Urban Traffic Management

Martin Fellendorf
Graz University of Technology, Graz, Austria

1 INTRODUCTION

1.1 Overview

While traffic management in general is considered to include transport planning issues for a sustainable future, urban traffic management is limited to traffic operations of road- and rail-based public transport networks within city boundaries. Traffic signal control is the most widely and most effective operational measure to control individual vehicle movements and guide traffic streams in dense urban networks by limiting capacity by green periods. Signal control measures can become operational by quick adaptation of green periods. While traffic signal control is usually understood to organize traffic flow between conflicting traffic movements at junctions, traffic signals are also applied at contraflow systems to allocated lanes according to traffic demand. While the majority of contraflow systems are implemented at highways, some can also be seen at urban areas. Parking guidance systems and contraflow systems aim to control traffic according to more general transport-related objectives (Figure 1).

1.2 Objectives of traffic signal control

At-grade intersections are the most complex locations within a road network. Vehicular and nonmotorized traffic movements have to cross or merge with numerous potential conflicts. Traffic engineers have to manage these conflicts to ensure safety and provide efficient traffic flow at the same time. While many other traffic management measures provide information to the driver, traffic signal control has to be obeyed. Because traffic signals are binding rules, control strategies have an instant impact on traffic flow. The first traffic signal was applied in London in 1868 by J.P. Knight with a revolving light. Between 1913 and 1920, several mechanical and electrical devices were patented in the United States to control the sequence of conflicting traffic streams. Quickly, in Salt Lake City and Houston, multiple signals were installed, which were interconnected and simultaneously controlled from a manual switch. Moreover, in the 1920s, several European cities adopted traffic lights such as Berlin at Potsdamer Platz. Thus, the principle of PROCEED at GREEN and STOP at RED exists since about 100 years. However, traffic signals are subjected to constant change because of advancements in control technology and transport requirements.

Primarily, traffic signals are built to improve traffic safety.

- Conflicting movements are not allowed to pass a junction at the same time.
- The frequency and severity of certain types of crashes, especially right angle collisions, can be reduced.

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition)
ISBN: 978-0-470-97402-5

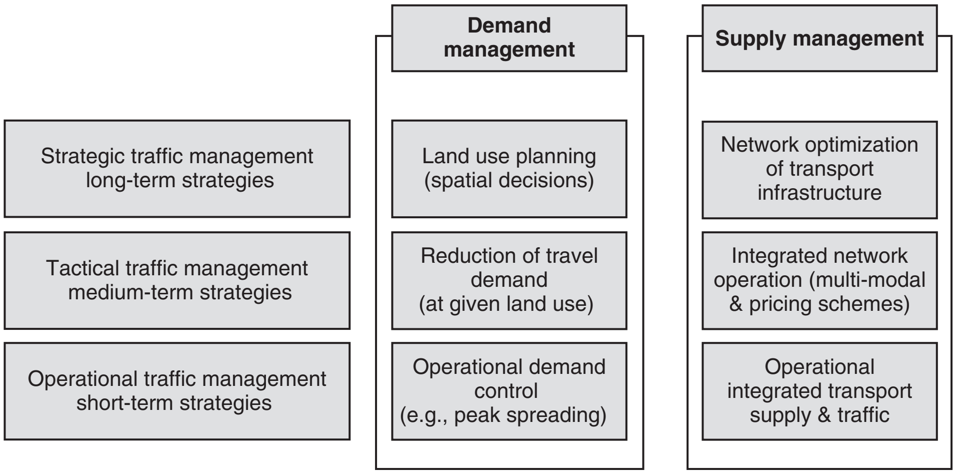


Figure 1. Classification of traffic management measures.

- Vulnerable road users (pedestrians and bicyclists) have a chance to cross roads safely.

Secondly, traffic signals are applied to improve road capacity and control movements according to politically accepted objectives:

- Roads with heavy traffic flow are interrupted to permit other traffic to turn or cross.
- Coordinated traffic signals along an arterial provide continuous movement at a given design speed along a designated corridor at undersaturated traffic conditions.
- Certain vehicle types such as buses and trams are prioritized to minimize delay of public transport users.

Traffic signals are controlled either at fixed time intervals (pre-timed signals) or with flexible time intervals according to traffic flow (actuated and adaptive signal control). While current traffic conditions will not influence the control patterns in the first case, signal durations may respond to traffic flow in the latter as part of a control circuit.

2 FIXED TIME SIGNAL CONTROL

2.1 Signal control terminology

Owing to different historical developments, traffic signal control hardware differs greatly between the North American market (mainly the United States, Canada, and Mexico) and Europe. This chapter follows standards defined in Europe with major recognition of the German standards RiLSA (FGSV, 2010), the Austrian RVS (FSV, 1998), and the British Technical Advice Notes (DoT, 2006). For specific US notation, please refer to the Highway Capacity Manual Chapter 18 (TRB, 2010) or Roess, Prassas, and McShane (2010). The terminology will be explained using a simple four-leg junction (Figure 2).

The major traffic runs East–West with a minor flow South–North. There is a separate left turn movement from the eastern direction, whereas a left turn from the western direction is prohibited. There are two through lanes from West to East because of high traffic volumes. All turning movements are permitted from the minor approaches. Pedestrians may cross on the western approach. Other pedestrian signalized crossings are omitted for simplicity. The scaled signal layout map contains relevant information for the traffic engineer. The road boundaries, islands, and lane markings allocate the available space for vehicular movements. The type and location of signal heads and poles indicate the complexity of control devices. For safety reasons, each signal display will be repeated using a second identical signal head. The signal layout map may also include boundaries of buildings and foliage to provide an indication of visibility. Furthermore, traffic signs, speed limits, and gradient may be added as supplementary information that is relevant to calculate performance measures.

Each signal head contains a fixed number of aspects determined by the national road regulations. Most likely vehicular signals contain three aspects: GREEN movement permitted; RED movement prohibited; AMBER transition period from GREEN to RED. The transition period has a fixed duration (commonly 3 s). Once AMBER is visible, the driver has to decide either to stop safely before the stop line or pass the stop line during AMBER. If neither one is possible, the AMBER period should be extended, which is done in some countries if the legal approach speed exceeds 50 km/h. Some countries use a 1 or 2 s

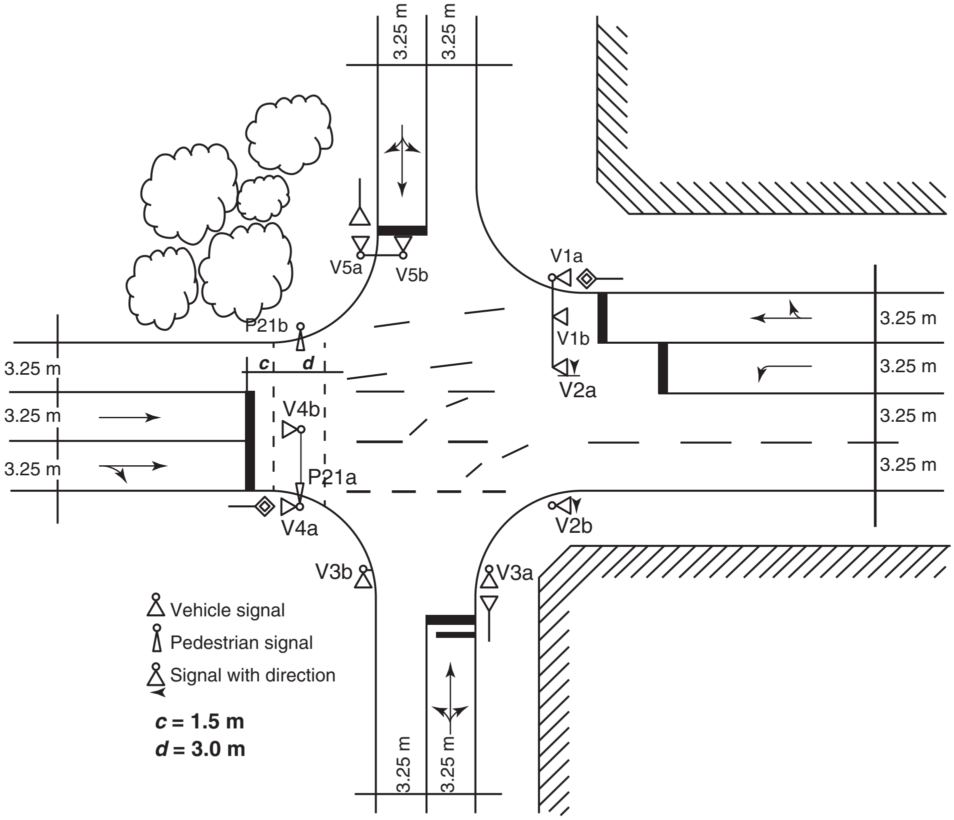


Figure 2. Signal layout map of a four-leg signalized junction with five vehicular and one pedestrian signal group.

RED-AMBER transition period between ending RED and starting GREEN as a preparation time. Commonly, there is no transition period for pedestrian because of instantaneous start or stop of walking. Some countries use FLASHING GREEN to indicate a clearance period without the right to newly enter the junction. Each signal has a fixed sequence of aspects.

The smallest control unit is called a *signal group*, which joins all signal heads with identical display at any given time. A stage contains a set of compatible signal groups. Signal groups are compatible if their movements can be released at the same time. For example, signal group V1 with signal heads V1a and V1b and signal group V3 may receive GREEN simultaneously during the same stage but not necessarily at identical times. Variations at beginning or ending green within one stage are produced by different intergreen times. Stages may also include semi-compatible movements; thus, movements with a conflict but a defined priority that has not to be signal controlled. For example, V2 and V4 may run simultaneously but the left turns have to obey the right-of-way of opposing through movements. In some instances, pedestrian movement P21 will also be activated in stage 3 with semi-compatible movements of turning movements. This will be an issue of less safety versus less delay time for the pedestrians. A stage sequence contains the periodic order of stages applied within a complete cycle (Figure 3).

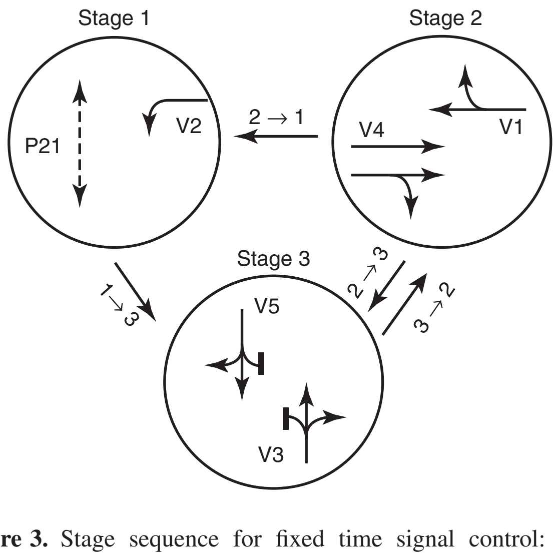


Figure 3. Stage sequence for fixed time signal control: stage 2–stage 1–stage 3–stage 2.

The signal timing plan contains information about the duration of one full cycle and the distribution of GREEN of each signal group. Stages, interstages, and intergreens can also be retained from a signal timing plan. (Figure 4) displays a signal timing plan with a 70 s cycle for the six signal groups of figure 2.

The transition from one stage to the next stage is called *interstage*. Each interstage starts with the first ending GREEN of the ending stage (e.g., V4 in stage 2) and the last beginning GREEN (e.g., P21 in stage 1) of the starting stage. The interstage period contains the intergreen times of conflicting movements. Intergreen prevents a conflict between the first vehicle of a starting signal and the last vehicle of an ending signal as long as both vehicles pass at their respective green signals.

2.2 Intergreen calculation

The minimization of potential conflicts between noncompatible movements is the major idea of traffic signals to improve traffic safety. Therefore, the identification of conflicts between noncompatible movements and the calculation of the intergreen times must be carried out with great care, especially in countries with irregular junctions of varying geometry. The intergreen time must be calculated for each pair of conflicting streams (Figure 5a) and the maximum value of the intergreen must be taken for each pair of noncompatible signal groups.

$$t_z = t_c + t_a - t_e \qquad (1)$$

with

t_z	intergreen time in (s) rounded to the next full second
t_c	clearance time calculated by clearance distance plus vehicle length divided by the slowest expected vehicle speed
t_a	transition period, usually the AMBER time between end of green and beginning of red
t_e	entering time calculated by entering distance divided by approach speed

Especially, the value taken for the clearance and entering speed varies between national guidelines, the calculated intergreen times is location dependent; for example, clearance speed of pedestrians may vary from 0.8 m/s at crossings near elderly homes up to 1.5 m/s. As the clearance speed of pedestrians is generally low, the intergreen time between clearing pedestrians and nearby car movements is long (Figure 5b, last row). In some countries such as the United States, the clearance time for pedestrians is indicated by a flashing signal allowing pedestrians to clear but not starting to cross the road. In general, intergreen times increase with junction size. Especially, long pedestrian crossings will lead to long intergreen times. In order to limit intergreen times and provide a safe area for pedestrians, islands are installed at large intersections. This may not be the most convenient for pedestrians, though.

The intergreen matrix is safety relevant. It is hard-wired within the signal controller to minimize a potential software error. The intergreen matrix has to be changed if stages

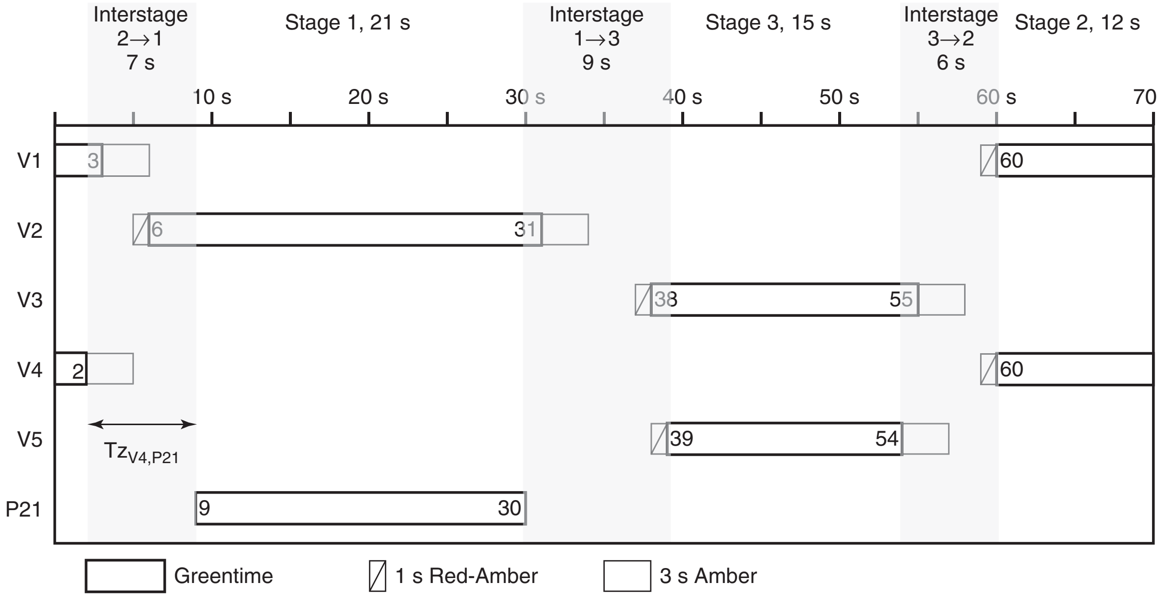
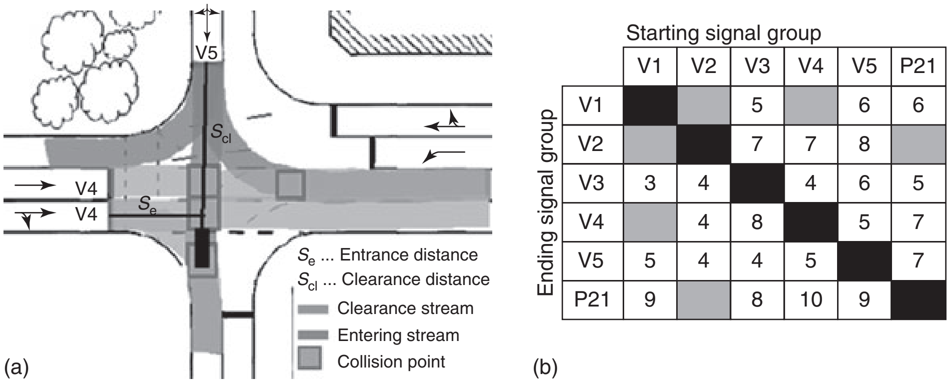


Figure 4. Signal timing plan with 70 s cycle length and three stages; stages, interstages and one intergreen are marked.



Starting signal group

Ending signal group	V1	V2	V3	V4	V5	P21
V1	■	▨	5	▨	6	6
V2	▨	■	7	7	8	▨
V3	3	4	■	4	6	5
V4	▨	4	8	■	5	7
V5	5	4	4	5	■	7
P21	9	▨	8	10	9	■

Figure 5. Conflict between clearance of V5 and entering of V4 (a) and intergreen matrix (b); example also includes intergreens of semi-compatible movements V2 and V4.

will be assembled differently. If the above example will be changed from a three-stage system to a two-stage system, then P21 would run parallel to V3 and V5 and the associated conflicts will be eliminated; respectively, V2 would operate simultaneously with V1 and V4 as permitted left turn.

In the United States and some other countries, intergreen calculation is replaced by an all-red transition period. Nonstandard junctions and individual signals for bicyclists, buses, and trams require a more detailed intergreen calculation, which is a common practice in Europe.

2.3 Cycle length and green splits

According to Roess, Prassas, and McShane (2010), a signal cycle is defined as a rotation through all indications provided. Its length is defined as the period between the same sequence of indications restarts (Figure 4). The cycle length significantly affects the quality of traffic flow. As the fixed sum of the intergreen periods does not provide intersection capacity, the throughput of vehicles will increase as cycle time increases. However, there are optimal cycle lengths that will minimize the average vehicle delay. The minimum cycle length is defined by the sum of the lost times and the required green time of each approach. According to Webster and Cobbe (1966), the delay minimizing cycle time can be estimated by

$$t_{\mathrm{c,opt}} = \frac{1.5 \cdot L + 5}{1 - \sum_i q_i / q_{\mathrm{s},i}} \tag{2}$$

with

$t_{\mathrm{c,opt}}$ optimal cycle time in (s)

L total lost time, which is typically the sum of the intergreens between the critical movements of each stage

q_{i} volume in (veh/h) of the critical movement within stage i

$q_{\mathrm{s},i}$ saturation flow for the critical movement within stage i

Among other national guidelines, the German RILSA has adopted this estimation with minor adjustments. The estimation is not necessarily the true optimum, as constraints such as minimum green times, pedestrian flow, and single public transport vehicles are not considered. At minimum cycle time and long cycle times, the average delay increases (Figure 6). As the risk of red-running of pedestrians increases at long delays, the selected cycle time $t_{\mathrm{c,select}}$ should range between 60 and 90 s. In Europe, cycle times of more than 120 s are not recommended, whereas cycles of up to 4 min can be observed at large signalized junctions in Asian megacities.

After fixing the duration of a cycle, the green duration of the critical movement of each stage is computed. The critical movement of each stage is defined by the highest ratio of volume divided by saturation flow (*v*/*c* or volume over capacity ratio in the United States). This ratio is computed for each signal group and can be interpreted as arrival rate (volume q) and departure rate (saturation flow q_s) at a queuing process. The saturation flow ranges roughly around 1800–2000 veh/h/lane. Vehicle features such as vehicle size, local driving habits, and green duration and geometrical features such as visibility, lane width, gradient, and curvature influence the saturation value. National guidelines provide a wide variety of computational definitions ranging from bulk park numbers for saturation flows to

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

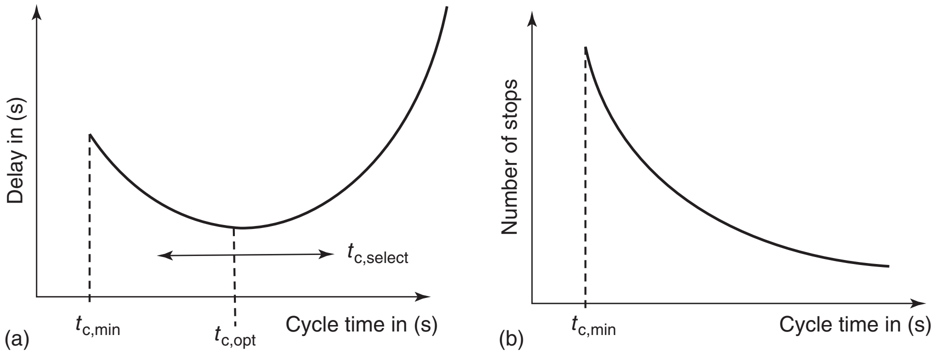


Figure 6. selected cycle length and average vehicle delay (a); (b) number of stops decrease with cycle length.

precise multinomial regression formulae. The green duration $t_{g,j}$ of each signal group j should satisfy:

$$t_{g,j} \geq \frac{t_{c,\text{select}} - L}{\sum \frac{q_i}{q_{S,i}}} \cdot \frac{q_j}{q_{S,j}} \tag{3}$$

After all green splits are computed, a signal timing plan (Figure 4) can be assembled, which has to satisfy the minimum green splits, the intergreens given in the intergreen matrix, and minimum green times.

The quality of traffic flow at a signalized junction is measured by performance indicators such as average vehicle delay, stopped delay, total person delay, maximum volume capacity ratio, queue storage ratio, and number of stops. These values listed may be classified by cars, public transport vehicles, pedestrians, and bicyclists. Furthermore, the quality of signalization is measured by a level-of-service (LOS), which is based on one or more of the above performance measures with delay of the highest significance.

At traffic signals, control delay is typically taken as base performance value. It is equal to the time-in-queue delay plus the acceleration–deceleration delay component needed to cross a junction compared to free flow travel. All analytical delay models consist of several terms, the first being uniform delay based on an assumption of uniform arrivals and stable flow with no individual cycle failures. Random delay is the additional delay, above and beyond uniform delay, because traffic flow is randomly distributed rather than uniform at isolated intersections. Overflow delay is the additional delay that occurs when the capacity of a green period is less than the arrival flow rate within this cycle. On the basis of early numerical simulations, Webster and Cobbe (1966) came up with a good delay estimation that contains the base uniform delay plus a term reflecting the stochastic effect of vehicle arrivals including also the probability of overflow arrivals. However, in the early delay formulae, delay increased to infinity in oversaturated situation. Various authors did measurements indicating that queues vanish after some time as demand will reduce after the peak period. This is considered in the following delay formula (RiLSA, 2010), which includes a term for the base delay and a second one to reflect spill back delay.

$$d_i = \frac{t_{c,\text{select}}(1 - t_{g,i}/t_{c,\text{select}})^2}{2(1 - q_i/q_{s,i})} + \frac{3600 \times N_{GE,i}}{q_{s,i} \cdot t_{g,i}/t_{c,\text{select}}} \tag{4}$$

with

d_i delay of signal group i in (s)

$N_{GE,i}$ average number of vehicles at signal group i, which remain queuing at end of GREEN

The calculation of N_{GE} depends on the degree of saturation and the number of cycles being considered. N_{GE} will be 0 if the degree of saturation ($q_i/q_{s,i}$) is less than 65%. The likelihood of a spill back at end of green increases heavily as the degree of saturation reaches 100%. However, the delay formula also operates in over-congested cases because experience shows that demand exceeds saturation flow only for a limited number of cycles. Therefore, the queue dissipates after a number of cycles. This dynamic effect is reflected in the value of N_{GE}, which is based on a variety of empirically based regression formulae shown in detail in the RiLSA (2010).

2.4 Coordination of traffic signals

Traffic signals are rarely operated independently from adjacent signalized junctions. At corridors with a sequence of signalized junctions (arterials), traffic signals should be coordinated so that the majority of vehicles traveling the main direction will not have to stop. The time–space diagram (Figure 7) depicts three junctions and the trajectories of several vehicles traveling from junction 1 to junction

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

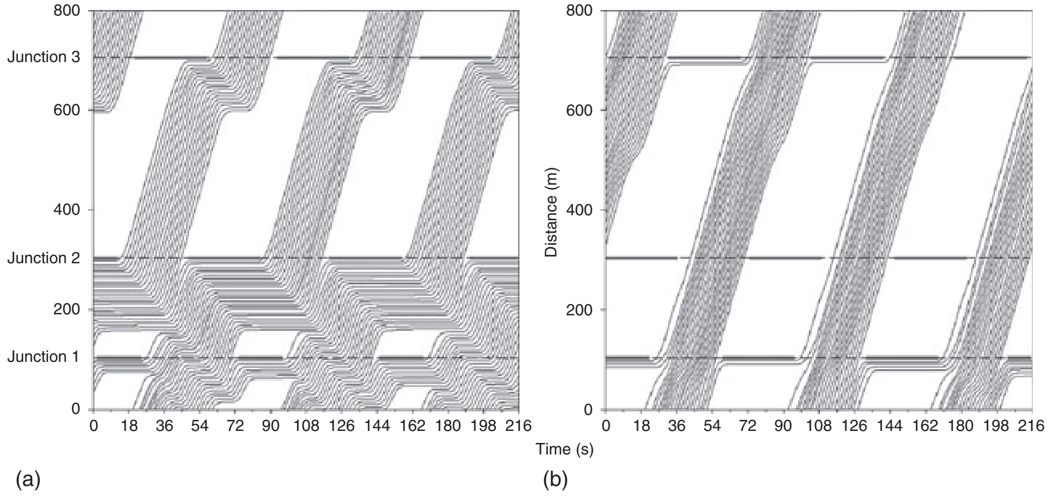


Figure 7. Trajectories in a time-space diagram for badly (a) and well (b) coordinated signals.

3. The vehicle highlighted in the left diagram has to stop three times, once because of a spill back from junction 2 across 1. As the green period is extended at junction 1 and the offset shifted at junction 2, very few vehicles have to stop.

Traffic engineers use offset variation as a standard method to provide coordinated traffic flow at arterials. However, coordination is not that easy if both traveling directions have to be considered. There is a geometrical dependency between cycle time, progression speed, and distance between junctions. As distance is fixed, the traffic engineer can only change the speed and the cycle time. The progression speed is usually slightly lower than the legal speed. Lower values will not be accepted by the vehicle drivers unless the legal speed is lowered with accompanied road design measures. Thus, cycle time variation is most common to improve coordination in both directions. The cycle time may differ significantly from the optimal cycle time according to Equation 2.

Constraints such as extensive traffic crossing from side streets, short cycles to minimize pedestrian delay, and geometrically fixed junctions require to break coordination. This interrupted flow should be done purposely at junctions with sufficient storage space for queuing. Coordination will only work on multilane arterials with extra pocket lanes for turning vehicles, closely spaced signals (<500 m), and no interruption by parked vehicles, on-street delivery or bicycles traveling on the road. Furthermore, the maximum degree of saturation (volume over saturation flow) should not exceed 90% to provide some spare capacity for the stochastic impact of traffic flow. In most cases of coordination failure, at least one of the above constraints is not considered so that traffic engineering by itself will not provide a good solution. There must be accompanied measures such as prevent parking and cycling, prohibit turns, and limit traffic flow demand (Brenner, 2005). Without these measures, coordination will not be possible as the system will be running above capacity.

Coordination is applied in two different methods: simultaneous and progressive coordination. The simultaneous system displays the same signal simultaneously at all signals on the coordinated arterial. A simultaneous system is suitable for short junction spacing (up to 100 m). In the progressive system, the green times are always shifted at each junction by the travel time from stop-line to stop-line (FGSV, 2010). Progressive systems are subdivided into simple, forward, flexible, and reverse progressive systems, according to the time difference (positive–negative) between green initiations (Roess, Prassas, and McShane, 2010).

The US literature and a variety of software systems such as Linsig, Sidra, Syncro, Transyt, Vistro, and HCM provide tools to identify the best possible coordination. However, these models are typically car-oriented without sufficient representation of urban constraints as found in traditional European cities. Therefore, coordination is manually adapted with tools such as Lisa or Sitraffic Office, which break coordination on purpose to improve uninterrupted flow on the remaining segments. The following figure gives an example of coordination planning in Germany with nonequally spaced junctions (Figure 8).

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

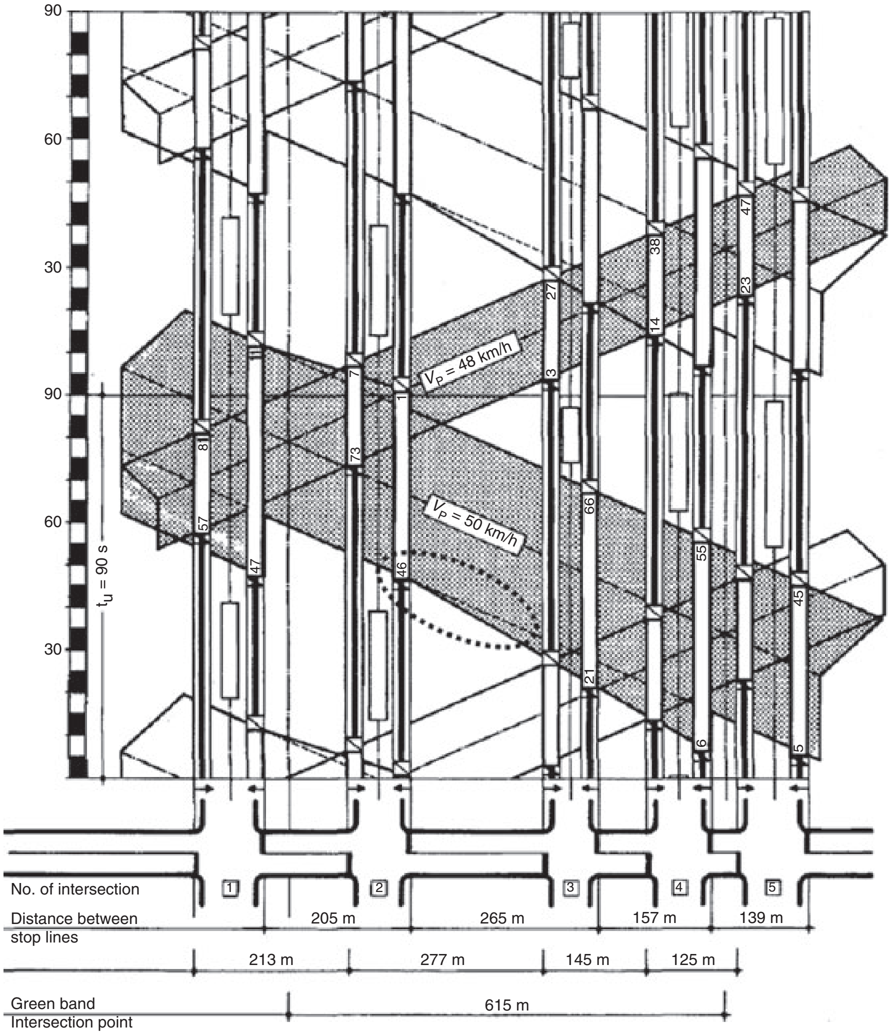


Figure 8. Space–time diagram of continuous coordination of eastbound traffic and interruption of the early platoon of the opposite direction at the marked signal. (FGSV, 2010. Reproduced by permission of FGSV Verlag GmbH.)

3 TRAFFIC-RESPONSIVE SIGNAL CONTROL

While fixed time signal control is either based on measured historical traffic volumes or forecasted design volumes, traffic-responsive signal control operates on real-time measurements. There is a closed loop between the demand of traffic and the signal timing plan. Traffic-responsive control is sometimes also called *traffic-dependent control*, which virtually indicates the feedback loop between travel demand and the signal indications given by the control device. Traffic-responsive signal control comprises (i) traffic-responsive plan selection (TRPS), (ii) actuated signal control, and (iii) adaptive traffic control systems (ATCSs).

3.1 Traffic-responsive plan selection (TRPS)

In fixed time signal control, a variety of signal timing plans is designed to fit different situations of travel demand. For each set of average travel demand, an optimal signal timing plan is computed. The signal timing plan is activated according to a weekly schedule. Usually, about four different programs are designed for an average workday and another four programs are available to fit weekend and holiday traffic. Additional programs are prefabricated to suit emergency situations, special events of highly directional flow, and other unusual situations.

TRPS may use the same set of prefabricated signal timings. However, a signal timing plan is not changed because of time of day but because of changes in traffic demand. Traffic-responsive mode was developed to ensure implementing appropriate timing plans recognizing actual traffic conditions in the network. Actual traffic demand is measured at strategic locations in the network. The urban traffic controller system (UTCS) compares the measured traffic demand against specific thresholds to determine the most suitable pre-timed timing plan. Subsequently, the traffic-responsive system implements a timing plan that is best suited to accommodate the current traffic demand. Generally, the cycle length, stage sequence, and green times at each junction remain fixed until a new timing plan is implemented, according to traffic demand.

TRPS is usually implemented in regional control systems with several signalized junctions each connected via a UTCS. Within a road network, a set of strategic detectors is distributed to measure volumes, density, and local speed. A well-designed algorithm needs to decide on the most suitable timing plan that will match the measured traffic conditions best (Gündogan, 2012). The decision process provides numerous options as traffic volumes will not likely fit the pre-timed signal plans exactly. In the United States, TRPS uses a computational channel to utilize a set of pattern selection parameters to select a timing plan based on the threshold values for each of these parameters (Abbas and Sharma, 2006).

In TPRS, the most suitable signal timing plan is selected based on conditional and temporal structured queries. TPRS can also be seen as a traffic-actuated control method at a macroscopic level because only strategic but no local detectors are used for the conditional queries. Various TPRS systems vary by the way they estimate the current traffic demand based on a limited set of strategic point measurements. Once the current traffic state is estimated, the most suitable signal time plan for each local traffic controller is selected out of the set of pre-determined signal timing plans. Usually, sets of signal plans for a region with several signalized junctions are considered simultaneously.

The known shortcomings of TPRS include the limited number of pre-timed signal timing plans and the switching process between two signal programs. In coordinated systems, it is difficult to switch between programs with different cycle length because the progression is interrupted. It may take several cycles until such a system stabilizes. Therefore, some TPRS implementations use sets of fixed cycle lengths with changes in green splits only in order to limit continuous changes in the offsets for the major movements on coordinated arterials. Changes in the base cycle length are only done as travel demand patterns change over time of day such as morning peak with directional inflow to the city centers versus more evenly distributed travel patterns during the day.

3.2 Actuated signal control

While TPRS looks at all signalized junctions within a regional control area at once, traffic-actuated control considers first individual junctions. Each controller is equipped with a software and hardware, which allows him/her to respond to traffic demand measured locally at each junction. The local controller is equipped with a software unit to process a logic by directly querying the logic such as green time extension conditions or temporal conditions such as predefined maximum green time durations of stages. Real-time evaluation and an optimization are not required in this type of control as all state evaluations are based on pure conditional clauses. Green splits will change according to the frame plans and measured traffic demands on the approaches. Cycle length may change but is usually fixed if the actuated traffic controller is embedded within a series of signalized junctions in an arterial or regional control area. Actuated signal control can be applied at a single junction either in semi-actuated or fully actuated mode. In semi-actuated control, detectors generally exist only for minor movements. These minor movements receive only GREEN if traffic demand is measured. In addition, green extensions or truncations can be activated for the minor movements. In semi-actuated mode, only a subset of movements are actually measured by detectors, whereas in fully actuated mode, all approaches (or signal groups) are furnished with detectors.

In arterial or network coordination, either the intersections are operated locally as actuated mode and coordination of intersection are provided centrally, or coordination with fixed cycle length and limited flexibility at each frame plan comes forefront with semi-actuated adjustments at each junction. Traffic-actuated control is the most flexible form of traffic-responsive systems as it can be programmed for each signalized junction individually. However, this may also require a large amount of customized programming effort Figure 9. Various vendors provide software libraries to reuse software modules for actuated traffic control.

Actuated control strategies can be classified in stage-based and signal-group-based systems. Stage-based systems use a stage as the smallest control unit. As in Figure 3, stage 1 may be omitted if there is no request by pedestrians P21 and the left turn movement V2. Then, the stage sequence may run stage 2−stage 3−stage 2 with the additional interstage (3,2). In a signal-group-based system, each signal group may be activated individually without stages and interstages. Taking the above example, pedestrians may have pressed the push button but no left turn vehicles are present at signal group V2. Then, a signal-group-oriented system may activate P21 parallel to V3 and V5, if these turns allow semicompatible movements.

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

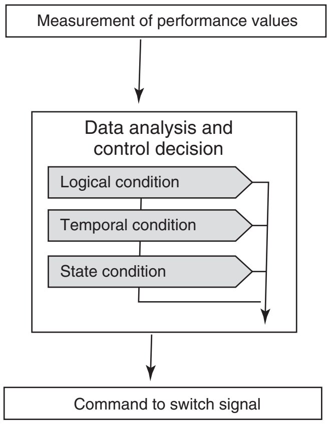


Figure 9. Actuated signal control architecture. (FGSV, 2010. Reproduced by permission of FGSV Verlag GmbH.)

Signal-group-based actuated control is the most versatile control technology at the cost of additional programming and maintenance effort. It is widely applied in advanced priority systems for buses and trams in many European cities. Figure 10 indicates a typical example of public transport priority with green extension dependent on detection at three sites. Within the approach, the bus has a passenger stop. The dwell time is distributed according to the historical data collection. The actuated algorithm computes earliest arrival and latest arrival by taking minimum and maximum expected approach speeds and dwell times. At the main detector, the arrival will be forecasted quite precisely and the green extension will be prepared precisely. Advanced public transport priority schemes also consider conflicting public transport routes by comparing vehicle delay and passenger loading. This requires precise vehicle identification with additional information on vehicle course, time table, and real-time passenger counts as defined in the German protocol R09/16 (Figures 9 and 10).

Public transport vehicles are prioritized by changing the structure of a signal program or only changing stage durations. In public transport priority systems, various options are widely applied such as (i) truncation/extension of cycle time, (ii) change of stage sequence using the same stages in a different order, (iii) change of stage sequence with additional bus/tram stages, and (iv) truncation/extension of stage durations to allow an early start or late finish of a stage with the respective public transport signal. All of these methods can also be applied on a signal group basis with more versatility at additional computational effort. In systems with extensive public transport, signal priority performance measures have to be evaluated to limit negative impacts on other modes such as delays of conflicting vehicle movements, queue lengths at approaches with limited storage space, and maximum delay times for pedestrians and bicyclists to prevent the likelihood of red-running.

3.3 Adaptive signal control

The development of traffic adaptive systems dates back to the 1980s. Instead of programming individual controllers for actuated control, a traffic-responsive system was needed to control a set of controllers. In the best way, all controllers of a city should be controlled from one central UTCS very much like TRPS but with more local flexibility. Within the early UTCS, timing plans were newly calculated (no pre-timed signal plans) every 15 min based on surveillance data. Later systems were conceived to implement and evaluate a fully responsive, online traffic control system with shorter update intervals of 3–5 min. Compared with traffic-actuated and TRPS systems, a traffic adaptive system uses a traffic flow model, an impact model, and an optimization model to decide on signal changes based on frame plans (Figure 11). The traffic situation is replicated in a dynamic traffic model with data that is continuously collected on point data. The detailed analysis of the traffic situation with different traffic parameters allows the classification of traffic conditions and timely detection of traffic disruption (RiLSA, 2010). Adaptive systems contain a traffic flow model to estimate the current traffic state. On the basis of this traffic state estimation, real-time optimization method is applied to identify the best suitable timings. 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 and cycle length to prevent negative transition effects (i.e., Split, cycle, and Offset Optimization Technique (SCOOT)), whereas 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).

The main advantage of traffic adaptive systems is the ability to adapt not only green time according to short-term fluctuations on traffic demand but also the automatic generation of traffic-related frame plans. Reduction in the number of stops and delays are proved in several comparative studies (Brilon and Wietholt, 2013). However, high installation and maintenance costs for software add to a large number of detectors and difficulties in process calibration. Adaptive control methods vary according to the software system, the detector network, and the local traffic engineering experience. In the best case, the system

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

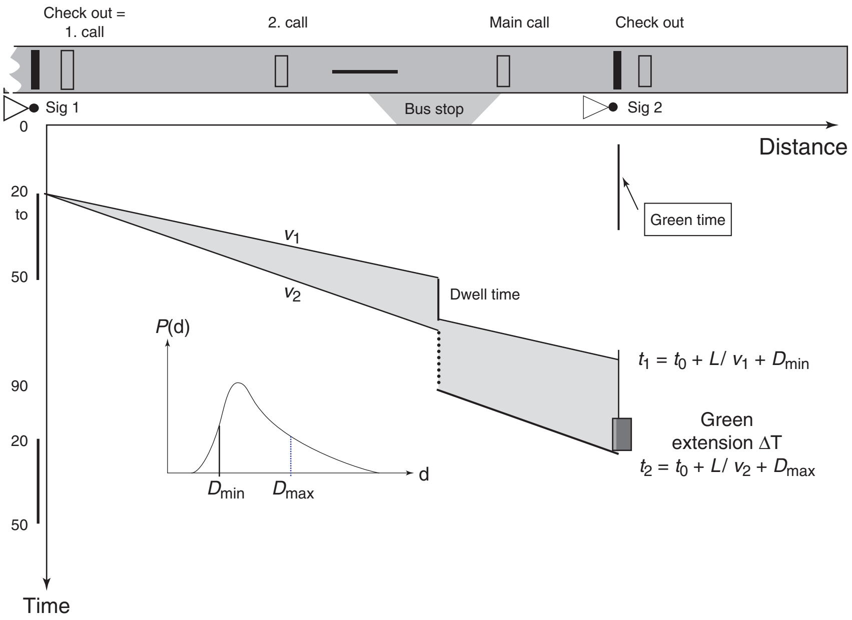


Figure 10. Arrival prediction at signal 2 of a bus measured at the 1. Call detector.

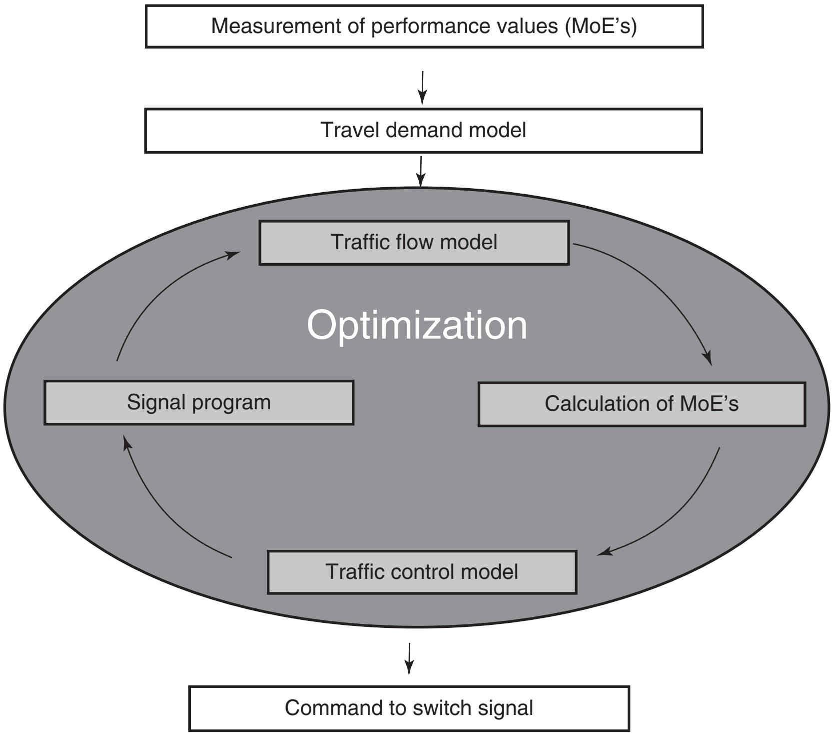


Figure 11. Adaptive signal control architecture. (FGSV, 2010. Reproduced by permission of FGSV Verlag GmbH.)

is self-calibrating in case of changes in travel demand patterns, and in the worst case, the system tends to operate like a fixed time system at much higher cost for the detection system. There are a number of ATCSs on the market, which are installed as UTCS; the most prominent ones are described in the following.

SCOOT was the first commercial optimization techniques for adaptive systems. SCOOT adjusts signal timing parameters in frequent, small increments to match the latest traffic situation. The main idea of SCOOT are cyclic flow profiles (Figure 12). The cyclic profiles consist of histograms, which record the variety of volumes during one cycle at the upstream signal (Hunt *et al*., 1981). SCOOT detectors are placed at each downstream link of each junction. SCOOT estimates queues. The current traffic state is recorded and vehicle arrivals are known from cyclic flow profiles. Thus, vehicle arrivals during the red time are added to the tail of each queue. The queue increases until the next green time (Hunt *et al*., 1981). SCOOT is the most widely used ACTS worldwide. Its traffic flow model predicts arrival patterns quite well but it cannot cope with public transport priorities except weighting buses and trams with a higher priority. As many cities require public transport priority based on one second accuracy, SCOOT is not the first choice for those cities.

SCATS (Sydney Coordinated Adaptive Traffic System) was developed and first installed in Sydney, Australia. Now, SCATS is used in more than 100 cities worldwide. Similar to SCOOT, SCATS adjusts 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. SCATS optimizes splits and cycles according to

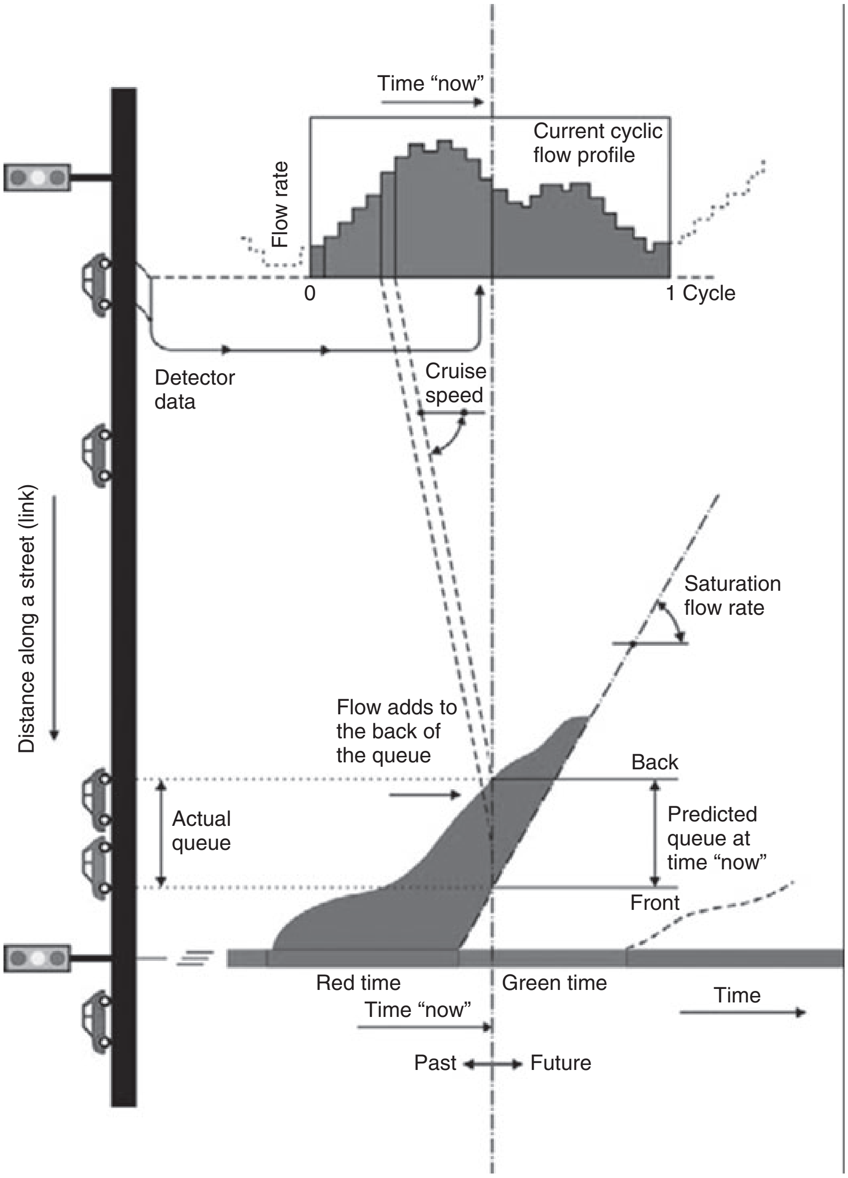


Figure 12. Traffic flow model in SCOOT. (Hunt *et al*., 1981)

Encyclopedia of Automotive Engineering, Online © 2014 John Wiley & Sons, Ltd.
This article is © 2014 John Wiley & Sons, Ltd.
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5

an objective function. Although the objective function is determined according to delay and number of stops, the degree of saturation is used as main parameter in the optimization. It controls the traffic in two levels: the strategic and the tactical levels, which are also used in newer adaptive systems. In the tactical level, the regional computer of SCATS controls up to 10 intersections. All regional computers of one city are connected to the main control system to decide on the strategic level.

There are two German adaptive systems called *Motion* and *Balance*. Both use a multilayer traffic control approach. Optimal cycle length, stage sequence, and coordination are defined every 5–15 min on a tactical level. Green splits are readjusted on a second-by-second basis to guarantee full flexibility for immediate interventions by public transport vehicles. At complex junctions, there is a high probability of urgent requests by single vehicles (trams, buses, and emergency vehicles). These requests intervene with the tactical level so that tactical decisions are most likely overruled in networks with public transport priority. The initial tactical plans are overruled in many instances. Utopia from Italy is another system with a two-level optimization process. Opac and Rhodes are two American systems with low practical impact but strong academic background. A feature list of the various adaptive systems may clarify (Table 1).

3.4 Signal control and V2X

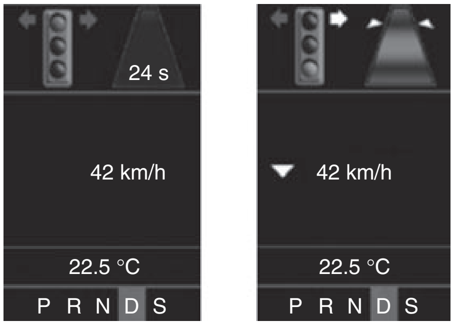
After the millennium turn, industry and science initiated research on vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication summarized as vehicular communication technology (V2X). Intelligent Transport Systems (ITSs) embrace a wide variety of communications-related V2X applications intended to increase travel safety, minimize environmental impact, and improve traffic efficiency to both commercial users and the general public. With respect to traffic signals in urban environments, the Green Light Optimized Speed Advisory (GLOSA) is the most likely candidate for implementation. First trials by Braun *et al*. (2008) showed the technical feasibility under real-world conditions in the city of Ingolstadt. The traffic signal controller communicates with approaching vehicles via WLAN. As the traffic signals were operated as actuated signals, pre-timed signal timings were not available. Under actuated or adaptive conditions, green and red times must be predicted either by detailed analysis of the control logic or by evaluation of the control history and statistical trend analysis. These methods are not fully fault-proven but provide additional information to the driver within an accuracy of a few seconds. This will be accurate enough so that the driver can adopt his or her driving behavior according to the traffic lights. If the remaining red indicates a time longer than a vehicle-dependent threshold, fuel will be saved if the engines start–stop device will be activated. Below the threshold, it will be beneficial if the engine will idle. The dynamic speed advisory GLOSA provides the driver with information to accelerate in order to reach the next signal at GREEN or to slow down. There is ongoing research to improve the Human–Machine-Interface. One possible implementation is shown in Figure 13.

On the basis of micro simulation studies, Katsaros *et al*. (2011) report up to 7% fuel saving due to GLOSA if 80% of all vehicles are V2X equipped. At lower penetration rates, the effect will be reduced. The GLOSA should be activated 300 m before the traffic signal in urban environments with

Table 1. Features of adaptive system.

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/5 min	3 s-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.
Gündogan, 2012 based on Stevanovic, 2010.



(a) (b)

Figure 13. On-board remaining red time display (a) and (b) dynamic speed advisory. (Braun *et al*., 2008)

50 km/h speed limit. There is more research ongoing with respect to GLOSA and display of remaining red time.

One of the practical difficulties is different standards in the automotive area compared to traffic engineering devices, which are not fully equipped with well-documented standards. V2X and traffic signals are still in its infancy and require further standardization and development.

4 OTHER URBAN TRAFFIC MANAGEMENT MEASURES

Although traffic signals are the most effective traffic control measure in urban areas there are several other measures to manage traffic. Most of these measures cannot be considered truly operational as they also effect on a tactical level.

4.1 Dynamic lane allocation

On several multilane highway bridges and tunnels contraflow systems are quite common. The limited number of lanes is allocated to each direction according to travel demand. The given infrastructure is most efficiently used by keeping the similar degrees of saturation on each side. This will prevent to add extra lanes at costly infrastructures such as bridges and tunnels. Examples can be seen on bridges such as Golden Gate (California) or the Elbtunnel in Hamburg. Contraflow systems in urban areas can be classified in

1. Systems of dedicated lanes for particular vehicle types for temporal use during particular time of day or traffic conditions such as dynamic bus lanes, which can be used also by taxis in several cities. In most cities, the temporal activation is indicated by static signs for certain times of day. However, there are ongoing tests to limit these lanes only if traffic conditions require extra space for certain vehicle modes.
2. Contraflow systems to allocate lanes to particular travel directions similar like the above-mentioned highway systems. These systems can be found in the vicinity of large stadiums and convention centers next to parking facilities. As travel demand is usually centered around the duration of an event, there is a high demand toward the parking facilities before the event and outgoing demand after the event. The lane allocation is usually marked by Variable Message Signs (Figure 14).

4.2 Parking guidance

Limiting parking space is one of the most effective measures to reduce vehicle travel demand in urban areas. However, this is considered a strategic traffic management measure according to Figure 1 and is not part of this chapter. There are some operational measures to manage the given parking space. In the 1980s parking guidance systems were installed in many cities indicating the available number of parking spaces. The systems are still in operation because they help to reduce the amount of park-search traffic. In large cities multilevel information is provided. On the outskirts general directions and the total number of empty parking spaces are given. The closer the driver gets to his/her destination the more precise the information gets listing all parking facilities in the area with information on distance and available parking lots. Parking guidance systems are usually limited to above ground and underground parking garages. The number of

Figure 14. Urban contraflow system next to Nordpark Mönchen-Gladbach. (Reproduced with permission from Grimm GmbH.)

entering and leaving vehicles must be counted which is common in all systems with surveillance of the parking period. Some parking facilities provide above ground infrared detectors which identify whether a parking space is occupied. These garages can provide the number of available parking lots without a gating system at the entrance or exit. These facilities are now available at some shopping centers. As the shopping centers by themselves do not maintain Variable Message Signs they provide the available number of parking lots via Internet sites and smartphone applications. Some research is being done to provide booking of parking space.

5 CONCLUSION

This chapter is limited to operational traffic management in urban areas. As such more strategic management schemes such as road pricing schemes and traveler information were not considered. Advanced traveler information systems (ATISs) provide information on real-time travel options. Recently multimodal journey planners implemented on portable mobile devices (smartphones) provide real-time travel information in unprecedented spatial and temporal accuracy. As such, ATIS will influence the behavior on departure time choice, mode choice, and route choice for long and short distance trips. While long distance trips are usually planned ahead, short daily trips are usually based on experience. Little empirical evidence is available to forecast the impact of ATIS on local trips in urban areas although the impact may be enormous because of the number of daily local trips for work, leisure, and education. It will be interesting to see whether ATIS will be provided by private providers with little incentive of traffic control issues or public agencies with an interest to control and guide traffic according to specific objective stated in urban traffic master plans. However, traffic signals will remain the most effective way to control road-based traffic in urban areas by limiting certain routes via green and red signals. Gating becomes a viable option to limit the amount of traffic passing corridors within traffic sensitive urban areas. This chapter provides some basic information on traffic signals, its functionality, and planning concepts.

RELATED ARTICLES

Intelligent Transport Systems: Overview and Structure (History, Applications, and Architectures)
Intelligent Transport Systems: Market and Policies
Data Acquisition by Roadside Detection
Advanced Highway Management Systems

REFERENCES

Abbas, M.M. and Sharma, A. (2006) Multiobjective plan selection optimization for traffic responsive control *Journal of Transportation Engineering*, **132**, 376–384.

Braun, R., Busch, F., Kemper, C. *et al*. (2008) TRAVOLUTION – Netzweite Optimierung der Lichtsignalsteuerung und LSA-Fahrzeug-Kommunikation. Straßenverkehrstechnik Heft 6, S. 365–374.

Brenner, M. (2005) in *Stadtverkehrsplanung Grundlagen – Methoden – Ziele*, 2nd edn (eds G. Steierwald, H.D. Künne, W. Vogt), Springer, Berlin.

Brilon, W. and Wietholt, T. (2013) Experiences with Adaptive Signal Control in Germany. Presented at TRB Annual Meeting, paper no. 13–1155, Washington.

DfT Department for Transport, Scottish Executive, Welsh Assembly Government (2006) Traffic Advisory Leaflet TAL 1/06 – General Principles of Traffic Control by Light Signals (Part 1–4), London.

FGSV Forschungsgesellschaft für Straßen- und Verkehrswesen (2010) Richtlinien für Lichtsignalanlagen (RiLSA Ausgabe 2010), Cologne.

FSV, Österreichische Forschungsgemeinschaft Straße und Verkehr (1998) Richtlinien und Vorschriften für das Straßenwesen RVS 05.04.32 Planen vonerkehrslichtsignalanlagen, Vienna.

Gündogan, F. (2012) Simplified traffic responsive signal control method for developing large cities. Inaugural dissertation at Graz University of technology.

Hunt, P.B., Robertson, D.I., Bretherton, R.D., and Winton, R.I. (1981) SCOOT – A traffic responsive method of coordinating signals. TRRL Laboratory Report 1014, Transport and Road Research Laboratory, Crowthorne, UK.

Katsaros, K., Kernchen, R., Dianati, M., and Rieck, D. (2011) Performance Study of a Green Light Optimized Speed Advisory (GLOSA) Application Using an Integrated Cooperative ITS Simulation Platform. *International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 918–923.

Roess, R.P., Prassas, E.S., and McShane, W.R. (2010) *Traffic Engineering*, 4th edn, Prentice-Hall, Upper Saddle River, New Jersey.

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

Transportation Research Board (2010) Highway capacity manual. National Academies of Sciences, Washington, D.C., London.

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

Encyclopedia of Automotive Engineering, Online
DOI: 10.1002/9781118354179.auto181
Also published in the *Encyclopedia of Automotive Engineering* (print edition) ISBN: 978-0-470-97402-5