikationsdatenbasierte, dezentrale Lichtsignalsteuerung in städtischen Netzen*, Shaker Verlag

PTV (2011a): *VISUM 12- User Manual,* PTV AG

PTV (2011b): *VISSIM 5.4- User Manual,* PTV AG

Retzko, H.G.; Boltze M. (1993): *Zielorientiertes Verkehrsmanagement – das Fallbeispiel FRUIT für Frankfurt am Main und die Rhein-Main-Region*, In: Varosi Közlekedes (1993)- (hungarian technical magazin in urban traffic), Volume 6

Riedel, T. (2003): *Workshop on Traffic Control and Traffic Management*, Adaptive Control Ltd., Winthertur

Roess, R.P.; Prassas, E.S.; McShane, W.R. (2004): *Traffic Engineering*, Pearson Prentice Hall Press

Russel, S.J.; Norvig, P. (2003): *Artificial Intelligence- a modern approach*, Pearson Prentice Hall Press

Schlaich, J.; Haupt, T. (2012): *Large-scale signal optimization: traffic platform framework and applications*, In: 4th Transportation research board conference on innovations in travel modeling (ITM), Tampa, Florida, USA

Schnabel W.; Lohse D. (1997): *Grundlagen der Straßenverkehrstechnik und der Verkehrsplanung*, Band 1, Verlag für Bauwesen

Schnabel, W.; Scholtz, T.; Pohl, K. (2005): *Sättigungsverkehrsstärke in lichtsignalgesteuerten Knotenpunktzufahrten*, In: Straßenverkehrstechnik (2005), pp. 501-506

Sharma, A. (2004): *Determination of traffic responsive plan selection factors and thresholds using artificial neural networks,* Texas A&M University, Master Thesis

Shen, L. (2008): *Freeway travel time estimation and prediction using dynamic neural networks,* Florida International University, Dissertation

Siemens (2002): *MOTION zentral V3.0 – Planungshandbuch*, Siemens AG

Slinn, M.; Matthews, P.; Guest, P. (2003): *Traffic Engineering Design - Principles and Practice*, Butterworth-Heinemann Press

Schmöcker, J.D.; Bell, M.G.H. (2009): *Traffic Control: Current Systems and Future Vision of Cities*, in: International Journal of ITS research, Vol. 8 (2010), pp.56-65

Stevanovic, A.; Martin, P.T.; Stevanovic, J. (2007): VISGAOST: *VISSIM-based Genetic Algorithm Optimization of Signal Timings.* In: Proceedings *86th Transportation Research Board Meeting (2007)*, Washington

Stevanovic, A. (2009): *Review of adaptive traffic control principles and deployments in larger cities,* In: Proceedings of mobil.TUM (2009), Munich

Stevanovic, A. (2010): *Adaptive traffic control systems: domestic and foreign state of practice*, Research Report Nr.403, Transportation Research Board, Washington, DC

Sunkari, S.; Songchitruksa, P.; Charara, H.; Zeng, X. (2010): *Improved intersection operations during detector failures*, Technical Report 0-6029-1, Texas A&M University

Teeuwse, S.P.; Erlich,I.; El-Sharkawi, M.A. (2002): *Feature reduction for neural network based small-signal stability assessment*, In: 14^{th} Power System Computation Conference (2002), Sevilla

Theodoridis, S.; Koutroumbas, K. (2003): *Pattern Recognition*, Elsevier Academic Press

TRB, Transportation Research Board (2000): *Highway Capacity Manual*, National Research Council

TRL, Transport Research Laboratory (2010): *TRANSYT 13 User Guide*, Transport Research Laboratory Press

Tsai, C.Y., (1999): *An iterative feature reduction algorithm for probabilistic neural networks*, In: The international Journal of Management Science, pp.513-524

Turner, J.; Harahap, G. (1993): *Simplified saturation flow data collection methods*, Transport Research Laboratory, UK

Unadkat, S.B.; Ciocoiu, M.M.; Medsker, L.R. (2001) *Chapter 1: Introduction,* In: Recurrent Neural Networks – Design and Applications (Editors: Medsker, L.R.; Jain, L.C.), CRC Press, 2001

Van Zuylen, H.J. (1979): *The estimation of turning flows on a junction*, In: Traffic Engineering Control, pp.539-541

Vasconcellos, E. A. (2001): *Urban Transport, Environment and Equity: The Case for Developing Countries,* Earthscan Publications

Von Mörner, J.; Leupold, A.; Vesper, A. (2009): *Befragung zum Betrieb von Lichtsignalanlagen und deren Steuerung in innerstädtischen Straßennetzen*, Shaker Verlag

Von der Ruhren, S. (2006): *Kurzfristprognosen von Verkehrszuständen auf Basis von Verfahren der Mustererkennung und von dynamischen Routensuch- und Umlegungsverfahren*, RWTH Aachen, Dissertation

Vössner, S. (2010): *Optimization Methods for Operation Planning*, Lecture notes at Graz University of Technology

Wang, W.; Jones, P; Partridge, D. (1998): *Ranking pattern recognition features for neural networks*, In: Proceedings ICAPR (1998), pp. 232-241

Wang, W.; Jones, P; Partridge, D. (2000): *Assessing the impact of input features in a feedforward neural network*, In: neural comput & applic (2000), pp.101-112

Watanabe, S. (1985): *Pattern Recognition: Human and Mechanical*, Wiley Press

Webb, A.R. (2002): *Statistical pattern recognition*, 2.Edition, John Wiley & Sons Ltd, UK

Webster, F.V.; Cobbe, B.M. (1966): *Traffic Signals*, Road Research Technical Paper Nr.56, Her Majesty's Stationery Office

Wen, H.; Yang, Z.; Jiang, G., Shao, C. (2001): *A new Algorithm of Incident Detection on Freeways*, In: Vehicle Electronics Conference, Proceedings of the IEEE International (2001)

Wietholt, T. (2009): *Einsatzbereiche Grüner Wellen und verkehrsabhängige Steuerungen*, Ruhr-Universität Bochum, Dissertation

Willumsen, L.G. (1978): *Estimation of an O-D-matrix from traffic counts – a review.* Working paper 99, Institute for Transportation Studies, University of Leeds, UK

Yuan, H.C.; Xiong F.L.; Huai, X.Y. (2003): *A method for estimating the number of hidden neurons in feed-forward neural networks based on information entropy*, In: Computers and Electronics in Agriculture (2003), pp.57-64

Yulianto, B. (2003): *Applicatio of fuzzy logic to traffic signal control under mixed traffic conditions*, In: traffic engineering and control (2003), pp. 332-336

Zackor, H.; Busch, F.; Höpfl, W. (1991): *Entwicklung eines Verfahrens zur adaptiven koordinierten Steuerung von Lichtsignalanlagen*, In: Forschung Straßenbau und Straßenverkehrstechnik, German Ministry for Transportation, Berlin

Zavitsas K.; Kaparias I.; Bell M.G.H.; Nocera S. (2011): *Urban traffic Management and Intelligent Transport Systems: A European Perspective*, In: Proceedings 90th Transportation Research Board Annual Meeting (2011), Washington, DC

Internet references

OECD Statistics: http://www.oecd.org/home/0,3675,en_2649_201185_1_1_1_1_1,00.html

Oxford Dictionary, available under: http://www.oxforddictionaries.com

UN, United Nations Statistics: http://hdr.undp.org/en/statistics/hdi

World Bank Statistics: http://data.worldbank.org/topic/infrastructure

Yandex Map Service: http://www.yandex.com.tr

Demographia [2012]: http://www.demographia.com

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Networks
ANNDE	Artificial Neural Networks Delay Estimation
ART	Adaptive Resonance Theory
ATCS	Adaptive Traffic Control System
BALANCE	Balancing Adaptive Network Control Method
CC	Computational Channel
COM	Component object model
DLL	Dynamic Link Library
FL	Fuzzy Logic
GA	Genetic Algorithm
FGSV	Forschungsgesellschaft für Straßen- und Verkehrswesen
HBS	Handbuch für die Bemessung von Straßenverkehrsanlagen
HCM	Highway Capacity Manual
HDI	Human development index
HUTSIM	Helsinki University of Technology Simulator
ICA	Intersection Capacity Analysis
ITCARI	Isolated Traffic Control at Arterial Road Intersections
LOS	Level of service
LPF	Link performance function
LQV	Linear vector quantization
MAXBAND	Maximal Bandwidth traffic signal setting optimization program
MIP	Mixed-integer linear program
MLP	Multi-layer perceptrons
MOTION	Method for the optimization of traffic signals in online controlled networks
NFACRL	Neuro-fuzzy actor-critic reinforcement learning
NFDE	Neuro-fuzzy Delay Estimation
OD	Origin-destination
OECD	Organization for economic co-operation and development
OPAC	Optimization policy for adaptive control
PASSER	Progression Analysis and Signal System Evaluation Routine
pcu/h	Passenger car unit per hour
PI	Performance Index
PNN	Probabilistic neural networks
PS	Pattern Selection
PT	Public transport
RHODES	Real-time Hierarchical Optimized Distributed and Effective System
RILSA	German Guidelines for Traffic Signals
RMSE	Root mean squared error
SC	Signal controller
SCATS	Sidney coordinated adaptive traffic system
SCOOT	Split, cycle and offset optimization technique
SFR	Saturation Flow Rate
SIGOP	Traffic signal optimization program
SOAP	Signal optimization analysis package
SOM	Self-organizing map

SPOT	System for Priority and Optimisation of Traffic
SPSA	Simultaneous perturbation stochastic optimization
TOD	Time of Day
TP	Timing Plan
TRANSYT	Traffic network study tool
TRPS	Traffic Responsive Plan Selection
UN	United Nations
UTCS	Urban Traffic Control System
UTOPIA	Urban Traffic Optimisation by Integrated Automation
VAP	Vehicle Actuated Programming
VBA	Visual Basic for Application
veh/h	Vehicles per hour
VISGAOST	VISSIM-based genetic algorithm optimization of signal timings
VISSIM	Verkehr in Städten: SIMulation
VISUM	Verkehr in Städten: UMlegung

Appendix A – Example for Learning in Neural Network

XOR (eXclusive OR) gate is well-known sample in electric/electronic area. The idea is that if only one of the inputs to the gate is true then the output is true. Otherwise the output is false. Since it is the first non-linear problem that could be resolved using MLP, XOR problem is also important in history of neural networks. One can say the multi-layer perceptrons could be developed using XOR gate. Therefore XOR gate is most used example to show how to learn the multi-layer perceptrons. The following example (adapted from [Öztemel, 2006]) with seven steps describes how the multi-layer perceptrons learn.

Step 1: Sample collection

The first step is the collecting the samples. An XOR problem has four samples as listed in Table 1.

Table 20: Input and outputs of XOR gate

	Input 1	Input 2	*Output*
Sample 1	0	0	*0*
Sample 2	1	1	*0*
Sample 3	1	0	*1*
Sample 4	*0*	*1*	*1*

Step 2: Determination of structure of the network

The network has two inputs and one output. Therefore the network structure consists of two inputs and one out layers. To keep this example more comprehensible only one input layer and only two input units are chosen.

Step 3: Determination of system parameters

Generally, multi-layer perceptrons have two adjustable parameters, momentum term and learning rate. Suppose that the momentum term α = 0.8 and the learning rate η = 0.5

Step 4: determination of starting weight

The weights of multi-layer perceptrons must be between -1 and +1. The starting weights can be assigned randomly. In this small example we will set these values as in following matrices.

Weight between input and hidden unit:

$$w_{IH} = \begin{bmatrix} 0.5 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}$$

Weights between hidden and output unit

$$w_{HO} = [0.5 \quad 0.4]$$

Step 5: Presenting the samples to the network and forward calculation

The first sample [0, 0] for input and [0] for output is presented to the network and calculated as following:

$$u_k = \sum_{i=1}^{I} w_{ik} * x_i$$

u_1 = 0*0.5+0*0.3 = 0

u_2 = 0*0.3+0*0.2 = 0

$$f(u) = \frac{1}{1+e^{-u+b_k}}$$

$$f(u_1) = \frac{1}{1+e^{-0}} = 0.5$$

$$f(u_2) = \frac{1}{1+e^{-0}} = 0.5$$

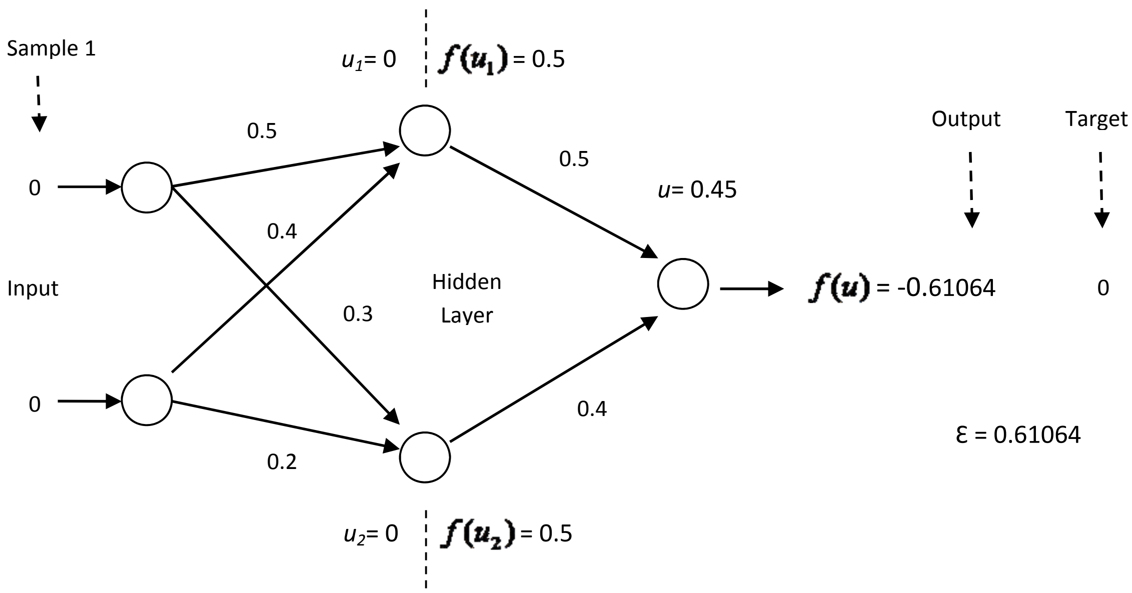


Figure 57: Forward calculation in multi-layer perceptrons (1.Iteration)

And the input of the output unit is calculated as following:

u = 0.5*0.5 + 0.5*0.4 = 0.45

The output of the network is:

$$f(u) = \frac{1}{1+e^{-0.45}} = -0.61064$$

Step 6: error calculation

The expected output was 0. So, the error of the network is Ɛ = 0-(-0.61064) = 0.61064

Step 7: back propagation

This error will be propagated backwards with correction of the weights. For this purpose, first the back propagated error of this layer is determined and weights are adjusted.

$$\delta_O = O_O(1-O_O)(T_O - O_O)$$

δ_O = -0.61064*(1-(-0.061064))*(0-(-0.61064)) = -0.14519

$$w_{jk}^{HO\ (t)} = \alpha w_{jk}^{HO\ (t-1)} + (\eta \delta_O O_H)$$

w_{11}^{HO} = 0.8*0.5 + 0.5*(-0.14519)*0.5 = 0.363704

w_{12}^{HO} = 0.8*0.4 + 0.5*(-0.14519)*0.5 = 0.283704

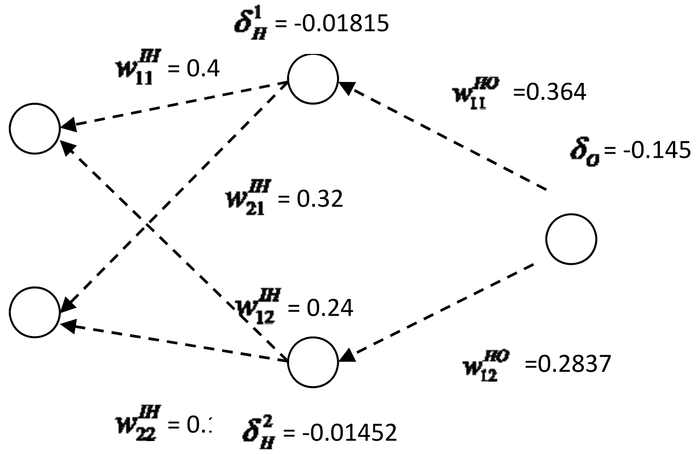


Figure 58: Training of system using back-propagation

Now, with the following formulas the back propagated error for input-hidden relation can be determined and the weight can be adjusted.

$$\delta_H = O_H(1 - O_H)\sum_O (\delta_O w_{jk}^{HO})$$

δ_H^1 = 0.5*0.5*[-0.14519*0.5] = -0.01815

δ_H^2 = 0.5*0.5*[-0.14519*0.4] = -0.01452

$$w_{ik}^{IH(t)} = \alpha w_{ik}^{IH\ (t-1)} + (\eta \delta_H O_I)$$

w_{11}^{IH} = 0.8*0.5 + 0.5*(-0.01815)*0 = 0.40

w_{12}^{IH} = 0.8*0.3 + 0.5*(-0.01815)*0 = 0.24

w_{21}^{IH} = 0.8*0.4 + 0.5*(-0.01452)*0 = 0.32

w_{22}^{IH} = 0.8*0.2 + 0.5*(-0.01452)*0 = 0.16

After the first iteration the new weight matrices are as following:

$$w_{IH} = \begin{bmatrix} 0.40 & 0.32 \\ 0.24 & 0.16 \end{bmatrix}$$

$$w_{HO} = \begin{bmatrix} 0.3637 & 0.2837 \end{bmatrix}$$

Each propagation of errors backwards and adjustment of the weights serves for minimizing the error. The training will continue until predefined iteration number or an acceptable error. As aforementioned, to keep the example more comprehensive only two hidden neuron is used. However, two hidden layer is generally not sufficient for multi-layer perceptron. The system cannot learn or take more time.

Appendix B – VBA Code used for neural network training

```
Option Explicit
'set number of iterations
    itera = Worksheets("main").Range("Z34").Value
'set parameters for the MLP Architecture
    b = Worksheets("main").Range("Z36").Value
    nu = Worksheets("main").Range("Z37").Value
    HiddenNeuron = Worksheets("main").Range("Z33").Value
    NumberofSamples = Worksheets("main").Range("Z30").Value
    NumberofFeatures = Worksheets("main").Range("Z31").Value
    NumberofTargets = Worksheets("main").Range("Z32").Value
'set variables
    ReDim input1(1 To NumberofSamples, 1 To NumberofFeatures)
    ReDim target1(1 To NumberofSamples, 1 To NumberofTargets)
    ReDim O(1 To NumberofSamples, 1 To NumberofTargets)
    ReDim w1(1 To NumberofFeatures, 1 To HiddenNeuron)
    ReDim w2(1 To HiddenNeuron, 1 To NumberofTargets)
    ReDim netout(1 To HiddenNeuron)
    ReDim E(1 To NumberofSamples, 1 To NumberofTargets)
    ReDim ER(1 To NumberofSamples)
    ReDim net(1 To HiddenNeuron)
    ReDim net2(1 To NumberofTargets)
    ReDim delta(1 To NumberofTargets)
    ReDim delta2(1 To HiddenNeuron)
    ReDim diff2(1 To HiddenNeuron)
    ReDim R(1 To HiddenNeuron)

'read input data
    For i = 1 To NumberofSamples
          For j = 1 To NumberofFeatures
                    input1(i, j) = Worksheets("input").Cells(i, j).Value
          Next j
    Next i

'read target data
    For i = 1 To NumberofSamples
          For j = 1 To NumberofTargets

target1(i, j) = Worksheets("target").Cells(i, j).Value
          Next j
Next i

'Assign weights randomly
      For m = 1 To NumberofFeatures
          For n = 1 To HiddenNeuron
                w1(m, n) = Rnd * 0.09
          Next n
    Next m
```

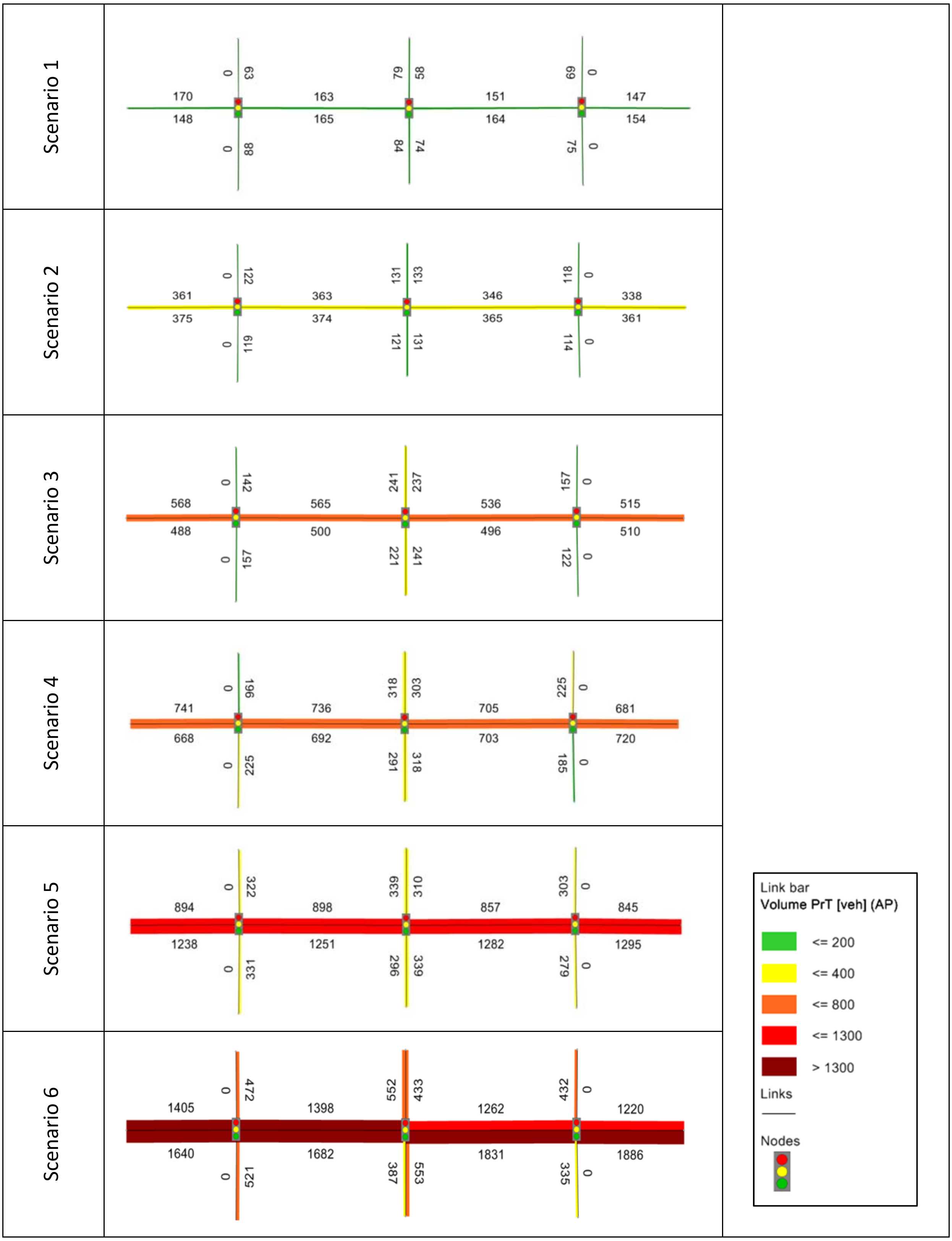
```
        For k = 1 To HiddenNeuron
          For l = 1 To NumberofTargets
               w2(k, l) = Rnd * 0.09
          Next l
     Next k
'number of iteration
       iter = 1
       While iter <= itera
'Or (alternatively) do until predefined error value
          'Do Until Worksheets ("main").Cells(5, 35).Value < 0.001
'Forward calculation
       For i = 1 To NumberofSamples
          For k = 1 To HiddenNeuron
               net(k) = 0
           For j = 1 To NumberofFeatures
               net(k) = net(k) + (input1(i, j) * w1(j, k))
           Next j
          Next k
          For k = 1 To HiddenNeuron
               R(k) = 1 / (1 + Exp(-(b + (net(k)))))
         Next k
       For m = 1 To NumberofTargets
           net2(m) = 0
                    For n = 1 To HiddenNeuron
                       net2(m) = net2(m) + (R(n) * w2(n, m))
           Next n
                       O(i, m) = 1 / (1 + Exp(-(b + net2(m))))
      Next m
'Back-propagation algorithm
'calculate delta
      For k = 1 To NumberofTargets
          delta(k) = O(i, k) * (1 - O(i, k)) * (target1(i, k) - O(i, k))
      Next k
      For k = 1 To HiddenNeuron
                    netout(k) = 0
          For l = 1 To NumberofTargets
                    netout(k) = netout(k) + delta(l) * w2(k, l)
          Next l
      Next k
      For k = 1 To HiddenNeuron
                 delta2(k) = R(k) * (1 - R(k)) * netout(k)
      Next k

'Calculate new weights
      For m = 1 To NumberofTargets
          For n = 1 To HiddenNeuron
                    w2(n, m) = w2(n, m) + nu * delta(m) * R(n)
          Next n
      Next m
```

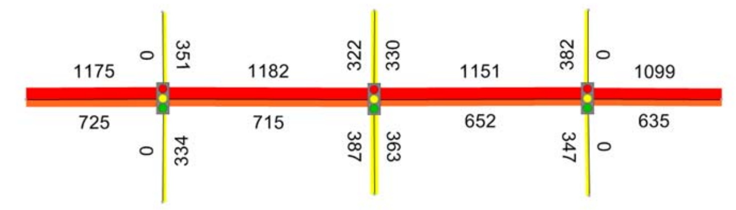
```
       For n = 1 To HiddenNeuron
            For m = 1 To NumberofFeatures
                    w1(m, n) = w1(m, n) + nu * delta2(n) * input1(i, m)
            Next m
       Next n
Next i

'Calculate error
    For i = 1 To NumberofSamples
            For n = 1 To NumberofTargets
                    E(i, n) = 0
                            E(i, n) = E(i, n) + (((O(i, n) - target1(i, n)) ^ 2) ^ 0.5)
            Next n
    Next i
    For i = 1 To NumberofSamples
                    ER(i) = 0
            For n = 1 To NumberofTargets
                    ER(i) = ER(i) + E(i, n) / NumberofTargets
            Next n
    Next i
    ERT = 0
            For i = 1 To NumberofSamples
                    ERT = ERT + ER(i)
            Next i
    ET = ERT / NumberofSamples
iter = iter + 1
Wend
MsgBox (ET)
'Write the outputs to MS Excel
    For i = 1 To NumberofSamples
            Worksheets("output").Cells(i, 8).Value = ER(i)
    Next i
            Worksheets("main").Range("AB30").Value = ET
    For i = 1 To NumberofSamples
            For j = 1 To NumberofTargets
                    Worksheets("output").Cells(i, j).Value = O(i, j)
            Next j
    Next i
              For m = 1 To NumberofFeatures
                    For n = 1 To HiddenNeuron
                            Worksheets("w1").Cells(m, n).Value = w1(m, n)
                    Next n
              Next m
            For m = 1 To HiddenNeuron
                    For n = 1 To NumberofTargets
                            Worksheets("w2").Cells(m, n).Value = w2(m, n)
                    Next n
            Next m
End Sub
```

Appendix C – Traffic scenarios for hypothetical arterial network

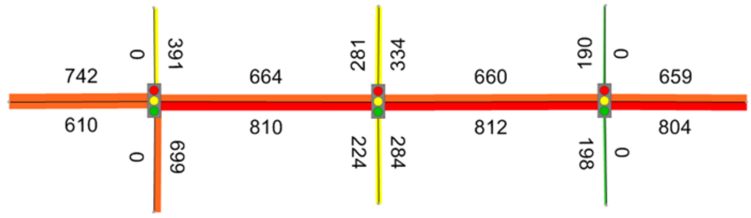


Scenario 7



Scenario 8

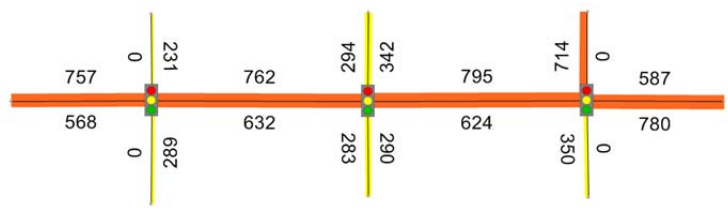
Scenario 9

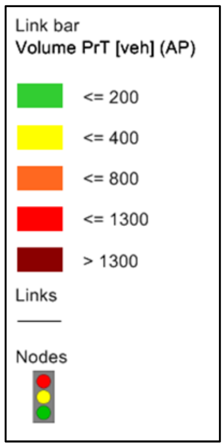


Scenario 10

Scenario 11

Scenario 12





Appendix D – Optimized timing plans for arterial networks

- **Optimized timing plans for corresponding scenarios for hypothetical three-intersection arterial using TRANSYT**

Timin Plan	Intersection	Cycle	Phase1	Phase2
1	1	80	43	19
	2	80	45	23
	3	80	51	19
2	1	39	20	7
	2	39	13	14
	3	39	22	7
3	1	46	21	7
	2	46	12	22
	3	46	25	11
4	1	77	42	17
	2	77	43	22
	3	77	46	21
5	1	39	24	7
	2	39	17	10
	3	39	22	7
6	1	60	31	11
	2	60	19	29
	3	60	39	11
7	1	81	26	37
	2	81	42	27
	3	81	54	17
8	1	130	76	36
	2	130	76	42
	3	130	87	33
9	1	53	23	12
	2	53	25	16
	3	53	19	24
10	1	105	61	26
	2	105	69	23
	3	105	68	27
11	1	44	19	7
	2	44	16	16
	3	44	27	7
12	1	41	16	7
	2	41	16	13
	3	41	23	8

- **Optimized timing plans for Fevzipasa Street and Sarigüzel Street using VISUM**

Intersection Nr.	Timing Plan Nr.	Cycle	Offset	Phase 1	Phase 2	Phase 3	Phase4
1133	1	90	28	46	43		
	2	90	30	38	51		
	3	120	119	36	83		
	4	60	50	24	35		
1134	1	90	4	89	11	9	
	2	60	0	48	11	17	
	3	90	23	89	15	44	
	4	60	42	59	8	14	
1135	1	60	26	59	17		
	2	60	6	59	29		
	3	60	25	59	30		
	4	60	27	59	15		
1136	1	60	56	42	17		
	2	60	38	32	27		
	3	60	0	33	26		
	4	60	49	37	22		
1137	1	120	9	42	29	34	14
	2	120	57	48	36	26	9
	3	120	17	57	26	26	10
	4	120	53	44	33	33	9
1158	1	60	20	26	33		
	2	60	14	37	22		
	3	60	54	37	22		
	4	60	0	59	24		
1159	1	90	16	89	23		
	2	60	19	41	18		
	3	90	2	89	26		
	4	60	11	59	16		
1256	1	120	0	119	54		
	2	90	79	89	45		
	3	90	2	89	44		
	4	120	56	89	52		

- **<u>Example TRANSYT optimization file</u>**

Stage Timings (TRANSYT 12 timings)

53s cycle time; 53 steps

Node ID	Number of stages	Stage 1	Stage 2
1	2	3	27
2	2	7	38
3	2	23	47

Stages

Node	Stage Index	Is Base Stage	Display ID	Links In This Stage	Stage Start (s)	Stage End (s)	User Stage Minimum (s)	Stage Minimum (s)	TRANSYT Stage Start (s)	TRANSYT Minimum Preceding Interstage (s)	TRANSYT Actual Preceding Interstage (s)	TRANSYT Stage Minimum (s)
1	1	Yes	?	3	15	27	0	7	3	12	12	19
1	2	Yes	?	1,2	33	3	0	7	27	6	6	13
2	1	Yes	?	10,5	13	38	0	6	7	6	6	12
2	2	Yes	?	7,8	44	7	0	7	38	6	6	13
3	1	Yes	32	11,9	28	47	0	7	23	5	5	12
3	2	Yes	31	12	52	23	0	7	47	5	5	12

Network Results

Run Summary

Results Upto Date	Analysis Set Used	Run Start Time	Run Finish Time	Modelling Start Time (HH:mm)	Cycle Time Used (s)	Total Network Delay (PCU-hr/hr)	Highest DOS (%)	Link With Highest DOS	Number Of Oversaturated Links	Percentage Of Oversaturated Links (%)	Network Within Capacity
Yes	Analysis Set 16 - CROSS4	13/12/2010 15:48:13	13/12/2010 15:48:14	2:00:00 AM	53	37.58	86	12	0	0	Yes

Network Results: Summary

Calculated Flow Into Link (PCU/hr)	Flow Discrepancy (PCU/hr)	Adjusted Flow Warning	Calculated Sat Flow (PCU/hr)	Degree Of Saturation (%)	DOS Threshold Exceeded	Effective Green (s (per cycle))	Unweighted Performance Index (£ per hr)	Performance Index (£ per hr)
11212	-1		N/A	86		222.00	705.69	705.69

Network Results: Stops And Delays

Mean Cruise Time Per PCU (s)	Mean Delay Per PCU (s)	Uniform Delay (PCU-hr/hr)	Random Plus Oversat Delay (PCU-hr/hr)	Unweighted Cost Of Delay (£ per hr)	Weighted Cost Of Delay (£ per hr)	Mean Stops Per PCU (%)	Unweighted Cost Of Stops (£ per hr)	Weighted Cost Of Stops (£ per hr)
14.61	12.06	26.11	11.47	533.33	533.33	69	172.37	172.37

Network Results: Queues And Blocking

Mean Max Queue (PCU)	Mean Max Queue EoTS (PCU)	Max Queue Storage (PCU)	Average Link Excess Queue (PCU)	Average Limit Excess Queue (PCU)	Excess Queue Penalty (£ per hr)	Max End Of Green Queue EoTS (PCU)	Max End Of Red Queue EoTS (PCU)	Wasted Time Starvation (s (per cycle))	Wasted Time Blocking Back (s (per cycle))	Wasted Time Total (s (per cycle))
N/A	N/A	575.30	N/A	N/A	0.00	N/A	N/A	6.00	0.00	6.00

Network Results: Fuel Consumption

Fuel Consumption Cruise (litres/hr)	Fuel Consumption Delay (litres/hr)	Fuel Consumption Stops (litres/hr)	Fuel Consumption Total (litres/hr)
99.75	43.17	78.68	221.63

Network Results: Journey Times

Distance Travelled (PCU-km/hr)	Time Spent (PCU-hr/hr)	Mean Journey Speed (kph)	Journey Time Per PCU (s)
1820.48	83.09	21.92	26.67

- **Example for ICA calculation in VISUM**

Node 101022:	
Control Type	Signalized
Method	HCM 2000
Average Delay	285.36
Average LOS	F
V/C	1.64
Loss Time	0

Volume and Adjustments by Movement							
Approach	N			E	S	W	
Movement	L1	T	R1	T		T	R1
Base Volume	239	223	209	550		682	457
PHF, Peak-hour factor	1.000	1.000	1.000	1.000		1.000	1.000
Peak 15 Volume	60	56	52	137		170	114
Adjusted Volume	239	223	209	550		682	457

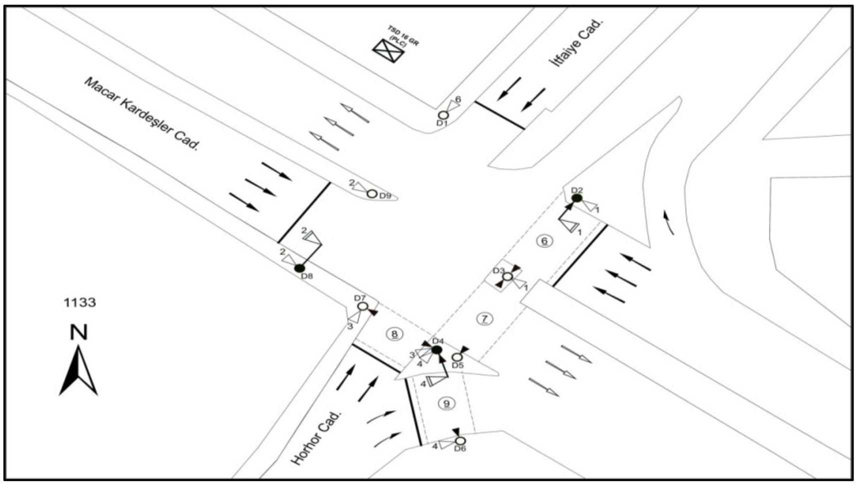
Saturation Flow Rate				
Approach	N	E	S	W
Lane Group	1335118	1335117		1335119
Control Type	Split	No Left		No Left
so, Base Saturation Flow Rate	1400	1400		1400
N, Number of Lanes	1	1		1
fw, Lane Width Adjustment	0.928	0.928		0.928
Phv, % Heavy Vehicles	0	0		0
fHV, HV Adjustment	1.000	1.000		1.000
fg, Grade Adjustment	1.000	1.000		1.000
fp, Parking Adjustment	1.000	1.000		1.000
fbb, Bus Blocking Adjustment	1.000	1.000		1.000
fa, Area Type Adjustment	0.900	0.900		0.900
fLU, Lane Utilization Adjustment	1.000	1.000		1.000
fLT, Left Turn Adjustment	0.983	1.000		1.000
fRT, Right Turn Adjustment	0.958	1.000		0.946
fLpb, Left Turn Ped. Adjustment	1.000	1.000		1.000
fRpb, Right Turn Ped. Adjustment	1.000	1.000		1.000
Has Short Lane				
s, Saturation Flow Rate	1101	1169		1106

Capacity, Control Delay, and Level of Service Determination				
Approach	N	E	S	W
Lane Group	1335118	1335117		1335119
Control Type	Split	No Left		No Left
V, Volume	671.13	549.87		1138.54
s, Saturation Flow Rate	1101	1169		1106
c, Capacity	385	702		664
g/C, Green / Cycle	0.35	0.6		0.6
X, Volume / Capacity	1.74	0.78		1.72
d1, Uniform Delay	39.00	18.12		24.00
k, Delay Calibration	0.50	0.50		0.50
d2, Incremental Delay,	344.61	8.55		328.39
d3, Initial Queue Delay	0.00	0.00		0.00
Rp, Platoon Ratio	1.00	1.00		1.00
P, Proportion Arriving on Green	0.35	0.60		0.60
PF, Progression Factor	1.00	1.00		1.00
d, Delay	383.61	26.67		352.39
LOS	F	C		F
dA, Approach Delay	383.61	26.67	0.00	352.39
Approach LOS	F	C	A	F
dI, Intersection Delay	285.36			
Intersection LOS	F			

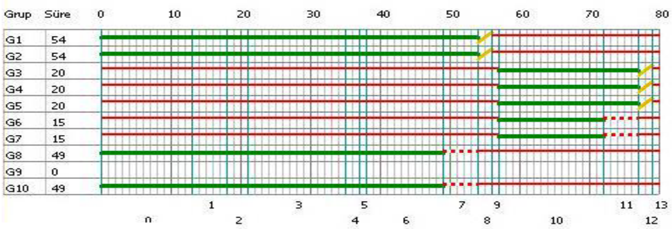
Signal Timing Information				
Approach	N	E	S	W
Lane Group	1335118	1335117		1335119
Control Type	Split	No Left		No Left
Signal Type	P	P		P
C, Cycle Length	120.00	120.00		120.00
Gp, Minimum Pedestrian Timing	0	0		0
L, Total Loss Time per Cycle	0.00	0.00		0.00
l1, Start Up Loss Time	2.00	2.00		2.00
l2, Clearance Loss Time	0.00	0.00		0.00
Green Time Start	78	3		3
Green Time End	120	75		75
G, Actual Green Time	42	72		72
gi, Effective Green Time	42	72		72
g/C, Green / Cycle	0.35	0.6		0.6
(v/s)i, Volume / Saturation Flow Rate	0.61	0.47		1.03
Critical Lane Group	Y			Y

Appendix E – Intersection layouts and timing plans of study intersections

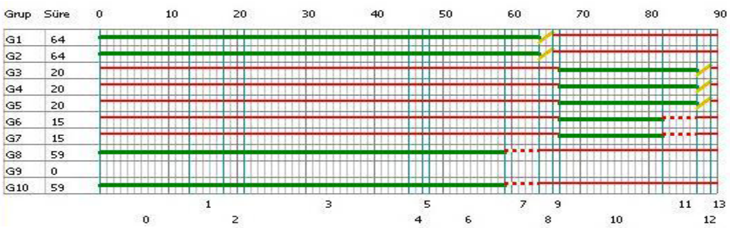
- **<u>Intersection Nr: 1133</u>**



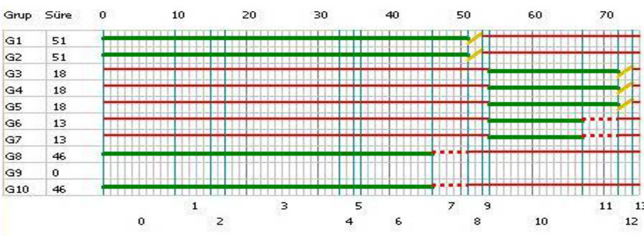
a) 07:00-10:00



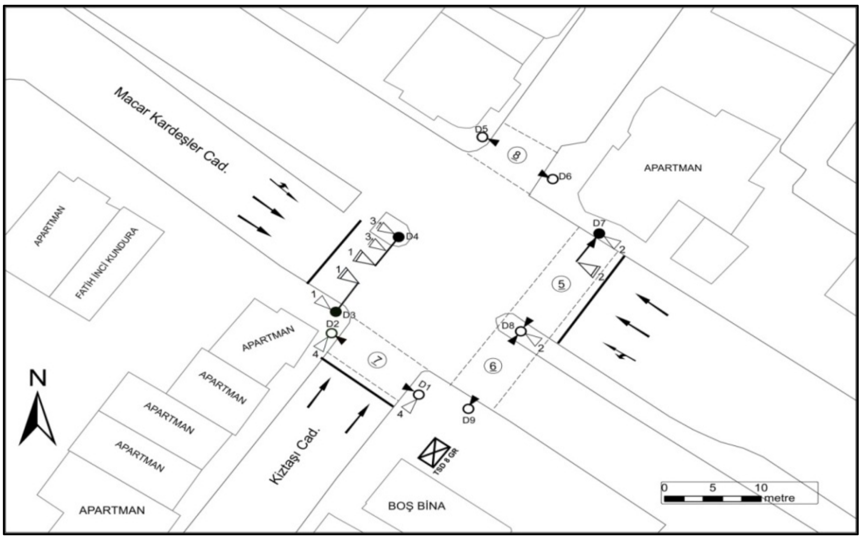
b) 10:00-16:30



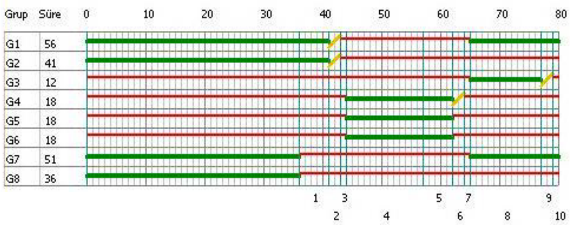
c) 20:30-24:00



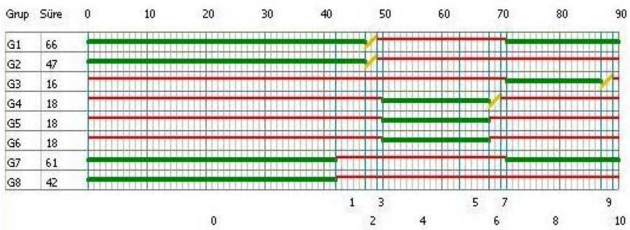
- **Intersection Nr. 1134**



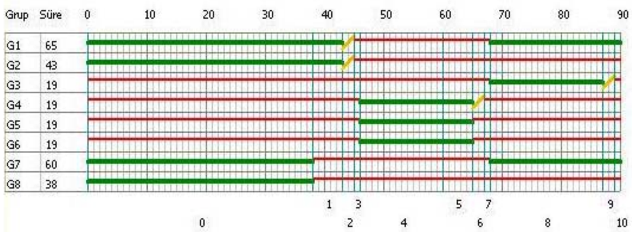
a) 07:00-10:00



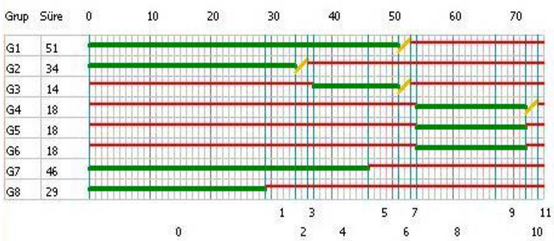
b) 10:00-16:30



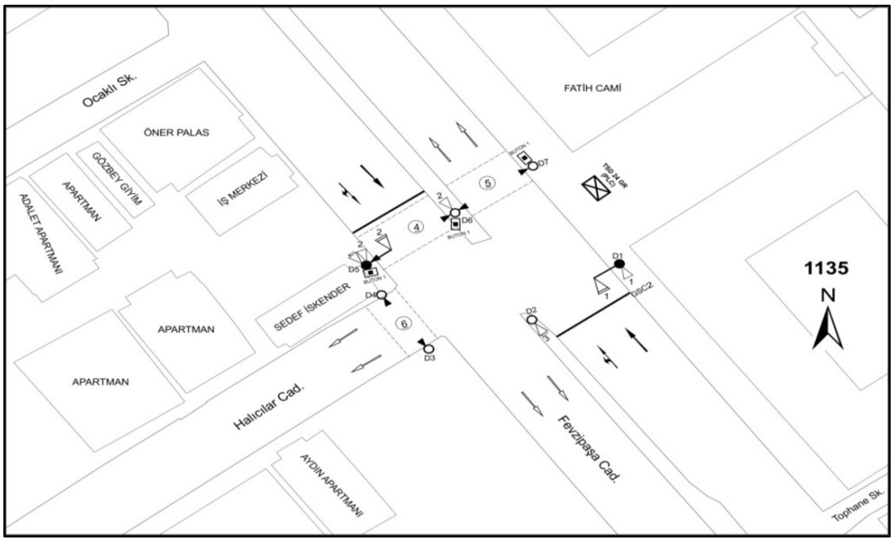
c) 16:30-20:30



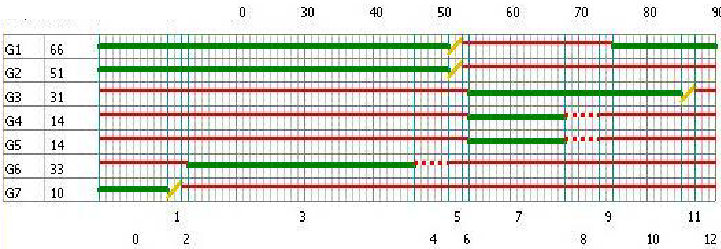
d) 20:30-24:00



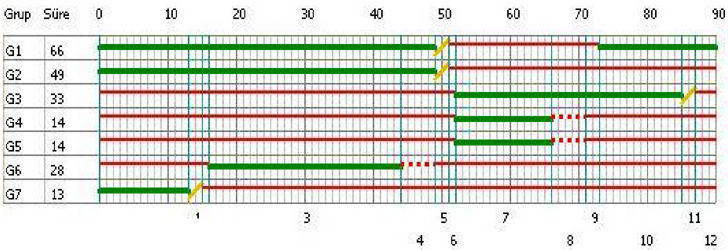
- **Intersection Nr. 1135**



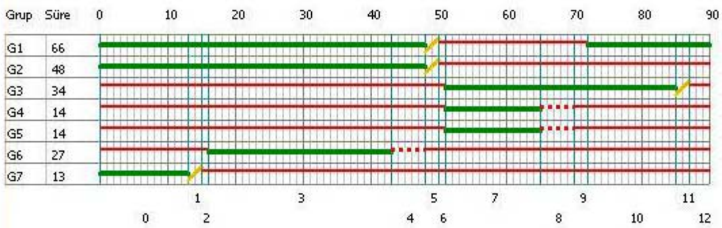
a) 07:00-10:00



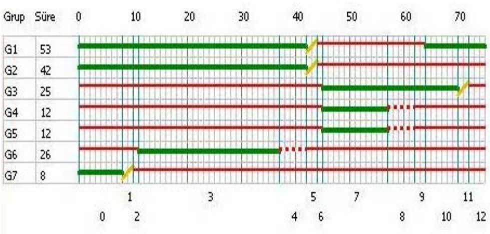
b) 10:00-16:30



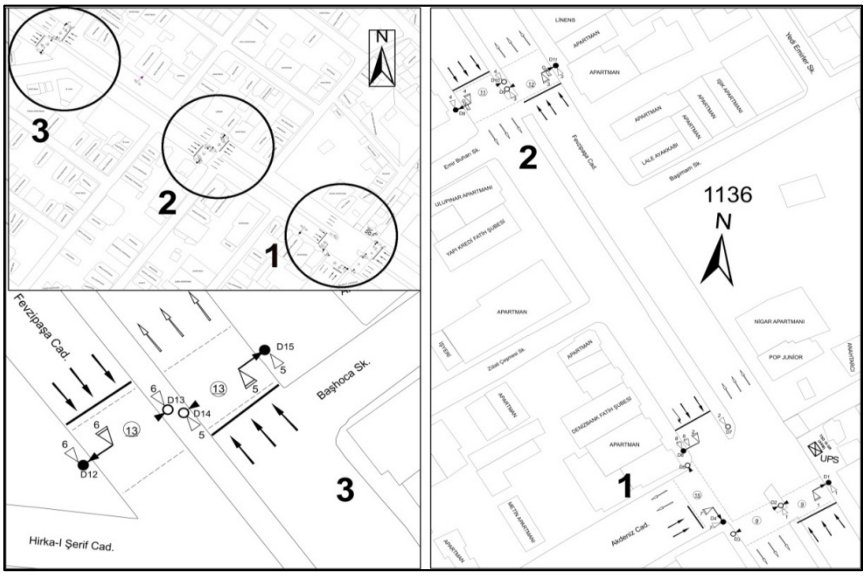
c) 16:30-20:30



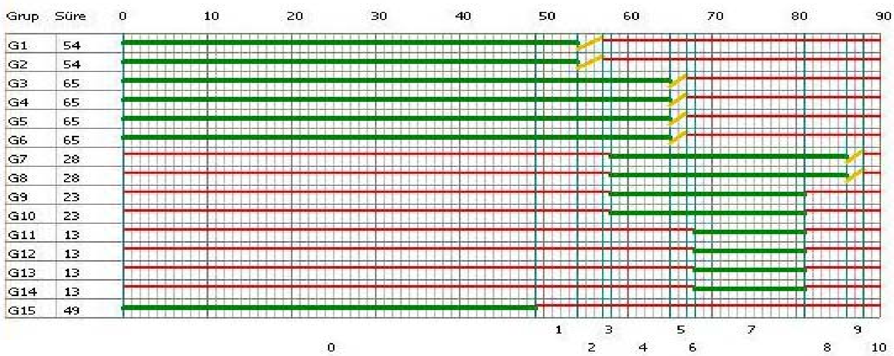
d) 20:30-24:00



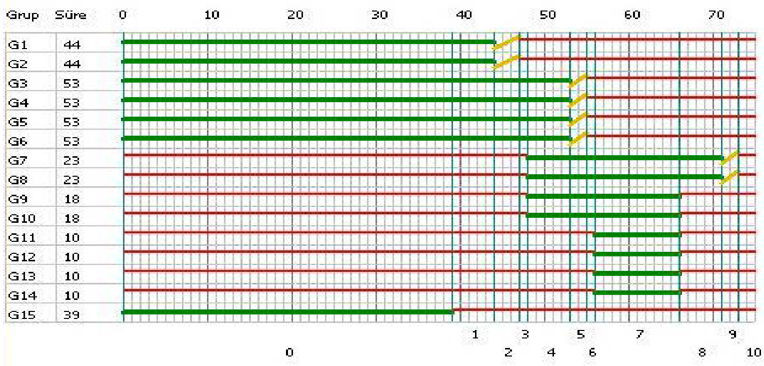
- **Intersection Nr. 1136**



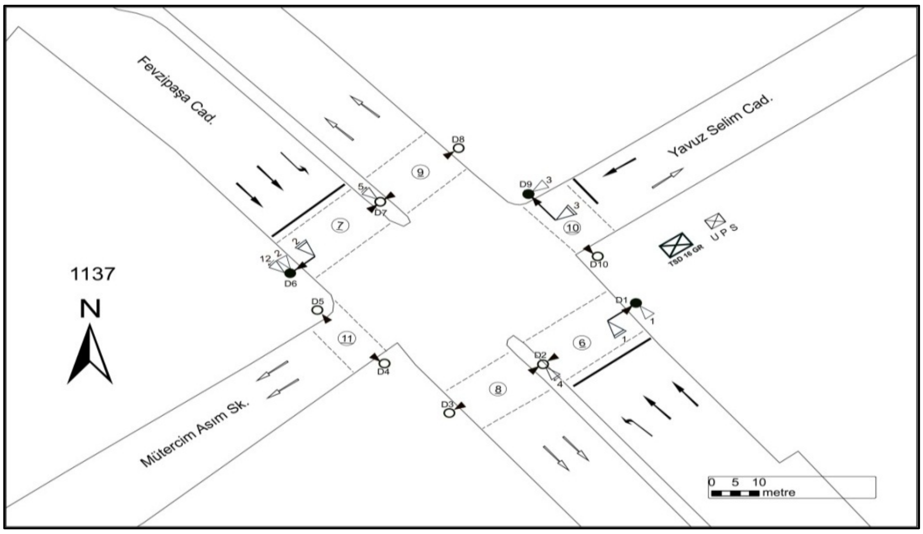
a) 07:00-20:30



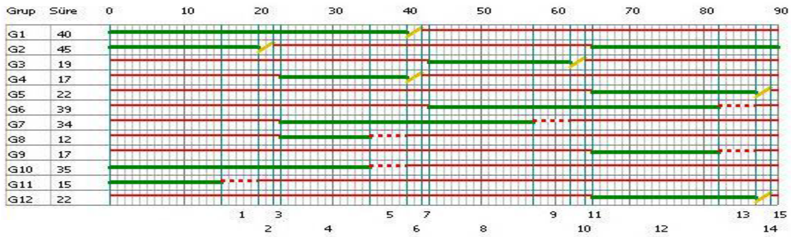
b) 20:30-24:00



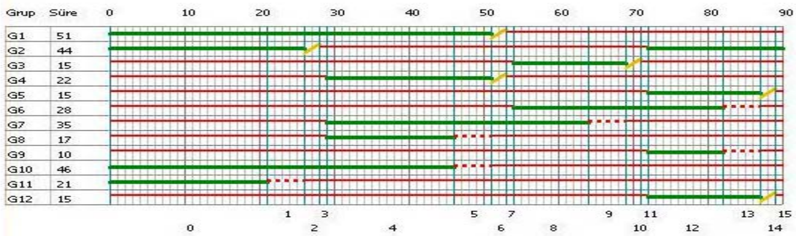
- **Intersection Nr.1137**



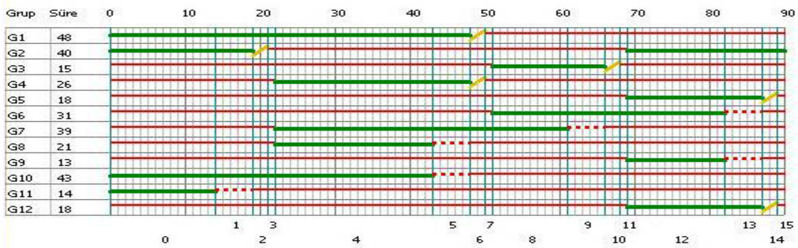
a) 07:00-10:00



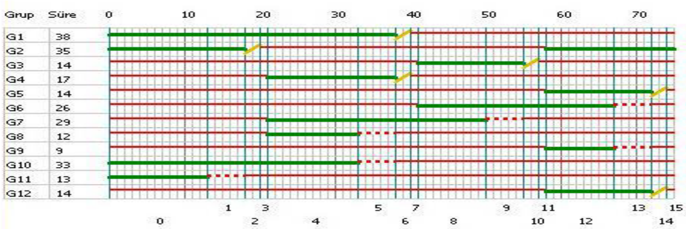
b) 10:00-16:30



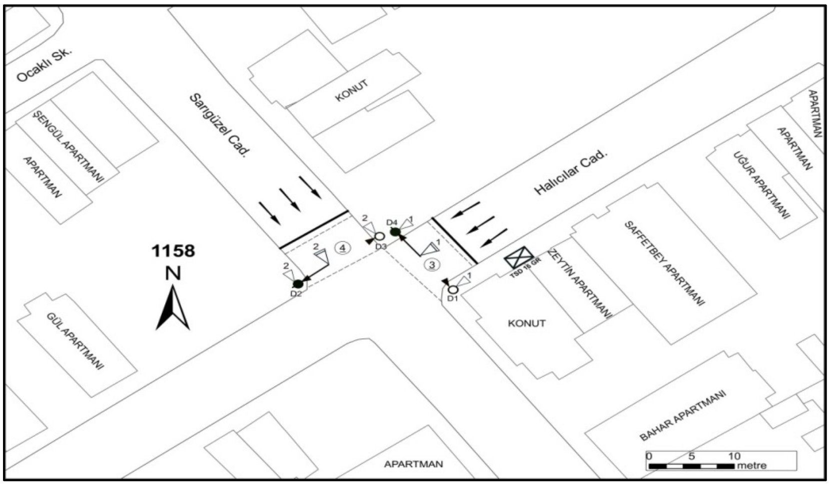
c) 16:30-20:30



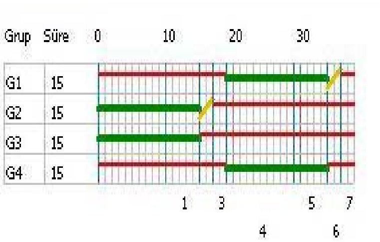
d) 20:30-24:00



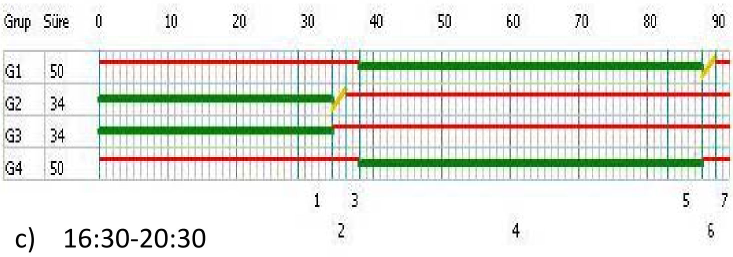
- **Intersection Nr. 1158**



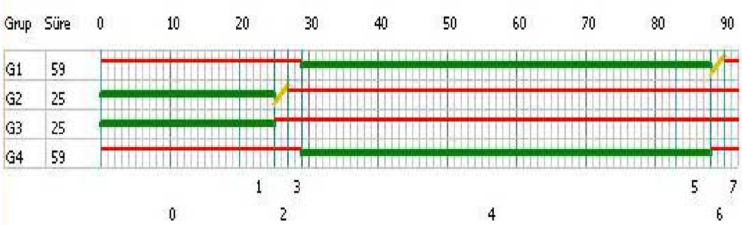
a) 07:00-10:00



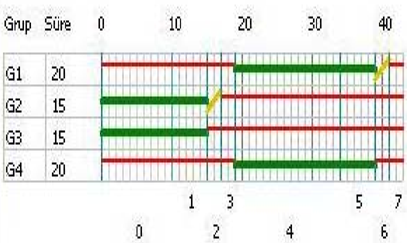
b) 10:00-16:30



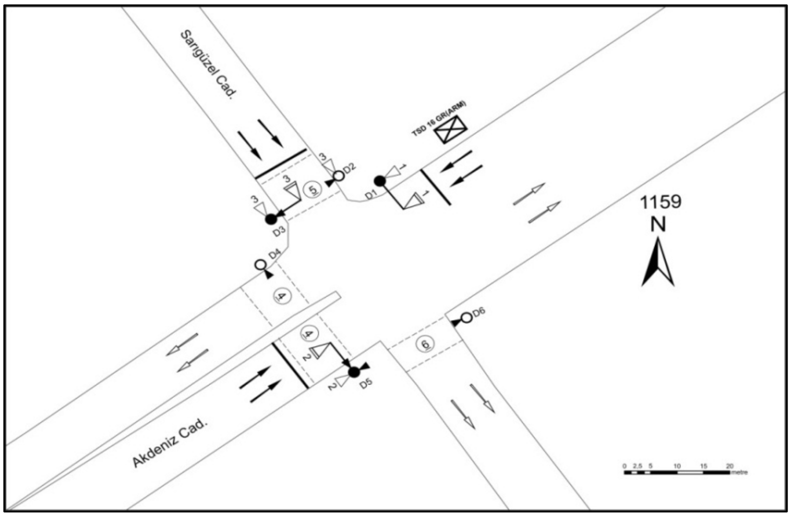
c) 16:30-20:30



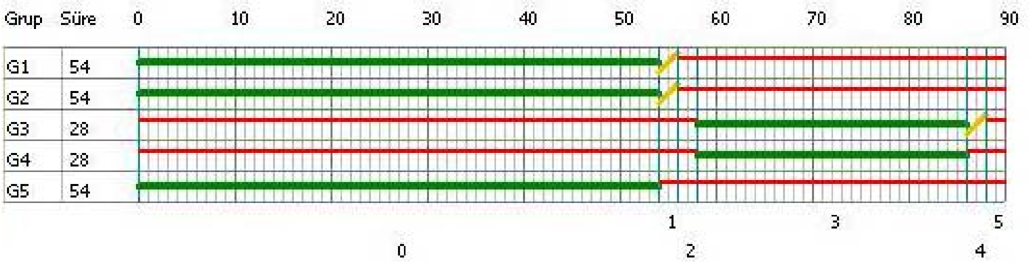
d) 20:30-24:00



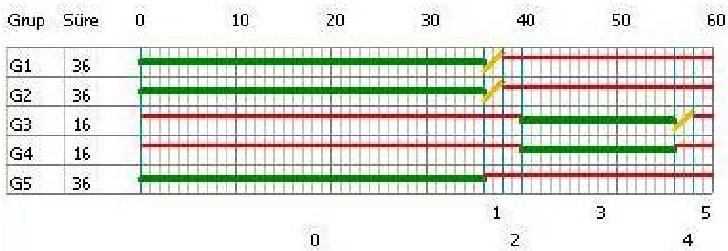
- **<u>Intersection Nr. 1159</u>**



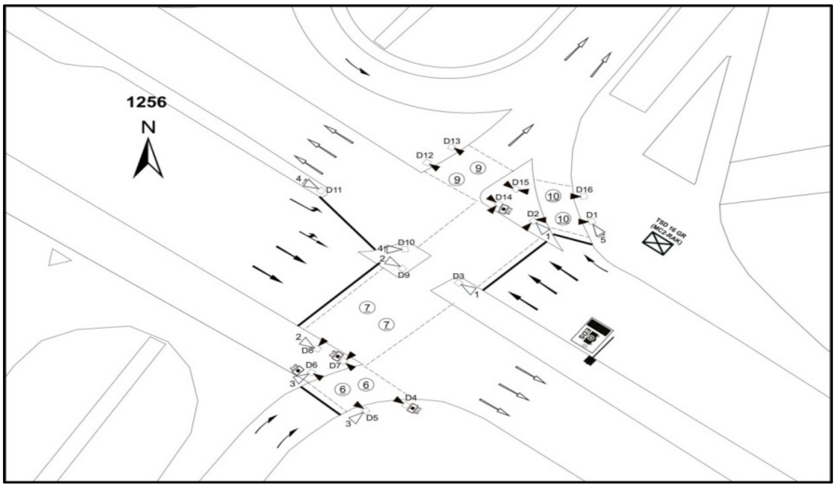
a) 07:00-20:30



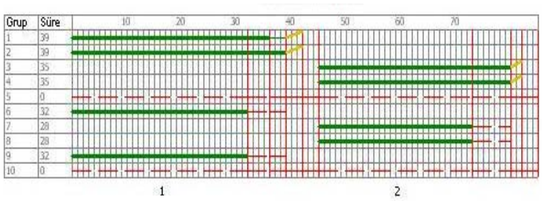
b) 20:30-24:00



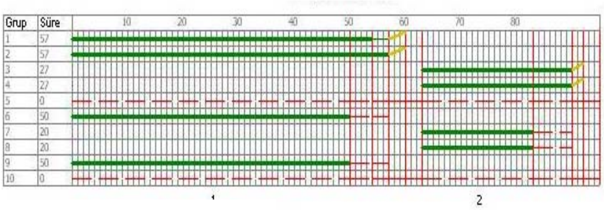
- **Intersection Nr.1256**



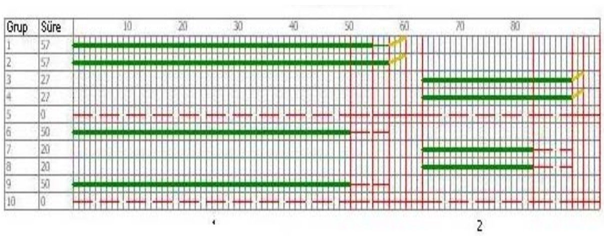
a) 07:00-10:00



b) 10:00-16:30



c) 16:30-20:30



d) 20:30-24:00