Evaluation of bus priority strategies in coordinated traffic signal systems

JOHAN WAHLSTEDT
Evaluation of bus priority strategies in coordinated traffic signal systems

Johan Wahlstedt

Licentiate thesis in Infrastructure
with specialisation in traffic and logistics

KTH Royal Institute of Technology
School of Architecture and the Built Environment
Department of Transport Science
Division of Traffic and Logistics
SE-100 44 Stockholm, Sweden
Akademisk avhandling som med tillstånd av KTH i Stockholm framlägges till offentlig granskning för avläggande av teknologie licentiatexamen onsdagen den 21 maj kl. 13:30 i sal V1, KTH, Teknikringen 76, Stockholm.
Abstract

Increasing congestion and environmental concerns have evoked an interest in promoting urban Public Transport (PT) the last decades. In 2012 the City of Stockholm adopted an “Urban mobility strategy” stating that public transport, cycling and walking should be prioritised over cars in central Stockholm. One of the most important factors influencing the modal choice is the travel time ratio between car and PT travel. According to earlier studies Public Transport Traffic Signal Priority (PTSP) can reduce travel times for public transport with only small negative impacts on other traffic. Conditional PTSP can also help to regulate the PT service. Thus PTSP may support drivers’ decision to change travel mode from car to PT, thus supporting adopted policy goals.

Conventional control strategies for coordinated traffic signals have pre-set timings based on traffic surveys. Some traffic adaptation based on real time detector actuations can also take place within the frames of the pre-set cycle time. PTSP changes the signal timings, within pre-set limits, when a PT vehicle is detected. Self-optimising control strategies use a traffic model to predict the traffic flows from traffic counts, and determine the signal changes in real-time by minimising a cost function including delay, number of stops etc. PTSP is included directly in the optimisation by giving PT vehicles a higher weight compared to cars.

In this thesis the fundamentals of signal control theory are reviewed as well as unconditional and conditional PTSP criteria and strategies. A simulation based method for evaluation of impacts of different PTSP strategies in coordinated controlled traffic signals is implemented. The simulation setup includes Software-In-the-Loop (SIL) signal controller simulators running the same control logic as used in field. Such simulation models can be useful to test and fine tune PTSP before being implemented in field. Simulations with a SIL setup also enable comparisons of signal control strategies or systems on equal terms, not practically or economically possible in field studies. The implemented SIL simulation model was used to evaluate the impacts on buses and other traffic from the different PTSP functions used in the “PRIBUSS” PTSP method. Short green time extensions showed travel time reductions for buses, with almost no travel time increase for other traffic.
Long green time extensions gave somewhat larger benefits for the buses, but more delay to other traffic. Red truncation gave less travel time savings to the prioritised buses and more extra delay for cross street traffic, compared to green extensions. Double red truncation and Extra phase showed some additional travel time savings to the buses, but had the largest negative impact on other traffic. A combination of PRIBUSS functions showed the best results. Depending on the structure of the signal coordination and the location of the bus stops different PTSP functions may be needed.

Based on the conclusions from the evaluation of the different PRIBUSS functions a conditional “differential on-time-status” based PTSP strategy was proposed and tested in the SIL simulation environment. The proposed method is focusing on direct travel time savings as well as on reduced bus bunching.

The two self-optimising signal control systems Utopia/Spot and ImFlow were tested, and their impacts were compared to conventional control including PTSP with the PRIBUSS method in a SIL simulation environment. The aim was to test if commercially available self-optimising control systems can reduce the overall delay per person by applying more sophisticated PTSP. Both systems reduced the delay for buses, cyclists and pedestrians at a cost of increased delay and increased number of stops compared to the existing conventional control used in field. The total delay for all road users was reduced substantially.

Keywords
Bus priority, public transport priority, transit priority, PTSP, TSP, traffic signal, traffic signal control, traffic signal control systems, self-optimising signal control, PRIBUSS, SIL simulation, traffic simulation
Sammanfattning
Intresset för att påverka resvanorna i våra städer så att kollektivtrafikandelen ökar har växt de senaste decennierna på grund av en ökad trängsel i gatunätet samt ökad miljömedvetenhet. Stockholms stad har antagit "Framkomlighetsstrategin" som innebär att kollektivtrafik, gång och cykel ska prioriteras framför biltrafik i centrala Stockholm. En av de faktorer som påverkar färdmedelsvalet mest är restidskvoten mellan bil och kollektivtrafik. Tidigare studier har visat att kollektivtrafikprioritering i trafiksignaler kan minska körtiden för kollektivtrafiken väsentligt, med små eller inga negativa konsekvenser för övrig trafik. Villkorig prioritering kan dessutom förbättra kollektivtrafikens regularitet. Kollektivtrafikprioritering i trafiksignaler kan på så sätt hjälpa till att förbättra kollektivtrafikens attraktivitet och därigenom öka kollektivtrafikandelen.


I denna avhandling beskriver teoretiska grunderna för trafiksignalstyrning, liksom metoder och kriterier för villkorig och ovillkorig signalprioritering av kollektivtrafik. En simuleringssbaserad metod för att utvärdera effekterna av olika signalprioritering har implementerats. Denna använder styrapparatsimulatorer med samma programmering som styrapparaterna på gatan, inklusive prioriteringsfunktioner. Sådana simuleringar kan vara ett användbart verktyg för att justera in prioriteringsfunktionerna innan dessa implementeras i signalstyrningen på gatan. Simuleringar med...

Baserat på utvärderingen av de olika PRIBUSS funktionernas effekter på bussar och övrig trafik har en tidhållningsbaserad differentierad prioriteringsstrategi föreslagits, som förutom att skapa direkta restidsvinster också försöker motverka ihopklumpning av bussar. Denna strategi har implementerats och testats i den framtagna simuleringssmiljön.

Acknowledgements
This work was largely funded within the METKAP project financed by The Swedish Transport Administration, which is hereby gratefully acknowledged.

I would like to express my gratitude to my supervisors Karl-Lennart Bång, Haris Koutsopoulos and Lars-Göran Mattsson. It was Karl-Lennart Bång who first made me interested in returning to KTH at part-time to write this thesis. He has guided me in the academic world for four and a half years through the not always straightforward process resulting in this thesis.

I would also like to thank the following persons and organisations:


Johan Steen and Anders Hällbom who have answered endless numbers of questions, explained how a signal controller works, and how to program it.

The City of Stockholm for financing the project resulting in paper IV, and for supplying data, documentation and information needed in the other studies.

My former and current managers at Ramböll, especially Carl Chytraeus, who allowed me to be on leave on part-time to write this thesis, and also contributed economically to many conference attendances.

My colleagues at Ramböll, especially Jevgenij Petoukhov and Farshad Farhat, who carried out a lot of the boring work for me, covered up for me in different projects, and also have assisted with data preparation in the simulation studies.

My colleagues at KTH for many stimulating discussions, which makes KTH the interesting environment it is. A special thanks to Albania (Bibbi) Nissan for all encouragement and generosity.

My family and friends for their patience all evenings and weekends. Sorry for being absent!

Stockholm, April 2014

Johan Wahlstedt
Publications

Paper I

Paper II

Paper III

Paper IV

The Swedish Transport Administration publication TRV 2013:79994
Wahlstedt, J. (2014), Handbok för kapacitetsanalys med hjälp av simulering, kapitel 6 Signalreglerade korsningar (Handbook on capacity analysis with simulations, Chapter 6 Signalised intersections), Trafikverket (The Swedish Transport Administration), publication TRV 2013:79994, March 2014, Borlänge, Sweden
Table of contents

1. Introduction __________________________________________ 1
   1.1. Background to public transport priority in traffic signals _______ 1
   1.2. Objectives _____________________________________________ 4
   1.3. Scope________________________________________________ 4
   1.4. Historical background from a Stockholm perspective __________ 5

2. Overview of traffic signal control methods _________________ 7
   2.1. Introduction ___________________________________________ 7
   2.2. Criteria for signal control________________________________ 8
   2.3. Isolated control _________________________________________ 8
       2.3.1. Fixed time isolated control ______________________________ 9
       2.3.2. Vehicle actuated isolated control_________________________ 13
   2.4. Coordinated control ____________________________________ 16
   2.5. Self-optimising control _________________________________ 19

3. Methods for bus priority _______________________________ 24
   3.1. The need for bus priority________________________________ 24
   3.2. Passive traffic signal priority ____________________________ 24
   3.3. Active traffic signal priority ______________________________ 25
       3.3.1. Conditional and unconditional priority____________________ 26
       3.3.2. Conflicting priority requests ______________________________ 27
       3.3.3. Conventional control strategies____________________________ 27
       3.3.4. Self-optimising control strategies__________________________ 28
   3.4. Bus detection _________________________________________ 29

4. Criteria for PTSP _________________________________________ 32
   4.1. Criteria for introducing PTSP ______________________________ 32
   4.2. Minimising total travel time with conditional PTSP __________ 32
   4.3. Improving PT service reliability with conditional PTSP __________ 33
   4.4. Differential PTSP ______________________________________ 34
5. Examples of systems for bus priority ____________________ 35
  5.1. PRIBUSS ___________________________________________ 35
  5.2. SCOOT _____________________________________________ 37
  5.3. SCATS _____________________________________________ 37
  5.4. MOTION ___________________________________________ 39
  5.5. Utopia/Spot _________________________________________ 40
  5.6. ImFlow ___________________________________________ 41

6. Evaluation of bus priority impacts in traffic signal systems __ 43
  6.1. Data needed for evaluation ___________________________ 43
  6.2. Empirical tests ______________________________________ 43
  6.3. Analytical models ____________________________________ 44
  6.4. Simulation as tool for evaluating signal strategies ________ 45

7. Evaluation study ______________________________________ 47
  7.1. Introduction and methodology __________________________ 47
  7.2. Simulation setup _____________________________________ 47
  7.3. Case study one ______________________________________ 50
    7.3.1. Study area ______________________________________ 50
    7.3.2. VISSIM model ____________________________________ 52
    7.3.3. Traffic signal control ______________________________ 54
    7.3.4. Results __________________________________________ 54
  7.4. Case study two ______________________________________ 56
    7.4.1. VISSIM model ____________________________________ 57
    7.4.2. Traffic signal control ______________________________ 58
    7.4.3. Results __________________________________________ 61
  7.5. Case study three _____________________________________ 67
    7.5.1. VISSIM model ____________________________________ 67
    7.5.2. Tested criteria for conditional PTSP __________________ 68
    7.5.3. Results __________________________________________ 70
1. Introduction

1.1. Background to public transport priority in traffic signals

In the last decade increasing congestion problems on urban streets combined with environmental concerns and reduced funds for adding road capacity have evoked an interest in promoting Public Transport (PT) all over Western Europe. There have also been shifts in focus from building metros to high quality surface transport, e.g. Bus Rapid Transit (BRT) and/or tram lines that share intersections with other road traffic. One of the most important factors that is known to influence modal choice is the travel time ratio between car and PT travel. It is also known that delay at signalised intersections constitutes a large part of PT journey time in urban areas. One way to reduce delay and improve service regularity for PT at a relatively low cost to other traffic is to introduce PT priority at traffic signals (Bång 1987; Al-Mudhaffar & Bang 2006; Zlatkovic, Martin & Stevanovic 2009).

Traffic signals are traditionally designed to minimise the delay per vehicle. However vehicles carry a very different number of passengers; a bus typically has 10 – 20 times the number of passengers compared to an average car and a tram contains even more (Bång 1987). In order to minimise the delay per person, PT vehicles need to be treated differently from cars at traffic signals, i.e. the PT vehicles need to be prioritised in order to minimise the delay per person. PT priority is not trivial to implement in conventional coordinated traffic signal systems, as “green waves” are often disrupted, and lengthy minimum times for pedestrian crossings restrict the possible signal changes. It may also be necessary to make green time compensations on other approaches once a PT vehicle has passed, which may have further negative impacts on the performance of the coordinated system (Wahlstedt 2013a).

Public transport traffic signal priority (PTSP) can be passive or active. In passive PTSP signal timings are set to favour PT in general. Active PTSP triggers priority measures when there is a PT vehicle present, thus requiring selective detection of PT vehicles. Active PTSP can be unconditional or conditional. Unconditional priority will always favour the PT vehicles without considering negative impacts for other road users, while conditional priority in one way or another tries to
restrict the impacts on other traffic. The criteria for conditional PTSP can be based on the impacts on other traffic and/or on the PT vehicles. Examples of criteria for conditional PTSP based on PT vehicles are: only give priority to certain lines; not giving priority to early buses/only giving priority to late buses. Criteria based on the on-time status of PT vehicles can improve PT regularity and service reliability (Furth & Muller 2000). A special case is prioritising between conflicting PT vehicles; two crossing lines, a line turning in a two stage controlled intersection, or even buses on the same line meeting at an unfavourable position. If the PT vehicles arrive at unfavourable timing and call for conflicting priority measures it will not be possible to serve both. Most PTSP systems do not handle the prioritising between PT vehicles in such cases (Christofa & Skabardonis 2010).

Bunching is a problem for bus, and to some extent tram, operations; if a bus is delayed there will be more passengers waiting at the next stops, resulting in longer dwell times, which make the bus even later. Furthermore the following bus will have a shorter headway and therefore fewer passengers at the stops and will catch up with the late bus. Bus bunching leads to an ineffective use of the bus fleet and causes delays and crowding for the passengers (van Oort 2011). The bunching increases the travel time to public transport users, which could be considered as delay in a broader sense. Much can be gained if the initial small delay is handled before it has escalated and causes bunching. Conditional PTSP only giving priority to late buses can reduce bunching significantly, compared to unconditional priority (Altun & Furth 2009; Ma, Yang & Liu 2010). Late vehicles can be “pushed” on time and early vehicles can be held back. With conditional priority the negative impacts on other traffic can also be reduced compared to unconditional priority (Furth & Muller 2000; Albright & Figliozzi 2011).
In conventional PTSP control strategies, the signal timings are adjusted with consideration to PT vehicle arrivals by local signal timing adaptations made within a fixed timed coordinated system. In some PTSP strategies the signal control is changed to isolated mode when prioritising a PT vehicle, and retained to coordination afterwards (Salonen 2010). Self-optimising signal control strategies minimise the overall intersection or area delay by minimising an object function (e.g. minimise total road-user costs) where the PT vehicles are given a higher weight than cars, and thereby prioritised.

Self-optimising signal control has been tested at a few locations in Sweden (Bang 1976; Kronborg 1992; Kronborg, Davidsson & Edholm 1997; Al-Mudhaffar 2006), but only one installation (Luthagsleden, Uppsala) is in operation today. The Italian self-optimising signal control system Utopia/Spot was tested at two sites in Stockholm 2000-2001 and 2003 (Kronborg & Davidsson 2004). The results showed travel time savings around five to ten per cent compared to updated conventional control (Al-Mudhaffar 2006). Despite the travel time savings the City of Stockholm chose not to implement self-optimising signal control at the time for several reasons: the extra investment and maintenance costs, concerns about implementation of Swedish control strategies in the self-optimising system and because there was only one supplier at the time.

In general it is difficult to analytically calculate the effects of dynamic signal timing or active PTSP. Although initial signal timing may be determined based on theoretical analysis, PTSP settings are usually conducted on the basis of traffic engineering experience. Analytical models for evaluation of PTSP in coordinated signals exist (Abdy 2010), but it is hard to calculate the impacts of active PTSP in coordinated systems in detail since the dynamic changes of signal timings in different signal controllers interact with each other. The platoon progression will be different in each cycle when the frequency of prioritised PT lines is high and/or when the coordinated system is large and containing many intersections. The impacts of a priority measure can also occur at a downstream intersection rather than at the one where the priority measure is carried out. These impacts are, however, possible to study using simulation. Microscopic traffic simulation connected to a signal controller simulator is often applied to evaluate dynamic changes in signal timings such as PTSP.
Many case studies of PTSP in coordinated systems can be found in literature, both field studies (Furth & Muller 2000; Al-Mudhaffar & Bang 2006; Albright & Figliozzi 2011) and simulation experiments (Wahlstedt 2005; Burghout & Wahlstedt 2007; Zlatkovic, Martin & Stevanovic 2009; Altun & Furth 2009; Wahlstedt 2011; Christofa & Skabardonis 2011). The effects of different priority measures (green extension, red truncation, etc.) have been studied at isolated intersections as well as the total effects of active PTSP in coordinated systems with several intersections. Green extensions are reported to be less disturbing to other traffic than red truncation or stage rotations (Furth & Muller 2000). However, there are few studies of the impacts of different priority measures within a coordinated system. The impacts of PTSP within a coordinated signal system are often different than that of an isolated signal, since vehicles tend to arrive in platoons rather than randomly considering that there is a coordinated traffic signal upstream.

1.2. Objectives
The overall aim of this thesis is to enhance the knowledge regarding effective strategies for conditional priority of public transport in traffic signals; i.e. how to give better benefits to public transport while minimising negative effects to other traffic.

Sub-aims, included in the licentiate part of the work, in order to achieve this are:

- Review of criteria for and fundamentals of public transport priority in traffic signals (PTSP)
- Analysis of the impacts of different PTSP priority functions/strategies in a coordinated signal system based on simulation trials
- Study of the possibilities to use conditional signal priority to avoid bus bunching

1.3. Scope
The licentiate thesis is focusing on strategies and systems for PTSP in general. The impact study is limited to conventional coordinated signalling systems currently used in Stockholm, and to self-optimising signal control systems.
1.4. Historical background from a Stockholm perspective

The world’s first traffic signal was installed in London 1868 with red and green semaphore arms and gas lamps for night use. After a few days of operation a gas-explosion occurred and the signal was taken away. It then took until the 1910s before traffic signals were reintroduced, this time in USA (Webster & Cobbe 1966). The first traffic signal in Stockholm was installed in 1925 at the intersection Vasagatan – Kungsgatan, four two-colour signals were set up in the street corners and controlled by a policeman at one of the corners (Dufva 1985). With the development of technology, the signals became automatic, fixed timed or vehicle actuated, and some signals where coordinated.

In a research project bus priority was tested in an isolated controlled intersection in central Stockholm 1972. A self-optimising signal control strategy including bus priority was tested in 1974 (Bång 1975). During the 70’s political demand emerged for priority of buses in traffic signals in Stockholm. Bus priority was tested in 1977 in a number of conventionally controlled intersections along the route Hantverkargatan – Tegelbacken – Norrmalmstorg. The tested priority included green extension, red truncation and extra bus stages; travel time reductions of five to eight per cent were reported. Bus-shaped inductive detectors, which were able to distinguish buses from cars, were used (Dufva 1985).

In the 1980s and 90s the centralised control centres for coordinated traffic signal systems were replaced with computerised signal controllers at each intersection. This new generation of signal controllers allowed more local vehicle actuation within the pre-defined coordinated signal plans, and made bus priority easier to implement. A bus priority method named “PRIBUSS” was developed by the City of Stockholm, see Section 5.1 for further description. After successful tests the PRIBUSS was implemented in a large number of intersections in central Stockholm. The buses operating on the inner city lines were equipped with transponders communicating with in-street detectors to request signal priority.
In the early 2000:s a trunk route bus-network with five lines operating with specially branded articulated buses was introduced in Stockholm. It was decided to only prioritise the five trunk lines in order to reduce the impact on other traffic, and to minimise the problems with conflicting requests for priority where bus lines intersect. A radio based detection system with AVL equipment in the buses on the trunk lines replaced the transponder system, and bus priority was removed from intersections outside the trunk bus network (Wahlstedt 2005).
2. Overview of traffic signal control methods

2.1. Introduction
The purpose of traffic signals is to increase capacity, traffic safety and/or give a fair accessibility to different road users by separating conflicts between crossing traffic streams in time. Under favourable conditions it is also possible to coordinate adjacent traffic signal controlled intersections in order to create “green waves” through a number of intersections. Furthermore it is also possible to prioritise PT in traffic signals, in contradiction to most other forms of intersection control. Traffic signals can be divided into isolated or coordinated controlled, and into fixed timed or vehicle actuated controlled. With isolated controlled signals the green periods and cycle time are based only on the conditions in the specific intersection. With coordinated controlled signals the green periods and cycle time are set also considering adjacent traffic signals, coordinating the green periods to facilitate passage through the system of coordinated traffic signals.

With fixed timed (FT) signals the green periods and cycle time are fixed and predetermined based on historical data. With vehicle actuated (VA) signals the green periods are related to the traffic demands using detectors of some kind (Webster & Cobbe 1966). Conventional VA traffic signals works with green demand and extension of on-going green based on the detections, and hence the cycle time will vary. Self-optimising traffic signals (sometimes called adaptive traffic signals) uses a mathematical algorithm to optimise the cycle time and green periods for all approaches based on the detected and/or predicted traffic conditions (Bång 1975 and 1976, Al-Mudhaffar 2006). Self-optimising traffic signals are actually a special form of vehicle actuated traffic signals, and unlike conventional VA signals it is possible to make self-optimising signals coordinated in larger areas.

Nowadays Swedish traffic signals are usually isolated-traffic actuated or coordinated-fixed timed. Isolated-fixed timed sometimes occurs when detectors are hard to place and maintain due to construction works etc. or as temporary solutions. Conventionally controlled coordinated signals need to operate with common cycle time and mostly pre-set timing and offsets. Some vehicle actuated local timing adjustments can be included within the fixed cycle,
but will in this thesis be considered as coordinated-fixed timed signals. Self-optimising signals are unusual in Sweden today, but are more common in some parts of the world.

2.2. Criteria for signal control
There can be many goals or criteria to optimise the design and settings of traffic signals. The most common are maximum capacity, minimum average delay or cost also including the impact of stops and emissions. Local regulations and traffic safety concerns also need to be considered, and can limit the possible design and potential to achieve other goals. In the Swedish LHOVRA control strategy (Al Mudhaffar 2006) traffic safety is of great importance and is allowed to increase minimum delay and number of stops if necessary. However the goals and criteria for optimal control are seldom clearly stated and local tradition and best practice is often applied when making the design. Traditional control strategies also often use heuristic methods to achieve the aimed goals, making it hard to balance between the different criteria.

2.3. Isolated control
More than 60% of signalised intersections in Sweden operate with isolated control at all times, and almost all traffic signals operate with isolated control at night.

Isolated controlled traffic signals base the signal timings only on the traffic conditions in their approaches. The arrival of vehicles is generally assumed to be random, with a negative exponential headway distribution:

\[ f(h) = q e^{-qh} \]  

(2-1)

Where:
\( h = \) headway
\( q = \) traffic flow

The performance of isolated traffic signals with random arrival in general, is possible to evaluate analytically.
2.3.1. Fixed time isolated control
With FT controlled traffic signals the green periods and cycle are fixed and preset based on historical- or estimated traffic flows. Although isolated fixed timed controlled signals are not very common, they are at the base of signal control theory.

The capacity of an approach to a signalised intersection depends on the green time available, and the number of vehicles able to pass the stop line per green time unit. The latter is normally expressed as saturation flow: the maximum number of vehicles that can pass the stop line under saturated, steady conditions. The discharge rate is lower during the first few seconds, when vehicles start to accelerate and during amber period when some vehicles stop and some carry on (Webster & Cobbe 1966), see Figure 2 below. The saturation flow includes those variations as well as lower discharge rate when yielding to pedestrians or other vehicles. In practice the saturation flow is both time dependent and site-specific, but experiential standard values can be used for calculations.

![Figure 2: Variation with time of saturation flow in a fully saturated green period (Webster & Cobbe 1966)](image)
The red-amber, green, amber and all red periods displayed by the signals can be replaced by "effective green", "effective red" and/or "lost time" periods in capacity calculations. During the effective green period the traffic flow is assumed to take place at the saturation rate. The effective green equals the actual green period minus the start-up loss plus the gain from the part of amber period used as green. Effective red is the part of the cycle not to be considered as effective green. Lost time is the part of the cycle where no approach has effective green.

Webster (Webster 1958) by field studies and simulation found that the interaction delay on any single approach to a signalised FT intersection can be expressed as:

\[
d = \frac{c(1-\lambda)^2}{2(1-\lambda x)} + \frac{x^2}{2q(1-x)} - 0.65 \left( \frac{c}{q^2} \right)^{\frac{1}{3}} x^{(2+5\lambda)}
\]  

(2-2)

Where:
- \(d\) = average delay on the arm
- \(c\) = cycle time [s]
- \(g\) = effective green time [s]
- \(\lambda\) = proportion of the cycle, which is effective green for the considered stage (\(\lambda = g/c\))
- \(x\) = degree of saturation (\(x = q/\lambda s\))
- \(l\) = lost time [s]
- \(s\) = saturation flow [vehicles/hour]

The first term \(\frac{c(1-\lambda)^2}{2(1-\lambda x)}\) represents the deterministic part of the delay at constant arrival rate and saturation flow, see Figure 3 below. The second term \(\frac{x^2}{2q(1-x)}\) represents the additional delay caused by random variation in arrivals and saturation flow. Those variations cause temporarily oversaturation when the traffic flows are close to the intersection capacity. The third term

\[-0.65 \left( \frac{c}{q^2} \right)^{\frac{1}{3}} x^{(2+5\lambda)}\]

is an empirical addition explained by irregularities in arrivals and saturation flow.
Figure 3: The deterministic part of the delay function (2-2)

Derived from the delay equation (2-2) Webster (1958) stated that the optimal cycle time for a FT controlled signal giving the least delay for all traffic, can be calculated as:

\[ C_o = \frac{1.5L + 5}{1 - Y} \]  

(2-3)

Where:
- \( C_o \) = optimal cycle time
- \( L \) = total lost time per cycle
- \( Y = \sum y_{max} \)
- \( y_{max} \) = largest \( y \) per stage
- \( y = \text{ratio of flow to saturation flow} \ (y = \frac{q}{s}) \)

1,5 and 5 are empirically determined constants, at “normal” conditions

The optimal cycle time will in practice, be limited by minimum green periods for pedestrians and maximum acceptable waiting time, and the practical optimal cycle time would be in the range of 25 – 120 seconds.
Webster also found that cycle times within the range of three-quarters to one-and-a-half times the optimum value the delay is within 10 – 20% higher than for the optimal cycle time (Webster & Cobbe 1966). As can be seen in Figure 4 below, the delay increases slowly with longer than optimal cycle time, and increases rapidly with shorter than optimal cycle time. The rapid increase in delay with short cycle times is due to over saturation at the intersection.

Figure 4: Effect on delay of variation in cycle time (Webster & Cobbe 1966)

Further, it can be derived from equation (2-2) that in order to minimise the overall delay, the ratio of effective green times (green split) in the approaches to an intersection should be set equal to the ratio of the y values:

\[
\frac{g_1}{g_2} = \frac{y_1}{y_2}
\]  

(2-4)

Where:
\(g_1, g_2, \ldots\) = effective green time in stage 1, 2, ...
\(y_1, y_2, \ldots\) = the highest y value in stage 1, 2, ...

For stage i it can also be expressed as:

\[
g_i = \frac{y_i}{\bar{y}} (c_o - L)
\]  

(2-5)
As can be seen in Figure 5 above, the delay is much more sensitive to optimal green time ratio than to optimal cycle time. With shorter cycle times the delay is even more sensitive to optimal green split due to oversaturation of the approach with too short green time.

2.3.2. Vehicle actuated isolated control
With VA controlled signals the green periods and thereby cycle time are variable, within pre-set limits, and related to the traffic demands using detectors. Most isolated controlled traffic signals in Sweden are vehicle actuated.
In principle, conventional VA controlled signals use detectors to recognise:

- Demand for a green phase
- Demand for extension of green

The signal will only change to green if there is a vehicle present in the approach demanding green. When the signal changes to green there is normally a short minimum green time, and after that the green time can be extended up to a maximum green time if there are further vehicles present in the approach (Kronborg, Davidsson & Edholm 1997). When there are no more vehicles in the approach, or the maximum green time is reached, the signal shifts to the next stage. The stages shift in a fixed sequence, but if there is no green demand for a stage, it will be skipped and the signal changes to the next stage with green demand.

In order to obtain the least overall delay, the signal should change stages when the queue accumulated at the red signal has been dispersed (Webster & Cobbe 1966). When this occurs, the time headways between the vehicles will increase and the discharge rate will drop. The increased headway is used to determine the optimal time to change stage in conventional VA control strategy. With the gap extension control strategy, the detector in the approach extends the green time only if the gap between the vehicles is smaller than the pre-set extension interval at the detector. In practice it is not simple to determine exactly when the queue is dispersed, so the extension interval needs to be a little longer than the distance between two vehicles at saturation flow.

Older VA signals usually had only one detector in each approach at some distance from the stop line. Nowadays there are usually two or more detectors in each approach; one at the stop line and one or more detectors at some distance. Each detector has a “detector interval”, or a time period in which the detector will be “active” after being actuated by an approaching vehicle. The detector interval should be long enough for the vehicle to reach the next detector in the approach, or to reach the stop line from the last detector. In this way the detectors work together to form the extension interval. The extension interval of the detectors working together can result in longer than ideal intervals for minimum delay. The motive for this is to reduce the number of stops. Accident reduction functions further add to the extension of green times and thereby, cycle times.
VA signals can be analysed analytically in only some special simplified cases, in other cases simulations are necessary (Bång 1975, 1976). According to simulations of a four-arm two stage intersection made by Webster (Webster & Cobbe 1966) the performance of VA control is similar to, and only slightly better than FT control optimised for the actual traffic conditions. If the extension intervals are too long VA will perform even worse than FT, see Figure 6 above. This implies that the FT signal timings is optimal, and in real life, the FT signal can of course not be constantly optimised and will consequently perform worse.

The HCM2010 includes an analytical model to calculate the impacts of a simple VA signal based on the probabilities for green demand and for extensions (HCM 2010). The model requires extensive calculations and is rarely done with pen and paper.
2.4. Coordinated control
With coordinated controlled signals, the green periods and cycle time are set considering adjacent signals, coordinating the offsets for the start of green periods to facilitate passage through the system in the main directions. With some exceptions, this implies a common cycle time for all coordinated signals. With conventional control strategies, the common cycle time means that the signals need to be mainly FT controlled, but some local traffic adaptation within the fixed cycle time is possible. This can include setting signal groups to green on demand only, green extensions or other functions as long as their overall influence is not too large.

Figure 7: Time-space diagram of a hypothetical system of coordinated signals showing the "green waves"

The signals along the coordinated section can be set to change simultaneously, but in most cases the signal changes are made progressive. The start timing (offset) of green periods are determined so that a vehicle traveling at the design speed of the coordination meets green at arrival at the next intersection. This creates a "green wave" for the platoon of vehicles released from the first
coordinated signal. It is not always possible to achieve a good green wave in both directions at reasonable cycle time; this is dependent on the distance between the intersections, and the minimum time needed for the side streets and pedestrian crossings.

If the street network is a grid where the signals should be coordinated in all directions, the green wave speed, cycle time and the intersection distance will be interrelated. The design of the coordination then needs to be a trade-off between cycle time and travel speed (Webster & Cobbe 1966).

The relation between intersection spacing, cycle time and resulting green wave speed can be expressed as:

$$D = v \cdot \frac{c}{2}$$  \hspace{1cm} (2-6)

Where:

- $D$ = intersection distance
- $v$ = green wave speed
- $c$ = cycle time

If the street network to be coordinated is not a grid, but a primary street with crossing side streets, then the possibility to obtain green waves is greater. The design speed of the green wave can be set to the desired speed of the drivers. The cycle time in the coordinated system of signals is normally determined by the intersection with the highest cycle time demand for isolated control, but under unfavourable conditions, an even longer cycle can be needed.

There can be different optimisation objectives when designing signal coordination:

- Minimise overall delay
- Minimise delay per vehicle
- Minimise number of stops
- Maximise the bandwidth of the green wave
- Improve the accessibility on the main road

The objectives are to a large extent shaped by local tradition and policy in different cities and countries.
The impacts of coordinated signals are theoretically possible to calculate analytically, but the calculations will be extensive since the vehicle arrivals are not random but dependent of the green periods for both primary and secondary streets in upstream intersections. In many cases, a simple simulation model is more suitable for this task. The design of the coordination can be made manually or with help of some coordination optimisation software. The optimisation methods can broadly be categorised into bandwidth maximising methods, and disutility (delay, stop etc.) minimising methods (Little, Kelson, & Gartner 1981). Several commercial coordination optimisation software are available, for instance TRANSYT, LinSig etc. (Binning, Burtenshaw & Crabtree 2009, Moore 2011).

Optimisation tools, such as TRANSYT, minimise a cost function by using a traffic simulation model and a signal time optimisation procedure using a hill climbing process, see Figure 8 above (Binning, Burtenshaw & Crabtree 2009). The cost function typically includes stop and delay for all or specified links.
2.5. Self-optimising control

Self-optimising control strategies aim to adapt signal changes to traffic flow demand in real time. Instead of detecting green demand and gap extension, as in conventional VA control, they optimise the signal changes by minimising a cost function for the calculated/predicted impacts of the signal timings. The cost function normally includes weighted costs for delay, stops etc. Self-optimising strategies can be used with isolated as well as coordinated controlled signals. Bus priority is often included directly in the optimisation process; a bus can be given higher weight than cars in the cost function. Since the bus priority is included in the optimisation the impacts can be balanced out before or after the bus has passed. Most self-optimising signal control systems count the vehicles at large distance upstream of the stop line and predict their arrival at the intersection by using a traffic simulation model as a part of the optimisation process (Tveit 1999, Al-Muhaffar 2006, Arveland 2006). In some respect they can be described as simulation based strategies.

Figure 9: Delay gained (1A) and caused (1B + 2) by an extension to be included in the control function \( \phi \) (Vincent & Pierce 1988)
The most used optimisation algorithm was first proposed by Miller (1963). At frequent intervals the predicted cost of changing stage is calculated and compared to the predicted cost of changing stages after an extension interval of h, 2h, 3h... The stages are changed immediately if the cost is lower than any later change, otherwise the stage is extended. The cost is re-calculated with new updated input data several times a second, and a new decision is made after each re-calculation of the cost of change stages immediately or not. The control function ($\phi$) calculates and weights the gain of the vehicles able to pass by an extension with the delay caused to the vehicles waiting longer, see Figure 9. For the stage change from North – South to East – West the control function can be expressed as:

$$\phi = \left[ \delta_N + \delta_S - q_N \frac{h - \delta_N}{s_N} - q_S \frac{h - \delta_S}{s_S} \right] * \left[ r_{NS} + l_{NS} \right] - h * \left[ n_W + n_E + \sum_{i=1}^{k_W} q_{W_i} + \sum_{i=1}^{k_E} q_{E_i} \right]$$

(2-7)

Where:
- $\delta$ = number of additional vehicles discharged if the stage change is postponed
- $q$ = traffic flow
- $h$ = the extension interval for which the control function is updated
- $s$ = saturation flow
- $r$ = effective red time
- $l$ = lost time at stage change
- $n$ = number of queuing vehicles
- $N, S, E & W$ = North, South, West & East approach
- $k$ = for each approach the smallest integer satisfying:

$$\left[ n + \sum_{i=1}^{k} q_i - \sum_{i=2+1}^{k} s_i \right] \leq 0$$

The first term in in equation (2-7) represents the reduced delay for vehicles traveling N-S able to use the extension, minus the extra delay for vehicles traveling N-S arriving after the extension has ended. Those will receive green signal later due to the extension. The second term represents the extra delay caused to vehicles traveling E-W due to the extension (Bång 1975). A drawback with the Miller algorithm might be that the calculations are done for one cycle only, and problems tend to be postponed (Kronborg, Davidsson & Edholm 1997).
The Miller algorithm was enhanced for practical implementation and tested through simulation and field studies at two intersections in Stockholm (Bång 1975, 1976). Bång put cost on the delay and added the cost of stops to the algorithm, and put in extra terms for buses and pedestrians. By giving the terms for vehicles, buses and pedestrians, different weights buses could be prioritised in the control strategy. The developed control strategy for isolated control was called TOL, Traffic Optimization Logic, and showed good results compared to conventional control in the field tests (Kronborg, Davidsson & Edholm 1997).

Self-optimising control strategies can also be used with coordinated control; the algorithms are inspired by the Miller and TOL algorithms, but are more complicated in order to handle coordination (Kronborg, Davidsson & Edholm 1997). One of the first commercial self-optimising systems for coordinated controlled signals was SCOOT, Split Cycle Offset Optimisation Technique developed in the UK (Arveland 2006). SCOOT is one of the most widely used self-optimising systems worldwide. It has a centralised structure with a central computer optimising the whole area communicating directly with every signal controller. The algorithms are developed from the ones in TRANSYT off-line coordination optimisation software.
The traffic flow profiles are counted at the exit of the intersections or a few hundred meters upstream from the first signalised intersection controlled by SCOOT, and the movements and queues are simulated similar to TRANSYT. The cycle time is optimised every five minutes and is kept the same (or a multiple) in the whole SCOOT area. Offsets are optimised every cycle and green splits a few seconds in advance of every stage change. SCOOT only makes small changes from the previous cycle in order to make the system stable. There are some, but compared to other systems limited possibilities of active bus priority in SCOOT. The emphasis in SCOOT is put on minimising the overall delay, according to British tradition (Arveland 2006).
Another system design is used by the Italian system Utopia/Spot. Here the optimisation is distributed to the intersections, which communicate directly with each other instead of with a central. Utopia/Spot optimises the signal timings by minimising a cost function inspired by the Miller algorithm locally in a Spot unit in each intersection. Spot units in neighbouring intersections communicate counted/predicted vehicle arrivals and predicted signal changes for the next minutes. Neighbouring intersections constantly adapt to the predicted vehicle arrivals from surrounding intersections by including this data in its time dependent cost function. A lower cost is generated if the arriving traffic for a particular direction is faced with green rather than red. This way coordinated green waves for platoons will be created most times, but not if causing a higher cost for some reason (Tveit 1999). In contradiction to the systems mentioned above a common cycle time is not necessary, the cycle time and coordination is decided in each controller and each cycle based on the cost function (Zlatkovic, Martin & Stevanovic 2009). The flexible cycle time give Utopia/Spot good possibilities to prioritise buses and balance other traffic streams before or after the bus passage (Kronborg & Davidsson 2005). Other decentralised self-optimising control strategies such as ImFlow work in a similar way.

Most self-optimising systems are stage based, and aim to optimise the overall signal changes rather than optimise the signal changes on when individual vehicles passes the stop line as VA control tend to do. The British stage based self-optimising MOVA strategy was tested and compared to signal group based conventional control with Swedish LHÖVRA strategy in field tests 1992. The results for LHÖVRA and MOVA where about the same, the gain from the optimisation based strategy was lost on lower flexibility due to stage control in MOVA (Kronborg 1992). Later a signal group based self-optimising strategy for isolated control "SOS" was developed and tested at the same site in 1995 - 1996. SOS worked well, particularly in reducing the number of vehicles in the dilemma zone at signal change, but was never commercialised (Kronborg, Davidsson & Edholm 1997).
3. Methods for bus priority

3.1. The need for bus priority
The configuration and timing of signalised intersections as well as physical design of streets are often optimised to minimise average delay for all motor vehicles. However, since buses and trams normally carry much higher number of passengers this will not minimise the delay per person. Public Transport Vehicles therefore need to be handled differently in order to minimise the overall delay per person. The bus lines sometimes use other streets than the main routes prioritised in the street network, and have a different speed profile due to bus stops compared to the design speed of signal coordination made for cars. Therefore buses and trams need to be specially prioritised in urban traffic to minimise the delay per person.

There can also be political reasons to “over compensate” buses and trams in order to promote travel by public transport instead of by car to reduce pollution and congestion. The City of Stockholm has adopted the “Urban Mobility Strategy” stating that the more surface efficient modes of mobility, such as public transport, cycling and walking should be prioritised over cars in central Stockholm (Firth 2012).

There are many ways to prioritise buses in the urban environment: physical measures such as bus lanes and good bus stop locations, and different types of bus priority in traffic signal control as described below.

3.2. Passive traffic signal priority
PT traffic signal priority (PTSP) can be passive or active. In passive PTSP, signal timings can be set in favour of approaches used by buses by giving them more than proportional green time. The speed of the progression in coordinated signals can be adapted to typical bus speed that includes dwell-time at bus stops. In a conventional FT coordination without bus priority there is a risk that the bus will meet a “red wave” if the “green wave” speed is too different from the travel pattern of the bus. The bus will “fall out” of the green wave when halting at a bus stop, see Figure 11.
If the green wave speed is adapted to the bus travel pattern other traffic will experience longer delays instead. Passive PTSP is often good when PT frequency is high and the dwell-time at stops is short and predictable (Smith et al. 2005). Since passive priority is always active the possible negative impacts on other traffic will occur also when no bus is present.

Figure 11: A bus falls out of the green wave optimized for car speed when halting at a bus stop (Björk & Dahlgren 1991)

3.3. Active traffic signal priority
Active PTSP triggers priority measures when there is a bus present, thus requiring the selective detection of PT vehicles. The signal timing can be changed by extending green or truncating red (see below) when a bus is detected in order to let the bus pass through the intersection with minimum delay (Smith et al. 2005). The priority measures will only be activated when a bus is present and can take advantage of it, thus active priority can disturb other traffic much if beneficial for the bus, and still be effective in reducing overall delay. Therefore active priority measures can give larger benefits to buses compared to passive measures.
The possibilities to provide good signal priority to the bus depend on how far in advance the request for priority is made. An upstream intersection or bus stop can make a reliable priority request difficult in advance. Other factors limiting the possibilities for active bus priority are:

- Long minimum green times for pedestrian crossings across the bus street
- Highly saturated intersections where bus priority may cause oversaturation and long queues that also affect the buses
- Buses on both main and side street causing conflicting requests for signal priority
- High bus frequency

Active bus priority is easier to arrange in isolated controlled signals compared to coordinated controlled signals where the “green waves” may be disrupted by the priority. Under isolated control, compensations to traffic streams affected by the priority are more easily arranged in the cycle after the bus has passed.

3.3.1. Conditional and unconditional priority
Active signal priority can be conditional or unconditional. Unconditional priority will always favour the bus without considering negative impacts for other road users, while conditional priority is restricted (Smith, Hemily & Ivanovic 2005). Unconditional priority is often used for trams only, or in isolated controlled intersections without capacity problems (Bång 1987). Conditional priority will try to restrict the negative impact on other road users by limiting the extent or number of priority measures. Some criteria are set for giving priority or not. Those criteria can be with respect to the buses, or with respect to the other traffic. Criteria for PTSP are further discussed in chapter 4.
3.3.2. Conflicting priority requests
There are several reasons for conflicting signal priority request to occur; there could be two bus lines crossing each other at an intersection with buses on both lines arriving at the same time requesting signal priority in different directions. Also buses on the same line can request for conflicting priority functions if the line makes a turn in the intersection, and buses in either direction therefore are served by different stages. Even two buses served by the same stage can cause conflicting requests for signal priority if they meet (or follow) at an unfavourable distance from the intersection, making it impossible for both buses to pass the intersection at the same green period.

Those conflicting requests for priority need to be handled by the PTSP strategy. With conventional control normally heuristic rules are used, first-come first-served is used in Stockholm. Algorithms prioritising green extension before red truncation when conflicting priority requests occurs have showed better results than first-come first-served (Zlatkovic, Stevanovic & Martin 2012). With self-optimising strategies both buses with conflicting PTSP requests are included in the optimisation algorithm, and the overall most beneficial signal timing can be chosen.

3.3.3. Conventional control strategies
Active bus priority in traffic signals with conventional control strategies works by changing the normal FT or VA signal timings when a bus is detected. The changes are typically extending green stage, truncation of red stage, skipping a stage and/or inserting an extra stage for the bus when a bus is detected (Smith, Hemily & Ivanovic 2005). This is relatively simple to arrange in an isolated controlled signal, but with coordinated controlled signals the common cycle-time and the coordination need to be retained after the bus has passed. A closer description of the PRIBUSS method for active bus priority in conventionally controlled signals is found in chapter 5.1 below.

The bus priority is often accomplished through local adjustment of signal timing by the signal controller in each intersection independently. In order to be able to return to the pre-set signal coordination, only small enough green time changes can be allowed so normal signal timings can be resumed in one or a few cycles. With larger green time changes stages can be skipped in cycles with bus priority.
Green extension can result in a large reduction in travel time for a few buses while early green ca results in a small reduction for many buses (Bång 1987). It may also be necessary to make green time compensations to other approaches after the bus has passed, without changing the cycle time. However, compensations made locally in the signal controllers may have negative impacts on the performance of the coordinated system if not carefully timed.

### 3.3.4. Self-optimising control strategies

Self-optimising control strategies predict the arrival of vehicles and their movements through the system. The optimisation is made in real time by minimising a cost function, in most cases a weighted combination of stops and delays for all vehicles. Buses can be selectively detected and given a higher weight in the cost function e.g. 20 times a car. The resulting signal timings will be more favourable for the bus if the total cost for other traffic is not too high. Different routes, directions or late buses etc. can be given different weights and thereby prioritised to different extents. In some self-optimising systems, e.g. BUS SCOOT, central optimised bus priority can be combined with conventional active bus priority directly in the signal controller overriding the centrally optimised signal timings.
3.4. **Bus detection**

The detection of the buses required for active signal priority can be made using different technologies including:

- Inductive detectors in the size of a bus
- Communicating transponder in the bus and roadside detector
- Radio based systems with AVL system in the bus

![Figure 13: Different techniques of bus detection, bus-shaped inductive loop, radio based with AVL system and communicating transponder (Wahlstedt 1997, Fäldt 1992)](image)

Inductive detectors with amplitude selective bus detection consist of an ordinary inductive detector in the size of a bus. A vehicle of the same size as the detector will give a stronger pulse, and a bus can thereby be distinguished from other vehicles. No extra equipment is needed in the bus, but the bus needs to be driven closely over the detector. With this technique reliable out-counting at the stop line is often not possible. Start of priority by in-counting and end of priority by a pre-set time-out is often used.
A transponder on the bus allows one or two-way communication between the bus and the signal controller enabling the system to distinguish between different bus lines, etc. The bus is equipped with a transponder communicating with a roadside unit when passing the receiver. The road side unit can be combined with the normal inductive detector loop connected to a special receiver unit, a separate detector loop or another receiver unit under or above ground. Several commercially available systems exist, using different techniques including inductive, infra-red light, microwaves and RFID. The detections are reliable and the vehicle position exact, but the road side units can be expensive to maintain. Usually in-counting at a distance of a few hundred meters and out-counting at the stop line is used to start and end the priority. In Sweden transponder based systems are used by trams and in Gothenburg also by buses. Stockholm earlier used transponders on buses, but has changed over to a radio based system due to high maintenance costs on the transponder detectors (Wahlstedt 1997).

Figure 14: Principal placement of in-counting and out-counting detectors
Radio based systems are becoming more common. The bus is equipped with an automatic vehicle location (AVL) system that keeps track of location, usually by a combination of GPS and odometer counting wheel turns. At pre-defined positions the bus sends a signal by radio to call for signal priority, and to end the priority after passing the stop line. Some systems use short-distance radio for communicating directly between the bus and the signal controllers, while other systems use a central server to handle all communication including calls for priority. Radio based systems requires less road side equipment and the detection points are easy to move, but there is often some uncertainty in the vehicle's position making the end of priority less effective. The AVL system also often also includes schedule data and is used for real time passenger information systems on board and at stops as well as for collection of statistical data. This information can also be used for conditional priority based on actual bus arrivals compared to time table accordance. A radio based system with direct communication between the bus and signal controllers is used in Stockholm.

Effective bus priority requires early detection of the bus, i.e. the detector/detection point needs to be placed at enough distance from the signal. An upstream bus stop is often limiting the possible distance to the detector/detection point since the uncertainty in dwell time at the stop prevents detection upstream of the nearest bus stop. Location of the bus stop after rather than before the intersections is therefore preferable if active bus priority is used.
4. Criteria for PTSP

4.1. Criteria for introducing PTSP
The motive of introducing PTSP is usually to reduce the bus travel time, or overall travel time for all road users. A certain intersection, the (coordinated) area or the traffic system could be considered. The effects to the system are especially relevant for PT as the service reliability is influenced also further along the PT line. If the system-wide bus travel times or the overall travel times should be reduced, some restrictions need to be set on the PTSP. The PTSP would be conditional as described in chapter 3.3.1. The criteria for the conditional priority are depending on the overall motive of introducing PTSP.

4.2. Minimising total travel time with conditional PTSP
If the overall travel time should be reduced the PTSP impacts on both buses and other traffic, pedestrians etc. need to be considered (Bång 1987). This means that the number and/or extent of the priority measures need to be restricted by setting criteria for giving priority or not. Those can be set with respect to the buses, with respect to other traffic or a combination thereof.

Criteria with respect to the buses are set in order to give priority only to the vehicles needing it most, and thereby reducing the number of priority measures. Examples of criteria with respect to the buses are:

- only give priority to specific bus lines
- only give priority in peak direction of travel by time of day
- only give priority to late buses/not to early buses

A combination of those criteria can be used. Different extents of priority can also be given based on the mentioned criteria. Restricting the priority to the buses needing it most also give those buses greater chance to gain priority, since giving priority to one bus also means less chance to give priority to another bus (Ma, Yang & Liu 2011). Conditional PTSP can thereby reduce the total travel time for bus passengers, if the PTSP criteria are set with respect to the buses.
Criteria with respect to other traffic are set to restrict the negative impacts on other road users, by limiting the extent of the PTSP. With conventional control strategies, this is made by setting time limits for green extensions, guaranteeing a minimum green times to un-prioritised streets, setting a minimum time between two priority measures in the intersection etc. (Wahlstedt 2005). The PTSP can also be restricted to uncongested time periods, or disabled by queue detectors if the effect on other traffic results in too long queues. With self-optimising control strategies the PTSPs negative impacts on other road users are included directly in the optimisation process (Bång 1975). The gains and losses for other road users are balanced with the gain for buses in the optimisation algorithm. Additional penalties can be set on blocking-back upstream intersections, and the weight on buses compared to cars can be adjusted to achieve the desired effect it the area.

4.3. Improving PT service reliability with conditional PTSP
Conditional priority where the criteria to give priority are set with respect to the buses can help to improve the service reliability of the bus line (Furth & Muller 2000). The criteria can be based on on-time performance, time from previous bus on the line or deviation from headway. The system effects of improved service reliability can reduce the total travel time for bus passengers also far away from the actual traffic signal.

The time table adherence can be improved with conditional PTSP, if priority is given to late buses (Albright & Figliozzi 2011). With conditional priority and criteria based on on-time performance the regularity can be improved (Kim, Park & Chon 2005). Time table adherence is important for bus lines with long headways, whilst constant headways are more important for bus lines with short headways.

Bus bunching is a well-known problem on high frequency bus lines, and conditional signal priority based on time table adherence can reduce bunching, especially in combination with appropriate holding strategies e.g. (Cats et al. 2011). The potential for avoiding bunching by smooth “holding” and “pushing” with conditional PTSP is higher in the beginning of the line, since the needed headway corrections is much smaller before the positive-feedback bunching phenomena has escalated. A conditional PTSP strategy focusing on avoiding bunching consequently needs to be direction dependent.
Number of passengers aboard the bus has been suggested as criteria for priority, but no on-street implementation of such strategy have been found in literature so far. However, passenger load aboard the bus have been tested as criteria for priority with simulation experiments in Stockholm (Morgan, Koutsopoulos & Ben-Akiva 2004).

Even headways to the previous and the next bus have also been proposed criteria for future implementation (Cats 2011).

4.4. Differential PTSP
The benefit of receiving priority will not be the same for all buses, especially if the system effects are considered. A differential priority strategy giving more priority to some buses will therefore give larger total benefits. There is also a need to handle conflicting requests for priority, see section 0 above. With conventional control strategies differential priority is complex to implement. However it is possible to restrict certain priority functions e.g. green extension, red truncation, extra stage etc. to some buses only. This is discussed in paper III where a new on-time-status based differential priority strategy is proposed and tested.

With self-optimising control strategies it is possible to make differential priority by giving different lines different weights, and thus give each bus different weight in the optimisation process. In Trondheim (Norway) such strategy is implemented in a Utopia/Spot installation. In this implementation bus lines with higher average number of passengers have higher weights, as well as buses in peak direction. When the bus requests priority, the weight based on bus line and direction of travel is multiplied with a factor depending on the lateness according to timetable from the bus’s AVL system. The resulting weight is put into the optimisation algorithm of Utopia/Spot.
5. Examples of systems for bus priority

5.1. PRIBUSS

The bus priority method PRIBUSS (Swedish acronym for Prioritising of Buses in Coordinated Signal systems) could be described as a toolbox of bus priority procedures for the traffic engineer to use when designing a traffic signal with active bus priority. It was developed by the city of Stockholm and TFK in the early 1990s. PRIBUSS is now included as standard in most signal controllers on the Swedish market (Wahlstedt 2005). PRIBUSS was primarily developed for coordinated control, but can also be applied for isolated signalised intersections (Björk & Dahlgren 1991).

PRIBUSS was developed for conditional bus priority on top of the normal mainly fixed-time coordinated control, or vehicle actuated isolated control. The engineer decides which PRIBUSS procedures, conditions and limitations to be applied by parameter programming, and is given large freedom to decide how extensive the priority should be (Björk & Dahlgren 1991). The priority and compensations are made locally in each signal controller and neighbouring intersections are not considered. It is up to the engineer to limit the PRIBUSS functions so that the negative impacts on the coordination are acceptable. There is no systematic method for prioritising between buses with conflicting priority calls within the PRIBUSS method, the first come – first served principal is used if conflicting calls for priority occurs.

PRIBUSS consists of six standardised procedures for bus priority:

- Extension: extends on-going green
- Re-taken start: return to green if conflicting groups have not yet have become green (e.g. under inter-green and red-amber)
- Early green: cuts the on-going stage, with optional limitations
- Extra stage: inserts an extra stage in or between the normal stages
- Double early green: cuts to subsequent stages
- Double extra stage: cuts to subsequent stages and inserts extra stage

In the signal controller the PRIBUSS programming is built up by a number of PRIBUSS functions starting one of the PRIBUSS procedures mentioned above.
Each PRIBUSS function has a number of conditions programmed, and if all conditions are met the function is started. When one PRIBUSS function is started all other is blocked until the first one is ended. There could be more than one PRIBUSS function with different conditions starting the same type of PRIBUSS procedure, but with different settings.

Each PRIBUSS function is conditioned by a “window” in cycle time, the status of certain signal groups, and conditioned to one or many time plans. The conditions ensure that the appropriate PRIBUSS procedure is chosen at any time. A guaranteed green time for conflicting signal groups could also be given before the PRIBUSS function is allowed to start. There is also a set of similar conditions deciding whether to make compensations or not after the priority is ended (Björk & Dahlgren 1991).

![Flowchart for PRIBUSS](image)

Figure 15: Flowchart for PRIBUSS (Björk & Dahlgren 1991)

The PRIBUS function will be ended by what occurs first of: out-counting of all in-counted buses, elapsed time-out time or last cycle-second reached. The last cycle second can, but must not, be specified in the design in order to avoid that the signal timings get to far from the coordination plan.

When the PRIBUSS function is ended, start-orders for signal changes blocked by the priority will be given by the PRIBUSS procedure. If specified conditions for compensation to conflicting signal group are meet, compensations s made by moving the start-orders in the next cycle. The PRIBUSS function can be programmed to block itself and/or another PRIBUSS function for a specified time after the PRIBUSS function is ended to reduce the impact to other traffic and avoid the signal getting completely out of coordination.
5.2. SCOOT

SCOOT is one of the first commercial, wide spread self-optimising signal systems for coordinated networks (Arveland 2006). SCOOT has a centralised structure with a central UTC computer that optimises the whole area communicating directly with every signal controller (Hunt et al 1981). The SCOOT algorithm optimises the cycle time every five minute, optimises offsets every cycle and optimises green split a few seconds in advance of every stage change. SCOOT only makes small changes from the previous cycle in order to make the system stable. Newer versions including BUS SCOOT however allows larger temporary changes for PT priority (TRL 2010).

There are two types of bus priority in SCOOT: central and local. SCOOT can make a central extension or recall; the bus approaching the intersection is taken into consideration in the optimization process given a higher weight than cars (Bowen et al. 1994). With central bus priority the communication with the central takes a few seconds that is a problem if the detection point is close to the intersection. For buses arriving late at green local priority in the signal controller can be made outside the optimisation. The local priority can be inhibited by SCOOT if the degree of saturation is too high. Also the central priority can be restricted to degrees of saturation, or recalls can be restricted to be allowed only under a specified degree of saturation since they disturb other traffic more than extensions. It is also a possible to restrict priority to late buses only. The Stage sequence cannot be changed, but in later versions of SCOOT (MC3) there is a function for stage skipping. SCOOT does not predict dwell time at stops, so the bus detector needs to be placed downstream of the nearest bus stop (TRL 2010). The average travel time reduction when introducing SCOOT is typically around 11% (Arveland 2006).

5.3. SCATS

SCATS is an Australian partly self-optimising, partly rule-based heuristic, hierarchical signal system (Kronborg & Davidsson 2004). It could be described as a mix of a self-optimising system and a plan selection system (Arveland 2006). The development started in the 1970’s and SCATS is mostly spread outside Europe. SCATS have a strategic level determining the optimum cycle length, phase splits, and offsets to suit the prevailing traffic conditions managed by the regional computers, and tactical level handling the cyclic variation in demand at
each intersection (SCATS 2014). The SCATS controlled network is divided in sub-
systems of one to ten signalised intersections around one critical intersection.
Cycle time and green splits are optimised for the critical intersection based on
real time vehicle counts. Timings in the other intersections in the sub-system
are then selected from a library of pre-set signal plans so that they are
compatible with the optimised signal plan for the critical junction. The local
controllers can skip stages if there is no demand, and shorten stages for side
roads if there is low demand within limits set by the strategic level. Optimal
offsets are calculated every cycle and are changed if three out of five
subsequent calculations suggest the new offset. At a strategic level a choice can
be made if two or more sub-systems should be joined with a common cycle
time at the current traffic situation (Arveland 2006).

Bus priority is made locally in each signal controller and includes green
extension, red truncation and extra stages. Green split compensations can be
made after a priority measure. Melbourne uses a town-wide SCATS system
(Currie & Shalaby 2008). The tram/bus priority mostly works with green
extensions and green truncations, but special stages can also be called;
clearance stages that clears right turning vehicles (Since Australia drive on left-
hand side) from the tram track and special tram stages. The extensions and
truncations are normally limited to 20% of the cycle time, and the cycle time
cannot be changed. The settings are different for the typically four time plans,
and generally is tram priority only given in peak direction during peak periods.
Conflicting priorities are handled by fist come – first served (Currie & Shalaby
2008). SCATS seem to be good for making green waves at the main road, and
have showed main road travel time reductions of 18%. The effects on side roads
and on PT are unclear (Wood 1993; Arveland 2006). In Melbourne tram travel
times where reduced six to ten per cent, and car travel times one to seven
percent when SCATS were introduced (Currie & Shalaby 2008).
5.4. MOTION

MOTION is a German partly self-optimising hierarchical traffic signal system (Kronborg & Davidsson 2004). It can be described as a selection system or automatic updating system, for fixed-time plans combined with some local traffic adoption.

![Figure 16: Optimisation workflow in MOTION (Siemens AG 2003)](image)

The optimisation is made in a central computer with some optional, non-essential intelligence for traffic adoption distributed to the local signal controllers. In the “MOTION Central” the traffic flows are predicted based on counting detectors and a traffic estimation model. Based on the predicted traffic flows can either predefined plans be chosen or new ones adaptively calculated. In adaptive MOTION optimal cycle time, green split, offset and basic Stage sequence (from predetermined, allowed stage sequence), calculated every five to 15 minutes for the whole coordinated system. From the calculated optimal signal times time plans for each intersection is created. If the new time plan is estimated to be significantly better than the currently used time plan, it is sent to the local signal controllers. If the signal controller is equipped with local logic, the basic Stage sequence can be modified in the current cycle by a vehicle actuated local control method by inserting special stages etc. (Siemens 2003).
Locally also some traffic-actuated changes can be made to react immediately to individual vehicles e.g. to give priority to PT vehicles. At central level only passive PT priority can be given, and active PT priority made locally in the signal controllers are not recognised by the central logic (Siemens 2003). The emphasis in MOTION is put on making good green waves for the main road according to German tradition.

5.5. Utopia/Spot

Utopia/Spot is a hierarchical distributed self-optimising signal control system developed in Italy and spread to a number of European countries (Mizar 2012). It is the only self-optimising system implemented in Sweden so far (Kronborg & Davidsson 2004). Utopia/Spot optimises the signal timings by minimising a cost function locally in a Spot unit at each intersection. Spot units in neighbouring intersections communicate counted and/or predicted traffic flows and predicted signal changes for the coming minutes. The neighbouring intersections are thereby adapting to each other’s predictions and thus creating coordinated green waves. In contradiction to the systems mentioned above a common cycle time is not necessary, the coordination is decided in each intersection and each cycle.

Utopia/Spot is developed to prioritise buses/trams. Their movements though the system, including stops, is predicted in Utopia and the intersections are “prepared” in advance for the arrival of the bus. At the local level in Spot, the bus can be weighted-up and given a value equal to a large number of cars in the optimisation process.
In Torino Utopia/Spot has resulted in reduced travel times of two to seven per cent for PT and 10% for cars. The average travel time reduction when introducing Utopia/Spot has been 10% for PT and 12% for cars (Arveland 2006). Field tests of Utopia/Spot in Stockholm have showed 10% travel time reduction for buses and 7% for cars (Al-Mudhaffar 2006).

5.6. ImFlow

ImFlow is a self-optimising signal system with distributed intelligence and a structure similar to Utopia/Spot.

A multi-criteria cost function is minimised. The optimisation is made in two steps; stage based optimisation at network/route (Network optimiser) level based on a cost function with user defined weights, and signal group based optimisation at intersection level (Intersection optimiser) based on logical rules (Peek Traffic 2013), see Figure 19 below. Both optimisations levels are distributed at the controllers. The actual signal controller (Traffic light controller) can be allowed to override the optimisation, and the controllers decisions are sent back to the intersection optimiser. Different number and different placement of detectors can be used. As well as in Utopia/Spot no common cycle time is used in the ImFlow network.
ImFlow is a relatively new system, launched 2012, and there is only a few installations implemented so far. According to the manufacturer ImFlow reduced the delay 24% compared to Utopia/Spot in the test installation in Helmond 2011. A few systems are planned to be implemented in European countries, including three systems in Denmark, in the near future.
6. Evaluation of bus priority impacts in traffic signal systems

6.1. Data needed for evaluation
Depending on the type of bus priority criteria to be studied, different types of data need to be collected to evaluate the effects (Wahlstedt 2014). Active bus priority in traffic signals not only gives benefits to the buses, it will also affect other road users. In most cases the impacts on pedestrians, cyclists, cars and lorries, need to be considered as well as the impacts on buses when the effect of the bus priority is evaluated.

If bus regularity is a criterion for giving signal priority; the time table accordance, headway distribution or other service indicator should be observed. The overall system effect on the bus line is of more importance than the local effects in each intersection. Reduction of total trip time for public transport users could also be used as overall criteria for bus priority, in which case both travel time and regularity needs to be observed in a larger area.

If minimising total travel time, or delay, is a criterion for the bus priority, the most important factor to observe is of course travel time for all modes of road users. Difference in travel time can in this respect be considered as almost equivalent to difference in delay caused by the signal control, but travel time is easier to observe. Number of stops for vehicles can also be of importance as an indicator on fuel consumption, emissions, noise, vehicle costs and other important factors that are difficult to measure directly.

6.2. Empirical tests
Field tests of bus priority tend to be expensive and time consuming (Al-Mudhaffar 2006). Data collection needs to be conducted under a long time period in order to account for short-time fluctuations in arrivals. Since the number of buses is few in most cases, the arrival time variation will be relatively large. For long data collection periods, e.g. including, before-and-after studies seasonal variations also have to be taken into account.
Traffic impact surveys for pedestrians and cyclists mostly require manual observations, which are expensive and therefore have to be limited to shorter periods. Bus travel times, however, can be collected through AVL systems used by most PT agencies nowadays.

6.3. Analytical models
The effects of active bus priority are hard to model analytically due to the stochastic nature of the bus arrivals (Morellato 2010). The arrival of buses at an intersection can best be described as events rather than traffic flow. In a traffic signal with bus priority each bus can have a large impact on the performance depending on its arrival time within the signal cycle. The effect of a bus arrival in each cycle second therefore need to be calculated, and combined with the probability that arrival.

Analytical models for calculating the effects of isolated FT controlled traffic signals without bus priority exist as described in Chapter 2, and are included in commercially available software such as Capcal, Dankap and Sidra (Linse 2013, Vejdirektoratet 2005, Akcelik & Associates 2012). Analytical models for the impacts of FT controlled traffic signals with bus priority can be found in the literature (Liu, Zhang & Cheng 2008, Abdy 2010). Abdy have developed an analytical model for impacts (delays for buses and cars) of bus priority in FT controlled coordinated systems (Abdy 2010). The model does not claim to give exact result, but an estimation to use for choosing the order of intersections to implement bus priority in. The intersections are considered one by one, and a progression factor is used to describe the proportion of vehicles arriving at green when no priority is active. Only parts of the effects occurring in other intersections than in which the priority measure is taken in will be considered by using the progression factor. More exact calculation of the impacts of bus priority in coordinated traffic signals is probably not feasible with an analytical model.
Also analytical models for effects of VA controlled signals without bus priority can be found in the literature, but require extensive calculations. HCM2010 chapter 31 includes such model, including probability calculation of extension interval lengths in a ring-barrier controller (HCM2010). An analytical model for the impacts of bus priority in an isolated VA controlled traffic signal would theoretically be possible, but would require extensive calculations of probability combinations for car and bus arrivals. No such model has been found in the literature.

The impacts of self-optimising control strategies are generally not possible to calculate analytically but can be estimated using simulation as described below.

6.4. Simulation as tool for evaluating signal strategies
Traffic simulation can be a suitable tool to study active bus priority in traffic signals, as well as to study other forms of adaptive signal control including self-optimising control strategies (Bang 1976). Purpose built simulation models can be used to study signal control strategies, but more commonly are commercially available microscopic traffic simulation software such as VISSIM, Aimsun, CoreSim or MitSim used. For general or large scale studies macroscopic or mesoscopic models can be used. For detailed studies of signal control are microscopic models needed due to the influence the stochasticity in the traffic behaviour has on signal control (Wahlstedt 2014). If system effects of bus priority on regularity should be studied a mesoscopic simulation model such as BusMezzo could be appropriate (Cats 2011). The network effects of changed signal control can be studied with hybrid microscopic-mesoscopic models (Burghout & Wahlstedt 2007).

Many of the commercially available microscopic traffic simulation software provide possibilities to program some signal control functionality integrated in the software package. If the signal control should be studied in detail, or more complicated signal control such as self-optimising control analysed, an external signal control simulator can be connected to the traffic simulation software (Zlatkovic, Martin & Stevanovic 2009; Čapek, Pitkänen & Niittymäki 2011, Wahlstedt 2014). A signal control simulator that is run on a normal PC (the same or another computer than the traffic simulation) it is usually referred to as software-in-loop (SIL), and a physical signal controller, or other special controller equipment connected to the traffic simulator, is referred to as
hardware-in-loop. If the software-in-loop signal control is a PC version of the real signal controller program it is often referred to as a signal controller simulator; and if it is a model of the signal controller, sometimes developed by a third part, often referred to as an signal controller emulator.

Figure 20: The hybrid mesoscopic-microscopic-SIL simulation model used in Paper I. The mesoscopic MEZZO (left), microscopic VISSIM (bottom right) and signal controller simulator EC1-simulator (top right)

One advantage with simulation studies is that a controlled experiment can be conducted without disturbance of external factors such as small traffic variations or incidents. Simulation also makes it feasible to evaluate a variety of control strategies (Morellato 2010) and future scenarios.

The disadvantage is that a model by its nature is a simplification of the reality, and not all aspects of the real world are covered. Sociological aspects and the road user’s perception and understanding of the signals are normally not covered, for instance the influences of a fixed Stage sequence known to the road users compared to a free Stage sequence. Technical reliability is seldom included in the models, for instance intermittent detector faults.
7. Evaluation study

7.1. Introduction and methodology
The impacts of different strategies for bus priority have been tested in a number of simulation experiments, conducted in software-in-loop simulation environments. Both active bus priority with conventional control and self-optimising signal control strategies was tested. Two areas located in central Stockholm were simulated. In the first two case studies on active bus priority in conventional controlled traffic signals, a simulation model of the urban arterial Fleminggatan was used. The third case study on self-optimising control strategies used a model of an X-shaped street network around S:t Eriksplan.

Travel times or delay for different groups of road users were used as key measurement in the evaluation of the strategies in the three case studies. In the study on self-optimising control the number of stops also was evaluated.

7.2. Simulation setup
Software-in-the-loop (SIL) setups were used in the simulation experiments. The SIL concept means that a software version of a traffic signal controller (or other traffic control device) is combined with a traffic simulation software. The advantage with SIL is that advanced control strategies such as active bus priority or self-optimising control can be modelled in a correct way.

The vehicular movements and interactions are modelled in a traffic simulator software as well as the detector loops. The detector status is sent to the signal controller simulator software running the signal control strategy. The signal controller simulator sends signal group statuses back to the traffic simulator. The signal group status changes the indication of the traffic signals in the traffic simulation model and thereby controls the simulated traffic flow. Output data such as travel times, delay, number of stops and other selected indicators is calculated in the traffic simulation software. The principal structure of the simulation set-up is shown in Figure 21 below.
In the simulation experiments the traffic simulation software VISSIM (versions 5.30 and 5.40 respectively) was used as traffic simulator (PTV 2011). VISSIM is a microscopic time-step and behaviour-based multi-purpose traffic simulator to analyse highway and urban traffic (Fellendorf & Vortisch 2010). Private and public transport can be modelled in detail. Different traffic signal logic can be modelled with the built-in fixed timed and vehicle actuated signal control or external signal control logic can be connected (PTV 2011). In all four case studies VISSIM was connected to external EC1-simulators, and in the fourth case study also to Utopia/Spot and ImFlow simulators (see section 5.5 and 5.6 above).

EC1-simulator is a PC version of the Peek EC1 signal controller, one of the more frequent signal controllers in Sweden today, possible to connect to VISSIM. Several EC1-simulators can be controlled and synchronised via the EC1SimulatorControl software (Starck 2007). In those studies VISSIM and the EC1-simulators was run on separate computers to make the setup more stable. There is no time synchronisation between the two simulators, but the simulation speed can be fixed at the same value. Care must be taken to ensure that the chosen simulation speed is not faster than VISSIM is able to run at the used computer at any time of the simulation. The simulation speed used was between two and five time real time in the different case studies.
The Utopia/Spot simulator is a number of programs forming the Utopia/Spot environment; it is run on a separate laptop and connected to VISSIM (Pasquero 2011). Utopia is the central level of the system, and a Spot instance controls respective signal controller in each intersection. The structure of the Utopia/Spot – VISSIM SIL setup is shown in Figure 23 below. The version of Utopia/Spot used in the simulation experiment was not able to run faster than real time, and there was no time synchronisation between the Utopia/Spot simulator and VISSIM. A newer version of the Utopia/Spot simulator is able to synchronise with VISSIM as well as running faster than real time.

The ImFlow simulator is PC version of the ImFlow system, designed to be connected to VISSIM (Peek Traffic BV 2013). The both software are run on the same computer, and the ImFlow simulator controls and run VISSIM via the com interface. The two simulators are automatically synchronised and was able to run up to 40 times real time in the used simulation setup.
7.3. Case study one
In the first case study the impacts of bus priority with the PRIBUSS method on different groups of road users was studied. The impacts on car and bus traffic in the direction opposite to the prioritised bus were also studied by limiting the bus priority to one direction at a time. Paper II (Wahlstedt 2011) was based on case study one.

7.3.1. Study area
Case study one, two and three were conducted in a simulation model of a part of the urban arterial Fleminggatan and the crossing street Scheeleagan on Kungsholmen in downtown Stockholm.

Figure 23: The study area in case study one and two with the signalised intersections and bus line 1 marked

The streets in the area are relatively congested during peak hours, and there is heavy bus traffic in general purpose lanes. Fleminggatan is a typical Stockholm main street, 18m wide, accommodating approximately 21 000 vehicles daily. It has one lane for general purpose in each direction, on-street parking and short turning pockets at some intersections. The intersection Fleminggatan - Scheeleagan is the critical intersection in the signal coordination, the other side streets are only of local importance.
Trunk bus line 1 runs at Fleminggatan in five minute headway with articulated buses and have signal priority. Bus line 56 with 15 minute headways follow the same route through the study area but have no signal priority. Line 40 with 10 minute headways operates on crossing Scheelegatan without signal priority. There are no dedicated bus lanes in the studied area.

The simulated area includes a system of six coordinated traffic signals, five on Fleminggatan, and one on Scheelegatan coordinated with the other in peak hours. The signals have a cycle time of 82 seconds in the morning peak hours (07.00 – 10.00) and are mainly fixed-timed, with some local traffic adoption. All but the main intersection have two-stage control in mixed pattern, the main intersection Fleminggatan – Scheelegatan also has a third stage with past-end green and protected left turn from west. The signals on Fleminggatan have PTSP with the PRIBUS method implemented.
The area has been studied several times the last decades and there is much traffic measurement data available. The impacts of local traffic adoption within the mainly fixed-timed coordinated signal control, and the benefits of TRANSYT optimized time-plans was studied with the Finish traffic simulation software HUTSIM (Al-Mudhaffar & Cunningham 2001). Field tests of self-optimising signal control with Utopia/Spot was conducted 2003 in a larger area on Kungsholmen, including Fleminggatan and Scheelegatan. At the same time the impacts of PTSP with the PRIUBUS method was evaluated in a short field test on Fleminggatan (Al-Mudhaffar 2006).

The impacts of PTSP with the PRIUBUS method was evaluated with SIL simulations in a master thesis using field data from the Utopia/Spot trials (Wahlstedt 2005). Later the same simulation model was used to evaluate Utopia/Spot in the MATSIS project (Kronborg 2008) before self-optimising control was abandoned in the project. The VISSIM model was then used as base scenario in a study of a tramway on Fleminggatan (Petoukhov & Wahlstedt 2010).

7.3.2. VISSIM model
The VISSIM model used in case study one was further developed from the one used in the master thesis, MATSIS project and tramway study mentioned above. The VISSIM network was cut down to 1000m x 500m and contains six signalised intersections, ten incoming and twelve outgoing links for vehicular traffic, pedestrians crossings, and three bus lines whereof one has signal priority.

As far as possible, data from 08.00-09.00 in the morning of Wednesday 2003-05-14 were used in the model. Traffic flows and turning percentages were collected from the Spot units in the self-optimising signal control system Utopia/Spot tested at the location (Spot collected data also when conventional control was in use) in 2003. The traffic counts from Spot were compared to manual counts from video recordings of the intersection Fleminggatan – Scheelegatan and Spot showed an overestimation of approximately 5% (Wahlstedt 2005). The flows and turning percentages reported by Spot in five minute intervals were aggregated to hour averages and complemented with manual counts. The percentages of heavy vehicles except buses were 2-3% on the streets according to manual counts from video recordings made by KTH.
Data on boarding and alighting passengers at each bus stop and occupancy of the buses on line 1 for the period 030505 – 030523 was collected from ATR, the AVL system used by the regional transport agency SL, and coded into the VISSIM model. The dwell times at bus stop were also collected from ATR, and a linear dwell time function depending on boarding and alighting passengers was estimated and coded into the vehicle model in VISSIM. The dwell time for the buses used on line 1; Scania Omni Link articulated low floor buses with boarding through a double front door and alighting through two double and one single rear door, was found to be (Wahlstedt 2005):

\[ t_d = 11.3 + 2b + 0.5a \]

Where:

- \( t_d \) = dwell time [s]
- \( b \) = number of boarding passengers
- \( a \) = number of alighting passengers

In the VISSIM model used in case study one the bus arrivals where coded according to the posted time table, and some disturbance introduced by a dummy bus stop, before the buses enter the actual network, with a normally distributed dwell time of 60 ± 20s.
This simulation model was calibrated and validated in 2004 with VISSIM 4.00 (Wahlstedt 2005), but have not been re-calibrated and validated after some later modifications of the model. The network can therefore not be considered as fully calibrated to current conditions, but as a realistic hypothetical network.

7.3.3. Traffic signal control
In case study one, the traffic signals were controlled by EC1 signal controller simulators in the SIL setup as described above. The programming of the signal controllers was the same as used on street in 2003, at the time of data collection for the simulation model, including PTSP with PRIBUSS. The signal programming used in the simulation study includes some errors present in the controllers in field in 2003, which were later corrected. This reflects the risk of discrepancy between the intended, projected function and the actual function of the signal control strategy when implemented in field. A modern Swedish microcomputer signal controller has approximately 10,000 parameters to be programmed, and there is a risk that programming errors occur due to the human factor. Errors in less frequently used functions, such as PTSP functions, are more likely to remain undetected than others. Two of the signal controllers (no. 3334 and 3338) were then of the older type ELC-3, and the programming was converted to EC1 programming by the City of Stockholm.

Four different PTSP scenarios were tested; no PTSP, PTSP in westbound direction only, PTSP in eastbound direction only and PTSP in both directions. The scenarios were formed by disconnecting the bus detectors in the SIL setup for buses in either direction.

7.3.4. Results
The travel times for buses and other traffic (cars and heavy vehicles together) along Fleminggatan and the major cross street Scheelgatan were evaluated in case study one. Ten simulation runs with different random seeds were conducted for each scenario, and results were averaged from the output of the ten runs.

With PTSP, the travel time for buses on line 1 was reduced 8.9% in westbound direction and 1.6% in eastbound direction throughout the simulated network. Travel times for other traffic on Fleminggatan, traveling in the same lane as the bus was reduced 2.1% westbound and increased 6.4% eastbound. This increase
in travel time for other traffic following the bus in eastbound direction is interesting, and gave rise to further analysis. On the cross street, Scheelegatan, travel times increased 11.2% northbound, and 12.9% southbound with PTSP. The travel times are given in Figure 27 below. The bus travel time savings are somewhat smaller than reported by limited field studies of the area, 11% (Al Mudhaffar 2006).

![Travel times with and without PTSP](image)

**Figure 27: Travel times with and without PTSP with the PRIBUSS method**

The number of stops in the whole simulated network decreased 23% for buses and increased 4% for all traffic with PTSP.

In the second step, PSTP was also tested for westbound buses only or eastbound buses only. In Figure 28 below the travel times between each intersection for the four scenarios are shown. Generally other traffic benefits from PTSP for buses in the same direction as it travels, but is disadvantaged by PTSP for buses in opposite direction of travel. PTSP for buses in both directions can have either positive or negative impact on other traffic. At intersections downstream a bus stop the travel time for other traffic are in many cases increased with PTSP; the time gained in upstream intersections can be lost when the normal coordinated green wave has to be awaited when the bus stops. Buses generally gain most from PTSP in its own direction of travel only, and gain less from PTSP in both directions of travel. The impact is very different in the different intersections.
Later a programming error of the PTSP in the signal controller at the intersection with Polhemsgatan was found by the City of Stockholm engineers and corrected in field. The error, obstructing the signal on Fleminggatan in eastbound direction to turn green for one cycle, only occurred after a long green extension in westbound direction.

Figure 28: Travel times for buses and other traffic (Car) between each intersection along Fleminggatan with PTSP in no, either or both directions. The results affected by the PTSP programming error marked with arrows.

7.4. Case study two
In case study two, the impact on buses and other traffic of the different bus priority functions (green extension, red truncation etc.) within the PRIBUSS method was studied. The impacts of each PRIBUSS function was tested one by one and all functions together. The impacts of the different PRIBUSS functions were studied first in the main intersection (Fleminggatan – Scheelegatan) and then in all signal controllers in the coordinated system on Fleminggatan.

Each scenario was tested with counted traffic flows, and with ten percent reduced traffic demand. Since the PTSP reallocates green time to the main street (Fleminggatan) the main intersection become temporary oversaturated without PTSP at counted traffic flows. With ten per cent reduced flows the intersection remains within capacity also without PTSP. Totally 22 different scenarios where tested in case study two. Paper three (Wahlstedt 2013a) was partly based on case study two, but some more results and details are presented here.
7.4.1. VISSIM model
In case study two, basically the same VISSIM model as in case study one was used but with some modifications. The bus detectors in the VISSIM model, and their connection to the inputs to the EC1-simulator, where changed in the way different PTSP functions could be called with separate detectors for each PRIBUSS function.

The bus arrival times in the VISSIM model were changed to actually observed departure times from the bus stops upstream of the simulated area, S:t Eriksgatan eastbound and Cityterminalen westbound marked in Figure 29 below. The actual departure times for line 1 at the two bus stops where derived from data in SLs AVL system (BussPC) between 08.00 and 09.30 on 2011-04-07, one random but typical day (i.e. a day where no major accident or other disturbance occurred). In the VISSIM model the 90 minute peak hour input data was repeated twice to get three hours output data per simulation run. It should also be noted that the bus data is from another year than the traffic flows, turning percentages, travel times etc. used in the model.

![Figure 29: The route of line 1 with the simulated area, bus stops where data were collected as well as the approximate city center marked (Cats 2012).](image)

The number of passengers on each bus entering the VISSIM network was calculated in proportion to the headway to the previous bus. The number of passengers varies from 5 to 62 when put into the VISSIM network. Since the dwell times at bus stops in VISSIM are calculated depending on the number of boarding and alighting passengers, as described in chapter 7.3.2 above, the
buses occupancy when entering the network will affect their travel times. Also
the headway will affect the number of boarding passengers in the VISSIM
model. Late buses, or actually buses with long headways, will therefore
systematically have longer dwell times in the VISSIM model, and the delay
increase. This is an important property of the model if the effect on regularity is
studied, as in case study two.

Figure 30: Headway distributions for buses entering the simulation network (nominal headway
cfive minutes), observations from field 2011-04-07.

The AVL data do show some bunching tendencies of the buses in the studied
area, with many short (and many long) headways, especially in westbound
direction, as can be seen in Figure 30 above. The travel time in westbound
direction from the first stop Frihamnen through the city centre to
Cityterminalen is 25 minutes, and the buses have larger risk to bunch compared
to in eastbound direction where the travel time from the first stop Stora
Essingen to S:t Eriksgatan is only 17 minutes and the traffic situation more
predictable.

7.4.2. Traffic signal control
The same SIL setup as in case study one described above was also used in case
study two. The programming of the signal controllers was basically identical
with the one from field used in case study one, but slightly modified. The signal
controllers where re-programmed in the way each group of PRIBUSS functions
could be started separately with different inputs. The programming error
mentioned in case study one was corrected, but after the case study was made
yet another error was found by the City of Stockholm’s signal engineers after
looking at the results from case study two. The function for Green extension
with maximum time (BF) was in practice not able to extend green in westbound direction. The signal controller programming in field has been updated, but not the simulation setup.

Due to local conditions all PRIBUSS functions (see chapter 5.1) are not implemented in every intersection, and the same functions are not always used both directions of travel the same intersection. The functions implemented in each intersection are given in Table 1, and the location of the signal controllers are shown in Figure 31 below.

Table 1: PRIBUSS functions implemented in each signal controller

<table>
<thead>
<tr>
<th></th>
<th>3331</th>
<th>3307</th>
<th>3338</th>
<th>3334</th>
<th>3321</th>
</tr>
</thead>
<tbody>
<tr>
<td>BF (Green extension)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>BF2 (Green extension during past-end-green)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ATS (Re-taken start)</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF (Extra phase)</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>AK (Red truncation)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>DAK (Double red truncation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 31: Traffic signals around the studied area, the study area marked and signal controllers numbered
Green extension during past-end-green (BF2) is only meaningful in approaches using past-end green (PEG), and not used where long green extensions should be avoided. Since the intersections are two-phase controlled, extra phase (EF) is only meaningful for westbound buses in 3334 between lagging green in the opposite direction and the cross street phase, see Figure 25 above. Red truncation (AK) is not possible in signal controller 3307 since the red time is only determined by the minimum green time and clearance time for the pedestrian crossing across Fleminggatan, which cannot be shortened of safety reasons. Re-taken start (ÅTS) and double red truncation (DAK) is only used at nearside bus stops, where the request for priority can only be made a few seconds ahead of the bus arrival at the intersection.

Due to the limited number of I/O inputs in the signal controller, some of the PRIBUSS functions needed to be grouped together and requested with a common input. The PRIBUSS functions with similar action were grouped together to form a number of scenarios to be tested.

Green extension with “max time” (BF) cannot be started after start order is given to a conflicting group, and the extension will therefore in most cases not change the green time too much. The BF function is supposed to be a short extension, and forms the scenario “short extension”.

Green extension during past-end-green (BF2) is however always started at the very end of green, and will in most cases result in larger changes of green times also in the next cycle. The Re-taken start (ÅTS) function, returning the signal to green immediately if no other signal group have turned green yet, could in this sense be seen as a green extension started extremely late, and will change the green times in the next cycle. The Extra phase (EF) function is in this case used to re-start green in westbound direction when the opposite direction is lagging green, and the effect of the function is an extension of green in the east-west direction. The BF2, ÅTS and EF functions together with BF function form the scenario “long extension”.

The Red truncation (AK) function alone forms the scenario “Red truncation”. All implemented PRIBUSS functions together form the scenario “All functions”.

60
Those scenarios as well as “No priority” were tested in the main intersection, Fleminggatan – Scheelegatan (3334). All implemented PRIBUSS functions were kept active in the other signal controllers in the study area to enable the buses to arrive outside the normal green wave. Totally four scenarios were evaluated:

1) No PTSP
2) Short extension; BF
3) Long extension; BF, BF2, ÅTS and EF
4) Red truncation; AK
5) All functions

Then the different functions were tested one by one in all signal controllers on Fleminggatan. In this case it was also possible to test the Double red truncation (DAK) function implemented in the signal controllers 3321 and 3338. Eight scenarios were evaluated:

1) No PTSP
2) Short extension; BF
3) Long extension; BF, BF2, ÅTS
4) Long extension and extra phase; BF, BF2, ÅTS and EF
5) Red truncation; AK
6) Double red truncation; DAK
7) All functions but DAK
8) All functions

7.4.3. Results
Travel times for buses and other traffic was evaluated for each scenario. The travel times were collected between each stop line for both through traffic and for vehicles from the side streets turning onto the main streets (Fleminggatan and Scheelegatan). Also the travel times for through traffic from one end of modelled part of Fleminggatan and Scheelegatan to the other was collected. The travel times for buses and other traffic around the main intersection are given in Figure 32 and the travel time difference compared to the no PTSP scenario in Figure 33 below.
Figure 32: Travel times at intersection Fleminggatan - Scheelegatan with different PRIBUSS functions

Figure 33: Travel time difference at intersection Fleminggatan - Scheelegatan with different PRIBUSS functions

No Priority
BF
BF+BF2+ÅTS+EF
AK
All Functions

Travel time [s] per vehicle, 3334

Travel time difference [s] per vehicle, 3334
In the intersection Fleminggatan – Scheelegatan the results show large travel time savings for westbound buses and other westbound traffic with all PRIBUSS functions. In eastbound direction, the travel time savings is much smaller. This is due to long queues in westbound direction, which are "pushed" though the intersection in front of the prioritised buses. This can be observed in field in the specific intersection. In order to get more generally applicable results a sensitivity analysis with 10% less traffic has been performed. The 10% reduction in traffic demand is of the same magnitude as the impact at the location of the congestion charges introduced in 2007.

![Travel time difference in person seconds, 3334](image)

Figure 34: Total person travel time difference at intersection Fleminggatan - Scheelegatan with different PRIBUSS functions

The travel times per person have also been calculated by multiplication of the average number of passengers per vehicle; 1.2 persons per car, 26 passengers per eastbound bus, and 38 passengers per westbound bus. The travel times per person are given in Figure 34 above. The travel time saving per person was 8% with Short extensions only, 12% with Long extensions, 7% with Red truncation and 15% with All functions together in the intersection Fleminggatan – Scheelegatan.
With 10% traffic flow reduction there are still queues in westbound direction but the oversaturation in the no PTSP scenario is relieved. The results still show the same pattern, see Figure 35. The travel time savings for prioritised buses are relatively large with a small increase in travel time for other traffic with both short (BF) and long extensions (BF+BF2+ÅTS+EF). The travel time savings for eastbound buses are limited by the short distance from the nearside bus stop to the intersection. Red truncation (AK) shows a reduction in bus travel time, but at a cost to the side street traffic. The short distance from the nearside bus stop to the stop line at the traffic signal in eastbound direction is of less importance for red truncation compared to extension. The combination of green extension and red truncation gives the largest travel time reductions to buses, and to other traffic following the buses. For the cross street traffic there is an increase in travel time between the increase with only extension and with only truncation.

The results when the different PRIBUSS functions were tested in all signal controllers on Fleminggatan show similar conclusions as in the main intersection, but the travel time savings for prioritised buses is not as large as in the main intersection. The results for the scenarios with counted traffic and with 10% reduced traffic flows are given in Figure 36 below. The patterns are the same, but the magnitude of the impacts different.
Green extensions give substantial travel time savings for the prioritised buses with small impacts on other traffic. Red truncation gives similar travel time reductions for prioritised buses; somewhat shorter travel times for other traffic following the bus route and increased travel time for cross streets. The double red truncation (DAK) function is only used for westbound buses at two nearside stops close to intersections 3321 and 3338 when the bus arrives late at green. A somewhat unexpected result is that the DAK function, when being the only allowed PRI BUSS function in the intersections, increases the travel time. But combined with other functions the DAK function show good results, especially in the less saturated scenario with 10% reduced traffic flows.
The travel time differences, with different PRIBUS functions active, from the upstream intersections to the stop line in each intersection is given in the diagrams in Figure 37 above. The increased travel times for westbound buses in the intersection with Polhemsgatan are partly due to the programming error in the PTSP mentioned above. The impacts on travel times are very different in respective intersection. All impacts are not easy to understand intuitively by looking at one intersection at a time; the impacts are depending on upstream intersections in both directions of travel. Only partly these impacts are possible to foresee by looking at the time space diagram of the coordination.
Some of the travel time saved in one intersection can be lost in a downstream intersection, or more time can be saved. This is depending on the structure of the coordination and the combination of PRIBUSS functions implemented in the intersections. Therefore a combination of different PRIBUSS functions is needed to achieve good results throughout the coordinated traffic signal system.

7.5. Case study three
In case study three different criteria for giving priority where tested, including a differential, on-time-status based strategy using the conclusions from case study two. With this strategy more buses were allowed to use the functions found to be less disruptive, and only a few buses allowed using the functions that showed to be the most disruptive for other traffic. Paper III (Wahlstedt 2013a) was partly based on case study three, and some more results from the case study are presented in this chapter.

7.5.1. VISSIM model
The same VISSIM model and SIL simulation setup as in case study two was used in case study three, with some small modification. The buses where classified in nine groups according to their on-time status compared to the time table based on the AVL data, see Table 2 below, and inserted as different vehicle types in the VISSIM model. Due to limitations in the simulation model each bus belonged to the same on-time class throughout the VISSIM network also if it catches up time or get delayed within the network. The bus detectors in VISSIM requesting the different PRIBUSS functions to start, were changed so that they were only activated by buses of the classes that were allowed to start the respective function in each scenario.

Most buses were found to be delayed compared to time table at the observed stops, but on time further down the line. The run times between each stop is probably not fully calibrated in the time table, but correct timed at the time points where the driver have to wait for departure time. In order to be able to test the regularisation effect of conditional priority, the on time classification were adjusted by offsetting the time table so approximately the same number of buses where late as early in the VISSIM model. This give a similar effect as if the timetable would be correct timed at the observed stops.
Table 2: Number of buses in each on-time class, positive value represents ahead of time table and negative value delayed.

<table>
<thead>
<tr>
<th>Deviation from time table [min]</th>
<th>Eastbound</th>
<th></th>
<th>Westbound</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>observed</td>
<td>adjusted</td>
<td>observed</td>
<td>adjusted</td>
</tr>
<tr>
<td>&gt; 2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>&gt; 1</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>&gt; 0.5</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>&gt; 0</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>&gt; -0.5</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>&gt; -1</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>&gt; -2</td>
<td>13</td>
<td>2</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>&gt; -3</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>≤ -3</td>
<td>18</td>
<td>4</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>

7.5.2. Tested criteria for conditional PTSP

In case study three five scenarios with different criteria for giving signal priority were tested:

1) No Priority
2) Existing conditional priority
3) Existing conditional priority + short extension (BF) for red (local) buses
4) Conditional priority with tighter on-time threshold + short extension (BF) for red buses
5) Conditional priority with new on-time-status based conditions + short extension (BF) for red buses

Scenario 1) is a reference scenario with no PTSP.

Scenario 2) uses the same criteria as in the PTSP is implemented in Stockholm today; only "Blue buses" on the trunk lines are given priority, and they are given priority only if not more than two minutes ahead of schedule. "Red buses" on local lines are not given signal priority.

In scenario 3) also Red buses were allowed to request for short green extensions (BF). In the evaluation of the different PRIBUSS functions in case study two, short green extensions had little negative impact on other traffic but gave large travel time savings for the prioritised buses. The hypothesis was that the total travel time could be decreased by also allowing buses with fewer passengers to use the PRIBUSS function least disruptive to other traffic.
In scenario 4) buses more than one minute ahead of schedule will not be given priority. The idea is that slowing down early buses will put them back on schedule and improve regularity, and at the same time cause less disruption to other traffic. Not giving priority to an early bus will also increase the possibility to give another bus priority. Red buses are also given short extensions.

In scenario 5) a new differential conditional priority strategy based on the buses on-time-status was proposed and tested. In this scheme the PRIBUS functions that in case study two showed to be the most disruptive for other traffic is only allowed to be used by the latest buses, functions somewhat disruptive are allowed to some buses and the least disruptive PRIBUSS functions are allowed to be used by most buses. The functions allowed for each on-time class is given in Table 3 below.

**Table 3: PRIBUS functions allowed for the different categories of buses in the proposed new differential, on-time-status based conditional priority strategy tested in scenario 5)**

<table>
<thead>
<tr>
<th>PRIBUSS</th>
<th>Deviation from timetable [minutes], Blue bus</th>
<th>Red bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>function</td>
<td>&lt; +2</td>
<td>&lt; +1</td>
</tr>
<tr>
<td>BF</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>BF2</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>ÅTS</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>EF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AK</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>DAK</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

By limiting the use of the most disruptive functions the impact on other traffic can be reduced. Instead functions only used by a few very late buses can be allowed to have larger impacts on other traffic. In case study three the restrictions for the PRIBUSS functions (guarantee time for cross streets, last time step etc. see chapter 5.1) was not changed from the programming from field though. Therefore the priority for the late buses could not be “harder” than in the existing strategy. In further studies this should be changed. As in scenario 4) buses more than one minute ahead of schedule is not given priority, and Red buses are given short extensions.
7.5.3. Results
Travel times for buses and other traffic was collected in the same way as case study two. The number of replications was extended to fifty simulation runs with different random seeds for each scenario, since the number of buses in each on-time class is small. Still the results for each intersection show too much variation to analyse each on-time class separately, and double the number of replications did not reduce the variance enough.

Figure 38: Travel times, including dwell time at stops, for buses on Fleminggatan with different on-time-status and for other traffic on Fleminggatan respective Scheelegatan with different criteria for giving signal priority

The travel times for buses and other traffic on Fleminggatan and Scheelegatan are shown in Figure 38 above; both directions of travel are combined for each category. The travel time for buses includes dwell time at stops, and since late buses in most cases have long headways they will have more passengers and thereby longer dwell times. This can be noticed in the results for Scenario 1) (Sc 1)) with no PTSP where the late buses have the longest travel times and the early buses have the shortest.
The criteria for priority used in Stockholm today, and in Sc 2), partly compensates for this by not giving buses >2 minutes ahead of schedule priority. The tightened threshold from two to one minute ahead of schedule for giving priority in Sc4) shows an effect, buses between one and two minutes of schedule previously having the shortest travel times will now have the longest travel time and smoothly hold back till they are on schedule. The new proposed on-time-status based conditions in Sc 5) show the same effect, early buses are hold back and late buses speed up. However this effect is not large enough to make up for the longer dwell times for the latest buses, they still have the second longest travel time.

Figure 39: Difference in travel times, including dwell time at stops, compared to PTSP with existing conditions (Sc 2)
The differences in travel time compared to PTSP with the conditions for giving priority used today are given in Figure 39 above. Giving Red buses moderate priority with short extensions (Sc 3)) increases the travel time somewhat not only for other traffic, but also to some extent for the Blue buses. Depending on the number of passengers per bus the total impact could be positive or negative, but no net effect calculation have been made. Changing the on-time threshold from two to one minute ahead of schedule (Sc 4)) of course increases the travel time of the buses between one and two minutes ahead of schedule, but it also decreases the travel time for other traffic since the number of green time changes made by the PTSP decreases. It also decreases the travel time for other buses since the chance to obtain a requested signal priority is larger if the requests are fewer. The proposed new differential on-time-status based conditional priority strategy (Sc 5)) acts mostly as supposed. Early buses are held back and late buses are speed up, but the effect is not much larger than in Sc 4). Other traffic experience reduced travel times compared to existing criteria in Sc 2) though.

7.6. Case study four

In case study four the impacts of self-optimising control, including PTSP, in coordinated traffic signals was studied. The work was part of a project called “Adapt” made for, and financed by the City of Stockholm. The main goal was to reduce bus delay without acceptable impacts on other road users, and to improve accessibility for pedestrians and cyclists, in order to minimise the overall delay per person. The objectives of the project were to make a beta test of the adaptive traffic signal control systems available on the Swedish market in a simulation environment and to compare them against each other and with the existing signal control scheme to see if self-optimising signals can be a tool to achieve the main goal; reduced over all delay per person.

Two commercially available self-optimising signal systems were compared with the conventional coordinated traffic signals including PTSP with the PRIBUSS method. Also the impact of PTSP with the PRIBUSS method was evaluated in the study area. Paper IV (Wahlstedt 2013b) was based on case study four, but some more details on the simulation model, input data and more results are presented here.
7.6.1. Method
The suppliers of signal control equipment and self-optimising signal control systems in the Swedish market (e.g. Peak Traffic and Swarco) were invited by the City of Stockholm to participate in a “competition” where their systems would be tested in a simulation environment against each other, and against the existing conventional coordinated control including PTSP. The suppliers were invited to participate in the project at a fixed remuneration for which they should provide a simulator version of their system possible to connect to VISSIM, as well as programming of the system with the settings the suppliers found optimal. One slightly modified version of the VISSIM model later used to evaluate the systems was provided to the suppliers, as well as the signal timing plans for the existing signal control, to set up and optimise respective self-optimising signal control system.

The final VISSIM models with different traffic flows used to evaluate the signal control systems, and the exact evaluation method were not revealed to the suppliers, only the goal to minimise the overall delay per person and that traffic flows from different time periods would be used. Based on these instructions the suppliers were free to set up and program their systems in the way they believed to be optimal. The evaluation was then made with SIL simulations where VISSIM was connected to signal control simulators for each signal control strategy tested. The adaptive signal control simulators were treated as a “black box” set up by the suppliers.

7.6.2. Study area
Case study four was conducted in a simulation model of an area around S:t Eriksplan in central Stockholm and consists of five intersections in an X-shape. The intersections are approximately 200 meters apart. The north east – south west S:t Eriksgatan is a major urban arterial with high traffic demand. The north west – south east Torsgatan is also an urban arterial with relatively high traffic demand. The east – west Karlbergsvägen and Odengatan are main streets in the neighbourhood.
The study area, from Google Maps, signal controllers with numbers and trunk bus routes with stops. Position of the count tube in marked in green bottom right, and location of photos below with yellow arrows.

The two trunk bus lines 3 and 4 traversing the study area have high passenger volumes and operate with five minute headway. Six local bus lines are also passing through the area, and operate with ten minute headway or less. There are large numbers of boarding and alighting passengers at the bus stops in the study area due to the interchange with metro and commuter rail, making the dwell times long. The trunk bus lines are given signal priority and traverse the study area in another route than the main traffic streams.
The area around the study area has many restaurants, shops and work places attracting large pedestrian flows. The metro (Tunnelbana) station S:t Eriksplan is situated under the study area, and the commuter rail (Pendeltåg) station Karlberg is situated 400 meter west of the study area. There are large pedestrian flows to and from those stations passing through the study area, as well as to the major bus stop at the S:t Eriksplan square.

7.6.3. Traffic signals
Due to geometrical reasons the central intersection (5515) needs an all-pedestrian phase, which limits the overall capacity of the coordinated system. Left turns are prohibited in the central intersection in order to keep the cycle time down, and the turning movements moved to the northern-most intersections (5516 & 5518). Those two intersections have complicated patterns with protected left-turns while the two southern-most intersections (5512 & 5514) have simple patterns with unprotected left turns and are less saturated.

The existing signal timing plan used in field today was developed 2006 in the Matsis project (Kronborg 2008) using basically the same simulation set up as case study four. One single signal time plan with 100s cycle operates 06.30 – 20.00, covering all four studied time periods. The signal plans were made by hand by an experienced traffic signal engineer, fine-tuned and evaluated with VISSIM simulations.
PTSP with PRIBUSS was introduced already in 1999 in four intersections, but the PRIBUSS settings were not updated with the new signal plans in 2006. PRIBUSS functions for Green extension, Red truncation and Extra stage (see chapter 5.1) are used. The PTSP is still functioning, but the PRIBUSS settings are not optimal with the new signal timings. The bus priority in the central intersection (5515) is disabled due to its major impact on green waves and capacity in the north-south main direction.

The existing signal control is not optimised with the same criteria as requested in the Adapt project. The goal of the Matsis project in 2006 was to minimise the emissions by reducing the stops within the coordinated signal system, as opposed to minimising delay as in this project. Delay for buses and pedestrians were not considered in the Matsis project. Furthermore, the traffic flows and traffic patterns have also changed since the signal plans were optimised. Traffic flows in central Stockholm have decreased approximately 10% due to the introduction of congestion charges (Eliasson 2009). Large construction projects have also affected the routes used in the surrounding area.

7.6.4. VISSIM model
The VISSIM model used in case study four was originally developed in the MATSIS project (Kronborg 2008) and includes the five studied intersections, surrounding streets and the major intersection S:t Eriksgatan – Fleminggatan 600 meters south of the study area. This intersection was included in order to get some platooning of the vehicles arriving to the study area. The VISSIM network used in case study four is shown in Figure 42 below.
Traffic flows and turning percentages in the area were counted in May 2012. Manual counts of turning traffic streams were conducted 07.00 – 09.00, 12.00 – 14.00 and 16.00 – 18.00 in each intersection, and an automatic flow count with tube was made for a full week. The average weekday traffic flow counted is shown in Figure 43 below.
From those counts, OD matrixes for each counted time period were developed and used as input in VISSIM. An evening traffic scenario for 18.00 – 20.00 was constructed by scaling the manual counts from 16.00 – 18.00 by the traffic flow differences from automatic counts. Also an increased PM traffic scenario was developed by increasing all vehicle flows by five per cent to represent possible future traffic growth.

VISSIM models for in total five traffic conditions were developed:

- AM peak (07.00 - 09.00)
- Mid-day (12.00 – 14.00)
- PM peak (16.00 - 18.00)
- PM peak with 5% increased traffic flows
- Evening (18.00 – 20.00)

![Average traffic flow, Torsgatan](image)

Figure 43: Average traffic flow per weekday at Torsgatan. Studied time periods marked with red, and the time period the fixed time plan with 100s cycle is active in existing control marked with green

Pedestrians were counted at every pedestrian crossing, but only in one ten minute interval per crossing. In the VISSIM model pedestrians are only modelled at the pedestrian crossings, and at one crossing at a time. Cyclists were counted in 2005 and are modelled throughout the network to be able to study the coordination effect on bicycles.
Actual departure times from the bus stops upstream of the simulated area for the trunk bus lines 3 and 4 were collected from “Buss PC”, SLs AVL system. Both routes have a headway of ten minutes before 07.00, then five minute headway throughout the day until 19.00, then ten minute headway according to timetable. The headway distribution of the actual departure times in the two hour PM peak period for line 3 and 4 at S:t Eriksplan is shown in Figure 44 below. The headway distribution shows bunching problems for both lines, with up to three times the scheduled headway for line 4 southbound in the PM peak period. The of-peak periods show somewhat less spread headway distributions, but there are still severe regularity problems. There seem to be a large potential to improve the service reliability on both bus lines, and developed PTSP strategies focusing on regularisation can perhaps be a helpful.

The buses on the trunk bus routes 3 and 4 (blue buses) were inserted in the VISSIM model according to the observed departure times. Buses on the local bus lines (Red buses) are put into the VISSIM model according to timetable with small random disturbance.

![Figure 44: Headway distribution for line 3 and line 4 at S:t Eriksplan bus stop, time table headway is five minutes](image-url)
Actual arrival and departure times at bus stops for the trunk bus lines were collected from SLs AVL system. The AVL system holds actual, as well as planned (time table), arrival and departure times for each bus stop, but no data on boarding and alighting passengers. Approximately ten percent of the buses only are equipped with passenger count equipment, and the passenger counts have not been paired with the dwell times from AVL data in this case study.

Figure 45: Dwell time distributions at bus stops Karlbergsvägen and S:tEriksplan, 06.00 – 20.00 2012-05-10

Dwell times where calculated from the observed arrival and departure times for line 3 at bus stop "Karlbergsvägen" and for line 3 and line 4 at bus stop "S:t Eriksplan". Their distribution is shown in Figure 45 above. The dwell time shows a large variation in the range 0 – 120 seconds with an average around 40 s, and most observations between 20 and 60 seconds. This large variation makes it impossible to use bus locations upstream of the nearest bus stop to prepare the intersections for a bus arrival in the self-optimising systems. Requests for priority must therefore be placed after the bus has left the upstream bus stop. The large spread in dwell times is probably depending on the variance in number of boarding and alighting passengers due to the headway variation, but this have not been verified in this case study.
Separate dwell time distributions per stop, line and direction were put into the VISSIM model based on the observed dwell time distributions. Since each bus will be given a dwell time randomly from the dwell time distribution when stopping at the bus stop, there will be no correlation between the headway of the buses and their dwell time at stops in the VISSIM model in contradiction to the VISSIM model used in case study two and three. This is of less importance in case study four since the regularisation effects of the signal control strategies are not studied. The dwell times for the local bus lines (Red buses) have not been collected, and a standard normal-distributed dwell time of 20 ±2 seconds where used for local buses in the VISSIM model.

The saturation flows and queue lengths in the VISSIM model was validated against observed saturation flows and queue lengths in two critical approaches to the intersection S:t Eriksgatan – Karlbergsvägen. The compliance was good and no re-calibration of the model was needed.

7.6.5. SIL simulation setup
The existing signal control was simulated with the same SIL setup with EC1 signal controller simulators as the earlier case studies. The programming of the signal controllers where translated to EC1 programming from the programming of the ELC-3 controllers used in field by the City of Stockholm in 2006.

The No PTSP scenario used the same programming and setup as the existing signal control, but the bus detectors to call for priority were disconnected.

The Utopia/Spot system consisting of a number of programs were installed on a separate laptop computer provided by the signal control supplier (Swarco). All programming and the setup of the Utopia/Spot system was made by the supplier, and the system was used as a “black box” connected to the VISSIM model used for the evaluation. The Utopia/Spot and VISSIM computers were connected via tcp-ip communication.

The ImFlow system simulator was provided by the other signal control supplier (Peek Traffic) and installed on the same computer as VISSIM. The ImFlow simulator is an integrated program that calls and runs VISSIM via the COM interface. Also in this scenario the ImFlow system was programmed and set up by the supplier, and the signal control system was used as a “black box” connected to VISSIM for evaluation.
Four scenarios with different signal control strategies were simulated and evaluated:
- Existing conventional coordinated control including PTSP with PRIBUSS
- Conventional coordinated control without PTSP
- Utopia/Spot
- ImFlow

7.6.6. Results
Average delay in the whole simulated network for different groups of road users was evaluated in case study four. The number of stops was also evaluated for each vehicle category. All results given below are averaged from ten simulation runs with different random seeds.

The outputs from VISSIM in delay per vehicles were recalculated to the total delay for all travellers by multiplying it with the average number of passengers per vehicle, see Table 4 below.

Table 4: Assumed number of persons per vehicle at different traffic conditions used to calculate total delay

<table>
<thead>
<tr>
<th></th>
<th>Walk</th>
<th>Bike</th>
<th>Car</th>
<th>Trunk line bus</th>
<th>Local bus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1.2</td>
<td>AM</td>
<td>Day</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>30</td>
<td>40</td>
<td>40</td>
<td>20</td>
</tr>
</tbody>
</table>

The average number of passengers per articulated bus used on the trunk lines is considerably higher than per normal bus used on the local lines. The number of passengers per bus also varies widely on a daily basis. Therefore different numbers of passengers were assumed for local and trunk line buses, as well as different numbers of passengers per bus in the different traffic condition scenarios; peak hour, midday and evening based on passenger counts from SL and assumptions.

The total delay was substantially decreased compared to the existing signal control with both self-optimising signal control systems. The total delay in person hours for each traffic condition scenario and the relative difference compared with the existing control in is shown in Figure 46 below. The VISSIM network also includes one FT signal controlled intersection not controlled by
the evaluated signal systems, which adds some extra delay to the results. The delay in this intersection is constant in all evaluated scenarios, but the relative change will be somewhat underestimated.

Figure 46: Total delay for all road users in the VISSIM network (left) and relative change in total delay compared to existing control, person hours

ImFlow showed the best results in all but one time period. After the evaluation was finished, it was found that the tested Utopia/Spot programming used PTSP calling points upstream of bus stops. The large spread in dwell times at stops destroys the optimised timings. Later tests with changed programming where the PTSP calls were made downstream of the bus stops only show better results.

Figure 47: Relative change in delay [%] for prioritised buses (Blue buses) and un-prioritised buses (Red buses)

The results are very different for different modes of transport, as can be seen in Figure 47 above and Figure 48 below. Generally buses, pedestrians and bicyclists gain from the self-optimising controls and cars gains in some scenarios and losses in some. The prioritised blue buses on the trunk lines show large
delay reductions, thanks to the more extensive PTSP, but also un-prioritised red buses on local lines show substantial delay reductions in all but one scenario. Some of the local lines partly follow the same route as the prioritised buses, but two local lines on Karlbergsvägen cross the prioritised bus route. The different local lines have not been separated in the analysis.

Pedestrians and cyclists (not shown in the figure, but following the same patterns as pedestrians) gain from the self-optimising control, especially from ImFlow control due to shorter cycle times compared to the 100s cycle in the existing control. The cycle time reductions and consequently delay reductions for pedestrians greater in the off peak scenarios. The delay for cars increased in all scenarios but the 5% increased flow PM peak period scenario. The self-optimising control manage to react on the increased flows, while the existing control suffers from oversaturation in a few turning relations causing long queues.
Both self-optimising systems showed a substantial increase in number of stops, compared to the existing control. The existing fixed timed control created good “green waves” that the self-optimising systems could not match. The number of stops for all road users in the simulated network is given in Figure 49 above. Another balance of the cost on stops versus delay in the settings of the self-optimising systems would probably give another result, at the cost of less delay savings. The results indicate that future work on self-optimising control should investigate optimal balance those costs.

7.6.7. Impact of PTSP with PRIBUSS around S:t Eriksplan

The impacts of the existing, poorly tuned PTSP implementation (see chapter 7.6.6) with the PRIBUSS method were also evaluated in case study four. In the whole simulated network where three intersections have PTSP, the average delay time for prioritised buses was reduced by 17 seconds AM, 25 seconds midday, 11 seconds PM, 14 seconds PM +5% and 9 seconds at night respectively. The delay for other traffic was increased under all but one traffic condition. The average delay per vehicle in the simulated network is given in Figure 50 below.
Figure 50: Average delay per vehicle with and without PRIBUSS PTSP around S:t Eriksplan

The total delay for all road users (weighted travel time) has been calculated in the way described above, and the results are shown in Figure 51 below. The PTSP reduces the total delay under all but one traffic condition. However, the travel time savings at S:t Eriksplan in this case study are much lower compared to the travel time savings on Fleminggatan in case study one, two and three. An updating of the PRIBUSS settings to join them with the 2006 signal plan would probably improve the outcome of the PTSP considerably, but it has not been studied in this thesis.

Figure 51: Total delay in person hours with and without PRIBUSS PTSP around S:t Eriksplan
8. Discussion and conclusions

8.1. General conclusions regarding PTSP

The case studies in this thesis support the conclusions, found in literature, that PTSP can considerably reduce the travel time for buses, at the expense of slightly increased travel time for other traffic. Principally PTSP reallocates green time to the bus approach from other approaches by green extension or red truncation, and one could therefore imagine that other traffic following the bus route would also benefit. In some cases this is true, but the reallocated green time is not necessarily useful for traffic other than the prioritised bus. If the signal is green, but no vehicle is present there is no use for this extra green time.

The changes in green times made by the PTSP strategy to prioritise buses do not only redistribute green time between the approaches, they redistribute green time over cycles as well. If the green time for the cross street is determined by the minimum time for pedestrians to cross, it cannot be shortened but only moved forwards in the cycle. This will affect the bus approach in the next cycle by delaying the green period, and deteriorating the green wave progression. This will mostly affect vehicles travelling in the opposite direction to the prioritised bus. Green time compensations to avoid cross street queue accumulation will have similar effects, and therefore need to be carefully timed. This shows one conceptual error in most implementations of PTSP in coordinated systems with conventional control strategies; the PTSP changes of green times are made one intersection at a time without explicitly making any changes of green times in adjacent intersections to maintain green wave progressions. However, a PTSP strategy applied within conventional control strategies that could make green time changes in several intersections would be complex though, while it would be one of the strengths of self-optimising control strategies.
8.2. The cost and benefits of different PTSP functions
The implementation of PTSP is unique in every intersection, but some general conclusions could be drawn from the simulation experiments made in case study two. The travel time savings for buses can vary significantly between different intersections and approaches depending on geometry, coordinated signal timings, bus stop location etc. In case study two, the travel time savings in the main intersection was fifty-five seconds in one direction of travel but only six seconds in the other.

The travel time costs and benefit are indeed different for the different PRIBUSS functions. Short green extensions give relatively large benefits for the prioritised buses at very small cost to other traffic. The detection point where the bus requests green extension needs to be far upstream from the intersection in order to execute a priority measure in time. The extension time cannot be longer than the driving time from the bus stop to the stop line if there is a nearside bus stop, and if priority request upstream of bus stops cannot be used. The usefulness of green extension in combination with nearside bus stops is therefore limited.

Long green extensions, including re-taken start, gives larger benefits for the prioritised buses, but also larger negative impacts on other traffic than short extensions. With long extensions cross street green time compensations may be needed, and they require careful timing to avoid extra delay. The number of buses gaining from green extensions is small, but the benefit for each bus is high.

Red truncation gives less travel time savings to the prioritised buses and more extra delay for cross street traffic compared to green extensions. Red truncation can be useful also when the detection point is close to the intersection, which is the case for approaches with nearside bus stops. The number of buses gaining from Red truncation is large, but the benefit for each bus is small.

The more complex functions, Double red truncation and Extra phase give some additional travel time savings to the prioritised buses but have the largest negative impact on other traffic, especially at saturated traffic conditions.
A combination of PRIBUSS functions gives better results however, compared to using only one of the functions. A bus given green extension in one intersection may need a red truncation in the next intersection; otherwise the travel time gained in the first intersection can be lost in the other. The use of the different PRIBUSS functions depends on the structure of the signal coordination in combination with the location of the bus stops and the dwell time at these.

8.3. To prioritise between buses

The “bunching” phenomenon is well known for routes with high frequency bus service. If the bus headways could be adjusted with smooth “holding” and “pushing” at the traffic signals by PTSP before the headway deviation becomes too large, then bunching can be avoided. To achieve this, the PTSP needs to be conditional with respect to the headway and/or on-time performance of each bus, and appropriate conditions should be set to give priority or not. For a bus line with five minute headways, a two minute threshold is probably too large to avoid bunching.

Giving signal priority to one bus implicitly reduces the possibility to give priority to another bus. This is due both to the restrictions on changing the signal timings too different from the normal coordinated plan, and the risk of conflicting requests for priority (see chapter 3.3.2). In case study one this effect can be observed by the larger travel time reduction when PTSP is allowed in the “own” direction of travel only. Since there is no bunching of the buses in this case study, the impact of conflicting requests for priority is caused by buses in the opposite direction of travel only. In case study three this effect could be observed by the increased travel time for Blue buses when Red buses were given moderate priority with short green extensions, as well as by the reduced travel time for other buses when priority for early buses where restricted. If buses with less need for shorter travel time, i.e. early buses, are given less priority, there are also larger possibilities to give priority to other buses with greater need for shorter travel time, i.e. late buses.
The proposed “differential on-time-status based conditional priority strategy” in case study two showed only slightly better results for buses, compared to when stricter thresholds for giving priority at all were implemented by changing the threshold from two to one minute ahead of their time table. The negative impact on other traffic was also reduced by restricting the use of the more disruptive PTSP functions. As implemented in the case study, the restrictions on the PRIBUSS functions (guarantee time for cross streets, last time step etc.) were the same in all tested scenarios. With more restrictions on which buses are given priority, there could be fewer restrictions put on the functions. Some functions would thereby be more disruptive for other traffic, but used by only a few buses. Limiting the priority to buses with a higher need for travel time reductions seems to give larger benefits with less negative impacts, but this need to be studied further.

If on-time-status based PTSP shall be used to improve regularity, a well-prepared timetable with correct timings at all stops is needed. This is not always the case today; in the example from case study three almost all buses were late on Fleminggatan, but on-time upstream and downstream of this rout segment. Obviously the time table was not correctly set over the whole route. If all buses are late according to the time table, it is not possible to distinguish the buses that need to be sped up, so all buses will get the same level of priority.

However, preparing a timetable with half minute accuracy at all stops requires a lot of work. If an even headway strategy as proposed by Cats et. al. (Cats, Larija, Koutsopoulos, & Burghout 2011) is used instead of a static time table for high frequency lines, the problem of preparing a timetable well timed at all points of the line is avoided.

### 8.4. The usefulness of SIL simulations

The impacts of PTSP seem to be rather different in relatively similar intersections, and they depend on the individual PRIBUSS settings in each intersection. Those are set by the traffic engineer based on experience, but there is no method to determine optimal settings. Due to the complex nature of dynamic changes, such as PTSP, in a coordinated traffic signal system the impacts are hard to foresee in detail. One possibility would be to use SIL simulations to test different settings of the PTSP systematically before implementing them in field.
Programming errors in the signal controller or improper design of the PTSP are hard to observe in field. Since the frequency of the use of each PTSP function is low, long observation periods are needed to observe malfunctions. In the signal controller in the intersection Fleminggatan – Polhemsgatan there were two different errors in the PRIBUSS programming, which were not spotted until several years later when the SIL simulations were made. The situation is probably similar in other areas that have not been studied. Systematically use of SIL simulations could be useful to find and correct such design or programming errors.

Simulations with a SIL setup also enable comparisons of signal control strategies or systems on equal terms under controlled conditions. A field evaluation of several different signal control strategies as in case study four would not be practically or financially possible.

8.5. Self-optimising control
The aim in case study four was to test if commercially available self-optimising signal control systems could be used to minimise the overall delay for all road users. Both tested systems; Utopia/Spot and ImFlow reduced the total delay for all travellers compared to the existing signal control. The self-optimising control managed to extend the bus priority and reduce the waiting time for pedestrians and cyclists, as requested by the Adapt project. However, with the tested settings the delay and number of stops for cars increased considerably. An increase in number of stops implicitly results in increased fuel consumption, emissions and noise. Another balance between the stop costs and delay in the programming of the self-optimising signal control systems might give a better performing signal control strategy in total, but less delay time savings. Future work should address the optimal trade-off between stop and delay.
A cost-benefit analysis can be done to evaluate the effect of self-optimising signal control. Standardised monetary values on stop and delay can be used to balance the control strategy in a socioeconomic "optimal" way. Using the standardised values on delay for different modes of road users do not necessarily give the same optimum that the City of Stockholm requests; prioritising surface efficient modes of transport in order to implement the Urban mobility strategy. To some extent it is a political decision which modes of transport should be prioritised in the urban streets. Both tested adaptive signal control systems seem to be useful tools to allocate the accessibility between different travel modes in a different way, and minimise the overall delay better than conventional coordinated signals.

The self-optimising control does not produce “nice” green waves as the conventional coordinated control does, and what traffic signal engineers traditionally look for when tuning the signal timings. Instead the control strategy tries to minimise the road user cost, which is close to minimise delay, at any given moment. This results in constantly changing signal timings, adapting the green times to arriving buses. However, this also results in an unsettled signal timing, which makes it difficult for the road users to learn and predict the signal behaviour.

Extra flexibility can be obtained if the self-optimising signal control strategy is also allowed to optimise the signal stage order in the system. ImFlow has this ability, but the impacts of optimised stage order needs to be further evaluated in terms of delay, capacity, and traffic safety.

Self-optimising control seems to have two major advantages over conventional coordinated control; it can adjust the cycle time to the capacity needed at the moment, and it can handle extensive PTSP. This leads to the conclusion that implementation of self-optimising control has it largest potential to give benefits at locations where there are fluctuations in traffic flows and/or where there is a need for extensive PTSP.

There should also be large potential with self-optimising control to prioritise between buses with different signal priority needs, but this needs to be evaluated further.
9. Future research needs

9.1. Criteria for signal priority
In this thesis “on-time status” has been used as criterion for giving priority or not in order to regulate bus operation and consequently reduce delay for passengers. However, the impact of irregularity on passenger delay is different for different bus lines, and for different locations on the bus line. A practical method to calculate average additional travel time (AAT) impacts on passengers, including expected passengers waiting at bus stops along the line, should be developed in future research. The AAT impact would then be used as a basis for criteria for giving priority.

The potential benefits of using PTSP for regulation, or to reduce additional travel time in general, should be further studied. The waiting time at bus stops along the line should be included as well as the travel time spent on the bus. To capture the effects on bunching and the boarding time dynamics from regularisation with PTSP, the model should cover more than a few intersections. Moreover the combined effects of even-headway holding strategies and regularisation with PTSP should be studied. A mesoscopic model such as Bus Mezzo is probably an appropriate tool for those studies.

9.2. PTSP in conventional control strategies
The proposed and tested new differential on-time-based conditional priority strategy should be further developed including fewer limitations on the PRIBUSS functions used only by few buses. The “on-time” criteria should also be revised and other values tested. There is also a need for some balancing between conflicting priority requests. This can be studied with further SIL simulations.

Moreover the results obtained from the case study simulations need to be verified by field tests. The benefits of stricter conditions for granting priority, i.e. changing the threshold from two to one minute ahead of schedule, could easily be tested in the field. If field studies also show good results the Stockholm PTSP implementation could be changed in near term future.
9.3. Self-optimising signal control

The self-optimising systems tested in case study four were compared with the existing conventional control used in field. However, the time plans used were developed some years ago with traffic flows other than the actual, and need to be updated to facilitate a more fair comparison. Such study could distinguish the effect of ageing time plans in the conventional control from the effect of the PTSP. In addition, some aspects of the self-optimising control could be updated and improved from the experience from case study four. This can be studied with further SIL simulations.

The possibilities to use PTSP for bus service regulation are probably greater with self-optimising signal control than with conventional control strategies. I.e. in the self-optimising system a bus that would benefit more from travel time reduction could be given a higher weight in the optimisation algorithm compared to a bus that would benefit less. This should be further studied.

There is still a need to evaluate the self-optimising control with field studies in order to verify the simulation results, providing that the City of Stockholm decides to continue with a field implementation of one, or both, of the tested systems.
10. Contribution of the thesis

In paper I "Hybrid Traffic Simulation with Adaptive Signal Control" (Burghout & Wahlstedt 2007) a hybrid mesoscopic-microscopic SIL simulation model including adaptive (VA) traffic signal control with PTSP was implemented. The entire Stockholm area were simulated in the mesoscopic simulation model MEZZO, while the area of specific interest around three intersections was simulated at microscopic level in VISSIM connected to EC1-simulator signal controller simulators running the signal control algorithms. With the integrated hybrid simulation model the network effects of better, adaptive traffic signal control was studied. The old FT control scheme was compared to a new adaptive control scheme also implemented in field. The capacity with old TF control was below traffic demand during the morning peak period causing long queues in inbound direction resolved with the new adaptive control. Travel times in peak direction through the area of specific interest were thereby reduced with more 50 per cent. Due to the shorter travel time traffic was attracted from other routes far away from the studies area. The traffic flow redistribution effects of the reduced travel time were captured by the mesoscopic model. Traffic flows in peak direction showed an increase by 23%, but the travel time was still 51% shorter despite the higher traffic volumes. The study showed the possibilities of using hybrid models to study network effects of improved signal control.

In paper II “Impacts of bus priority in coordinated traffic signals” (Wahlstedt 2011) the method with SIL simulations is developed and implemented in Stockholm street network. The impacts of PTSP with the PRIUBUS method on travel times for buses and other traffic are evaluated. Also the effects on traffic in the same or opposite direction of travel as the prioritised bus are studied. The results show travel time reductions for the prioritised buses, and for other traffic following the bus route in one direction. The travel time increases for traffic on the major cross street, and other traffic following the bus route in the other direction. With PTSP for buses in one direction only, the travel time savings are generally larger for buses in the prioritised direction of travel compared to PTSP in both directions. Buses in the opposite direction of travel show longer travel times also compared to the scenario without PTSP. The impacts on other traffic following the bus route are similar as on the buses.
In paper III “Evaluation of Different Bus Priority Functions in Coordinated Traffic Signals” (Wahlstedt 2013a) the impacts of the different PRIBUSS functions (green extension, red truncation etc.) were evaluated. The impacts of each PRIBUSS function was tested one by one and all functions together, first in one intersection and then in all intersections at Fleminggatan. Green extensions give the buses considerably shorter travel times with very little impact on other traffic, while red truncation gives less travel time reduction for buses and longer travel times for cross street traffic. Double red truncation and extra phases decrease bus travel times, but have larger negative impacts on other traffic. A combination of functions gives the best results. Furthermore, different criteria for giving signal priority is evaluated, and a new on-time-status based differential priority strategy is proposed. Tightened threshold for giving signal priority holds early buses back smoothly and reduces travel time for later buses as well as other traffic. The proposed new criteria show slightly better results than tightened thresholds, and need to be further developed to achieve the full potential.

In paper IV “Evaluation of the Two Self-Optimising Traffic Signal Systems Utopia/Spot and ImFlow, and Comparison with Existing Signal Control in Stockholm, Sweden” (Wahlstedt 2013b) the impacts of self-optimising control, including PTSP, in a network of coordinated traffic signals were studied. The suppliers of signal control equipment on the Swedish market were invited by the City of Stockholm to participate in a “competition” where their systems would be tested in a SIL simulation environment against each other, and against the existing conventional coordinated control including PTSP. The suppliers should provide a simulator version of their system possible to connect to VISSIM, as well as programming of the system with the settings the suppliers found optimal. The aim was to test if commercially available self-optimising signal control systems could be used to minimise the overall delay for all road users. The self-optimising control managed to extend the bus priority and at the same time reduce the waiting time for pedestrians and cyclists, as requested by the project. The total delay for all road users was considerably reduced with the self-optimising control. However, with the used settings the delay for cars and number of stops increased significantly. ImFlow performed better than Utopia/Spot in the evaluation.
In the Swedish Transport Administration guidelines on capacity analysis with simulations, chapter 6 on signalised intersections "TRV 2013:79994, Handbok för kapacitetsanalys med hjälp av simulering, Kapitel 6 Signalreglerade korsningar" (Wahlstedt 2014) guidelines are given for when and how microscopic simulations should be used to analyse the capacity and delay of traffic signal controlled intersections. Swedish regulations for traffic signal control are introduced. Some advice on how to simplify the signal control under certain conditions is given, and a method to model vehicle actuated traffic signals as fixed timed are presented. The data needed for microscopic simulation of traffic signal controlled intersections are specified. Practical advice is given on how to model traffic signal controlled intersections in VISSIM, and how to verify and calibrate the model.

The main contributions of this thesis are:

1. A comprehensive review and study of unconditional and conditional PTSP criteria and strategies
2. Development and implementation of simulation based methods for evaluation of impacts of different PTSP strategies in coordinated traffic signal networks in Stockholm
3. Evaluation of the impacts on buses and other traffic of different PTSP functions in conventional control strategies
4. Proposal for an “on-time-status” based PTSP method focusing on direct travel time savings as well as reduced bunching
5. Analysis and comparison of impacts of conventional control strategies including PTSP with self-optimising signal control methods
6. Implementation of a hybrid mesoscopic-microscopic -SIL simulation model capable of including the traffic redistribution effects of changed signal control
11. Glossary of terms

Some specific traffic signal terms are described below with their meaning used in this thesis.

**Active bus priority** – Bus priority measures in traffic signals made only when a bus is present. Requires selective detection of the bus.

**Adaptive signal control** – A control strategy adapting to the actual traffic. Sometimes also used as a synonym for Self-optimising control, see below.

**AVL system** – Automatic vehicle location system; many buses have a positioning system on board used for real time information systems, automatic announcing of stops, calls for traffic signal priority, statistic collection etc.

**Bus/buses/bus priority** – What is stated for buses in this thesis will also be valid for trams, or public transport vehicles in general. The shorter term “bus” is used to make the text easier to read.

**Capacity** – The highest stationary flow of vehicles (or people) that can traverse an intersection or other part of the road network under given, saturated conditions.

**Coordinated controlled signals** – Traffic signals controlled with consideration of adjacent traffic signals in order to facilitate the passage through the system of coordinated signals. Contradiction to isolated controlled signals.

**Cycle time** – The length in time of a complete stage sequence before it is repeated. Normally given in seconds.

**Degree of saturation** – Ratio of flow to capacity.

**Delay** – Additional travel time experienced by the road user. In a signalised intersection the delay can be divided into geometrical delay due to the physical design (eg. The driver needs to slow down to turn around a corner), or control delay due to the signal indication shown and interaction with other traffic (eg. Waiting time at red light and/or queue in the approach).
Detector – Device used to detect the presence or passage of a vehicle (or person).

Detector interval – The time period a detector (logic) will be active after being actuated by a vehicle.

Discharge rate – The number of vehicles per time unit passing the stop line.

Extension interval – The time period all detectors working together in an approach will be active after being actuated by a vehicle. If the distance between vehicles is shorter than the extension interval the detector will extend green (if allowed by the control logic).

Effective green – The time when traffic can be discharged at saturation flow rate. Actual green time minus start loss plus end gain plus the part of amber time used to drive.

Effective red – The part of the signal cycle not effective green. Cycle time minus effective green.

Fixed timed (FT) signal control – A control strategy where the signal timings are fixed and pre-set.

Isolated controlled signals – Traffic signals controlled independently of other traffic signals. Contradiction to Coordinated controlled signals.

Late release – An arrangement of stages whereby the green period in one direction starts after that for the opposed traffic.

LHOVRA – The most common control strategy for isolated VA traffic signals in Sweden. It is a conventional signal group based strategy using demand of green and green extension in each signal group individually. LHOVRA is an acronym (in Swedish) for the included functions.

Lost time – The time of the signal cycle where no discharge flow takes place in any conflicting signal group.

Optimal cycle time – The cycle time giving the minimum delay for all vehicles (or all road users) under given conditions.
**Optimal green split** – The ratio of effective green time in the approaches (or stages/signal groups) giving the lowest delay for all vehicles under given conditions.

**Passive bus priority** – Bus priority measures made regardless of a bus is present at the moment or not. Can be physical measures such as bus lanes or passive bus signal priority.

**Passive bus signal priority** – Bus priority in a traffic signal made regardless if a bus is present at the moment or not. Can be extra green time in approaches with bus traffic or signal coordination made for typical bus speed.

**Past end green (PEG)** – Green extension after the normal time for signal change from green to amber.

**Phase** – In British English equivalent to “Signal group” in American English, in American English equivalent to “stage” in British English. Not used in this thesis to avoid confusion.

**Platoon** – A group of vehicles (or pedestrians) moving together.

**Progression factor** – The proportion of vehicles arriving at green light. Often used as a term to describe how “good” coordinated a signal is.

**PTSP** – Public transport traffic signal priority.

**Saturation flow** – The maximum number of vehicles per time unit that can pass the stop line under steady, saturated conditions at continuous green signal. A continuous upstream queue is supposed to exist and no lost time is considered.

**Self-optimising signal control** – A signal control strategy with a mathematical algorithm minimising a cost function (including delay, stops etc.) that is used to determine optimal timings of the signal changes.

**Signal group** – Used in the control logic as the “smallest part” controlled individually. All signal heads controlled by the same signal group will always change signal at the same time.
Signal head – Hardware device that displays signal information, normally red, amber and green, to the road users.

Stage - Part of the signal cycle when a set of movements are given green signal simultaneously. No signal changes can occur in a stage.

Stage sequence – The order in which the stages are given green signal.

Vehicle actuated (VA) signal control – A control strategy where the signal changes are set by the actual traffic detected by the signal controller.
12. References


Altun, S, Furth, P. (2009). “Scheduling Buses to Take Advantage of Transit Signal Priority”. In Transportation Research Record: Journal of the Transportation Research Board, No. 2111, Transportation Research Board of the National Academies, Washington, D.C. USA, pp. 50-59

Angus P. Davol (2001), “Modeling of Traffic Signal Control and Transit Signal Priority Strategies in a Microscopic Simulation Laboratory”, degree project, MIT, Boston, USA


Linse, L.(2013). “Capcal 4.0 Användarhandledning” (Capcal 4.0 user guide), Trivot2013:87, Lund, Sweden


SCATS (2014), “SCATS 6 Functional Description”, Roads and Maritime Services, New South Wales, Australia


Trafikverket (2013). (The Swedish Transport Administration) “Metodbeskrivning för beräkning av kapacitet och framkomlighetseffekter i vägtrafikanläggningar” (Method description of capacity calculation on roads) publication TRV2013/64343, Sweden

Tveit, M., Ø. (1999). “Bruk av adaptiv trafiksignalregulering i byområder” (Use of adaptive traffic signal control in urban areas, Phd thesis), Doktoravhandling ved NTNU 1999:15, Trondheim, Norway


Wahlstedt, J. (2005). “Simulering av bussprioritering i trafiksignaler” (Simulation of bus priority in traffic signals), Master’s theses, KTH Royal Institute of Thecnology, Stockholm, Sweden


Paper I

Hybrid traffic simulation with adaptive signal control

WILCO BURGHOUT AND JOHAN WAHLSTEDT

Published in
Transportation Research Record:
Journal of the Transportation Research Board, No. 1999,
Transportation Research Board of the National Academies, Washington, DC. USA, pp. 191 – 197
Hybrid Traffic Simulation with Adaptive Signal Control

Wilco Burghout and Johan Wahlstedt

A hybrid mesoscopic–microscopic model is implemented that applies microscopic simulation to areas of specific interest while simulating a large surrounding network in lesser detail with a mesoscopic model. The hybrid model integrates VisSim, a microscopic traffic simulation model, and Mezzo, a recently developed mesoscopic model. The hybrid model is applied on a network in which Mezzo simulates the entire area (6,000 links) of Stockholm, Sweden, and VisSim simulates the area of specific interest, containing three intersections with adaptive signal control with bus-priority functions. The adaptive signal control and bus-priority functions are simulated by a separate signal controller simulator (EC-1 simulator) that interacts with the hybrid Mezzo–VisSim model and thereby provides the actual signal changes that would take place in the field. Two alternative control schemes are evaluated with the hybrid setup: the original fixed-time control and the new adaptive control. The results show clear improvement in terms of travel times, delays, and stops with the new adaptive control scheme. They also show that although these improvements for the local (microlevel) area attract additional traffic from the surrounding (meso) area, the net effects both locally and networkwide remain positive in terms of travel times, average number of stops, and delay. Moreover, this study demonstrates the advantages of hybrid simulation in evaluation of complicated adaptive traffic control in which both local detailed effects and network effects need to be studied.

Traffic simulation has become popular for modeling the operations of dynamic traffic systems. Traffic simulation models can be classified as macroscopic, mesoscopic, or microscopic. Macroscopic (macrolevel) models such as Strada (1) and Metacor (2) tend to model traffic as a continuous flow, often using formulations based on hydrodynamic flow theories. Mesoscopic models such as DynaMIT (3, pp. 19–36) and Dynasmat (4) model individual vehicles but at an aggregate level, usually by speed–density relationships and queuing theory approaches. Microscopic (microlevel) models such as MITSimLab (5) and VisSim (6) capture the behavior of vehicles and drivers in great detail, including interactions among vehicles, lane changing, response to incidents, and behavior at merging points. Because of this level of detail in the representation of traffic dynamics, microscopic models are appropriate for evaluation of intelligent-vehicle systems (ITS) at the operational level, since the representation of many dynamic traffic management systems requires such fine-grained modeling of the traffic process.

However, because of the detailed nature of microscopic models, the preparation of input data (e.g., network coding and representation) can be time-consuming and tedious. In addition, microscopic models are highly sensitive to errors or variation in input demand data, especially under congested conditions, and their calibration is not trivial. Therefore, microscopic models are usually applied to smaller networks. Macroscopic and mesoscopic models usually have fewer parameters to calibrate and are less sensitive to errors in network coding or demand variations. However, because of their more aggregate nature, such models are limited in their ability to capture the detailed behavior needed to study traffic networks with dynamic traffic management capabilities.

Recently, hybrid macroscopic–mesoscopic models (7–12) and mesoscopic–microscopic models (13–15) have appeared, attempting to combine the strengths of macroscopic or mesoscopic simulation (large networks, less sensitivity to network coding errors, easier calibration) with those of microsimulation (greater detail, ability to model and evaluate ITS and adaptive traffic control). Such hybrid models enable the simulation of large-scale networks, incorporating the effects of local microphenomena. This feature increases accuracy and validity while reducing the required data collection and calibration effort of the overall model.

Until now, these hybrid models have been tested mostly by using very simple networks, often consisting of a number of consecutive links, where one or more links in the middle of the sequence are modeled in microsimulation and the others in mesoscopic or macro-simulation (7–12). In most cases the microscopic models are kept simple, not even including lane changing. However, in the work of Burghout et al. (13) and Burghout (14) the hybrid model (Mezzo–MITSimLab) was applied to a small Stockholm, Sweden, network in which the south part consisting of intersections and a roundabout was simulated in the microscopic model, MITSimLab (5), and the remainder in the mesoscopic one, Mezzo (14).

In this study the hybrid framework presented by Burghout et al. (13, 14) is implemented by using a different microscopic model, VisSim (6, 16). The replacement of the open-source MITSimLab with a commercial microscopic model has had a number of consequences for the implementation of the framework, which will be discussed here. In addition, a simulator program for the adaptive signal control (EC-1 simulator) is interfaced with the hybrid model, thereby providing the actual signal changes that take place in the field. The new implementation of the hybrid model (named InterMezzo) is applied on a network in which Mezzo simulates the entire Stockholm area (6,000 links) and VisSim simulates the area of specific interest containing three intersections with adaptive signal control with bus-priority functions. The city of Stockholm has changed the signal control scheme to improve traffic performance of the intersections. This new scheme is evaluated against the old signal control by using...
the advantages of the hybrid model combined with the signal control logic as it is implemented in the field.

**Mesoscopic–Microscopic Integration Framework**

In the work of Burghout et al. (13, 14) the conditions for a consistent hybrid mesoscopic–microscopic simulation model were presented, and a hybrid integration framework was proposed that satisfies these conditions. In this section the requirements and framework are introduced briefly. The requirements for a successful hybrid model are

- Consistency in route choice and network representation,
- Consistency of traffic dynamics at mesoscopic–microscopic boundaries,
- Consistency in traffic performance for mesoscopic and microscopic submodels, and
- Transparent communication and data exchanges.

**Consistency in Route Choice and Network Representation**

In earlier work (13, 14) a general architecture is proposed consisting of a separate module that contains elements common for the mesoscopic and microscopic models: a database with the network graph, the travel time tables, the set of paths and the origin–destination (O-D) flows, as well as a travel behavior component with route choice models and path generation algorithms. This general architecture is applicable when new models are developed with integration in mind.

In addition, the earlier work (13, 14) presents a simplified framework for the integration of existing models that minimizes intermodel communication overhead and uses functionalities in both models (such as route choice) in a consistent manner.

In the simplified framework the mesoscopic model includes the O-D matrix for the entire network. This assumption is by no means restrictive, since an origin or destination node in the microscopic area can always be designated as a boundary node in the mesoscopic area connected directly to the microscopic subnetwork.

In addition, the mesoscopic network includes virtual links for each path connecting boundary nodes in the microscopic network. This representation guarantees that each relevant path through the microscopic model is represented correctly in the mesoscopic route choice. The mesoscopic model collects travel times for the virtual links from the corresponding paths in the microscopic network and uses them in the route choice like any other link in the network.

Under this simplified architecture the mesoscopic model is solely responsible for all pretrip routing decisions. En route decisions are the responsibility of the respective subnetwork and paths inside the mesoscopic model. In this study the simplified architecture is implemented by using the mesoscopic model for the overall routing.

**Consistency of Traffic Dynamics at Mesoscopic–Microscopic Boundaries**

In the earlier work by Burghout et al. (13, 14) the modeling of traffic dynamics at mesoscopic–microscopic boundaries is discussed in detail. One of the main sources of potential inconsistencies at this interface is the correct determination of crossing-vehicle attributes in both directions (microscopic → mesoscopic) and (mesoscopic → microscopic). Attributes such as speeds, accelerations, headways, and so forth should be consistent with the prevailing conditions in the new segment to which the vehicle is moving, otherwise unnecessary shock waves may propagate upstream.

On the boundary from the mesoscopic to the microscopic model (mesoscopic → microscopic), submodel information is exchanged in both directions. Information about vehicles (with a certain speed and at certain time intervals) and about blocking of boundaries and downstream density needs to be communicated from the mesoscopic to the microscopic model. If the entry to the microscopic link (downstream of the boundary point) is blocked, the mesoscopic model needs to stop vehicles from exiting. The microscopic model informs the mesoscopic model when the blockage is removed so that vehicles can start flowing again (over that specific boundary). The microscopic model also sends the density in the vicinity of the boundary to the mesoscopic model, where it is used to calculate the speed of the shock wave that propagates upstream.

A more complicated issue is the generation of information that is needed in the microscopic representation of traffic but missing in the mesoscopic model. The microscopic characteristics that need to be generated at the entry to the microscopic model are divided into vehicle and driver attributes and model variables. Attributes such as desired speed are generated independently in the microscopic model based on the distribution of these characteristics in the general driver population assumed by the microscopic model. Model variables such as the vehicle’s speed, acceleration, and time headway to the vehicle in front need to be in accordance with the traffic situation upstream and downstream of the boundary.

In the interface from the microscopic to the mesoscopic model (microscopic → mesoscopic) similar conditions need to be met as mentioned for the mesoscopic-to-microscopic case. First, the mesoscopic model needs to inform the microscopic one each time the downstream mesoscopic link becomes blocked or unblocked, so that the microscopic model can stop or start sending vehicles at the right moments. In addition to the blocking, the vehicles in the microscopic model that move toward the exit to the mesoscopic model need to react to the downstream traffic conditions, as they would if the downstream link were microscopic as well. In that case the vehicles would react to vehicles in front by using their car-following logic, and those vehicles would react to vehicles in front of them, and so forth. In the mesoscopic model, however, the position and detailed behavior of vehicles are not usually modeled, but it is known when the last vehicle arrived and the (average) speed it was assigned. With this information a virtual vehicle is projected in the “imaginary” continuation of the microscopic link. Microscopic vehicles that are near the exit react to the virtual vehicle as if it were a normal vehicle in front of them. This concept can be extended to provide one virtual vehicle for each lane on the microscopic link.

**Consistency in Traffic Performance and Transparent Communication and Data Exchanges**

In addition to the conditions discussed in the two previous subsections, the mesoscopic and microscopic submodels need to be calibrated carefully to ensure that facilities that can be modeled by both submodels have similar capacity in both mesoscopic and microscopic models. Furthermore, the models need to communicate information (vehicles, traffic conditions, etc.) efficiently without too much overhead. More information may be found elsewhere (13, 14).
IMPLEMENTATION OF HYBRID FRAMEWORK

The framework described in the earlier work (13, 14) and briefly introduced in the previous section is implemented by using Mezzo (14), an event-based mesoscopic model especially developed for hybrid modeling, and VisSim (16), a commercially available state-of-the-art microscopic model.

Burghout and coresearchers (13, 14) used as the hybrid prototype the MITSimLab microscopic model, for which the entire source code was available. This method enabled a full integration of the common components of the microscopic and mesoscopic models (such as the route choice) and easy adaptation of the models for vehicle dynamics to allow for improved vehicle loading at the microscopic → mesoscopic boundaries and virtual vehicles at the mesoscopic → microscopic boundaries. Although this implementation was versatile, MITSimLab is for the most part an academic model with few commercial users. For hybrid modeling to become available to a wider public of microscopic model users, the hybrid framework was redeveloped and reimplemented by using VisSim.

Although it is not possible to change the source code of VisSim to obtain the desired hybrid functionality, VisSim does offer a component object model (COM) application programming interface (17), which is an interface in which the user gains access to all internal objects with any type of programming language that can manipulate COM objects (such as C++, Visual Basic, or Visual Basic for Applications). This feature means that most objects available within VisSim (links, vehicles, paths, etc.) can be accessed, and their attributes manipulated, but no new object types can be added nor can any objects (such as vehicles and their car-following behavior) be modified (added attributes, different functionality). This aspect implies that although most of the proposed architecture could be implemented directly, some of the functionality (such as virtual vehicles and paths) had to be implemented in a different way.

Network and Route Representation

The routes in VisSim consist of a series of user-defined “abstract” nodes on top of the link-based network, usually one for each intersection. The standard network representation of VisSim is based on links and link connectors only. Routes are part of the dynamic traffic assignment add-on, which enables users to define O-D demand via parking lots in the network and routes between these parking lots by using superimposed nodes. The implementation of the hybrid framework uses the mesoscopic virtual links as described earlier. These virtual links represent the paths inside the VisSim network, connecting inbound (mesoscopic → microscopic) and outbound (microscopic → mesoscopic) boundaries. The virtual links are then part of the mesoscopic network representation and are treated by the route choice model in the same way as the regular mesoscopic links. In particular, each virtual link description consists of

1. The outbound and inbound mesoscopic boundary nodes,
2. The start and end parking lots in VisSim, and
3. A sequence of nodes representing the path inside VisSim from the start to the end parking lot.

In the mesoscopic model (Mezzo), a list is maintained of all vehicle objects currently in VisSim. When they reenter the mesoscopic network, their travel times are logged for the virtual link they have been on. This log allows the mesoscopic model to account for the pretrip routing throughout the model [more details may be found elsewhere (14)]. Burghout and coworkers (13, 14) presented an additional construct, microscopic virtual links, to handle en route diversions inside the microscopic model, since such diversion may require a different exit point into the mesoscopic network. Although it is possible to implement this construct in the future, it was deemed outside the scope of this implementation, especially since the size of microscopic networks used in a hybrid setting (a few intersections) usually does not require this functionality.

Traffic Dynamics at Mesoscopic–Microscopic Boundaries

Vehicles arrive in the microscopic model from the mesoscopic model with time headways determined by the node servers in Mezzo, where each outgoing lane has its own (stochastic) server process [more details may be found elsewhere (14)]. Whereas in MITSimLab the vehicle loading was modified to generate initial speeds that were in accordance with the time headway from the leading vehicle in the selected lane, in VisSim these adaptations are much more difficult to make because in VisSim only existing objects and their attributes can be accessed and modified through the COM interface, not the internal functionalities such as the vehicle loading mechanism. In addition, as shown by Burghout and Koutsopoulos (18), the standard loading of vehicles in VisSim generates much less artificial deceleration and fewer capacity problems than the original MITSimLab method does. When a vehicle crosses the boundary to VisSim, it is created in a parking lot and assigned its (VisSim) path, whereupon the standard loading mechanism assigns it a lane and a speed [details may be found elsewhere (18)]. The parking lots are modeled as “abstract lots,” meaning that the vehicles do not start from a standstill but leave immediately at an appropriate initial speed (16, 17).

Queue spillback across the mesoscopic–microscopic boundaries is taken care of by checking at the end of each VisSim time step to see if there are any vehicles in entry parking lots that have not been able to leave (because of congestion ahead). If this is the case, the virtual links to which the queued vehicles belong are suspended (blocked), so that no more vehicles for these virtual links are sent to the microscopic area until the parking lots become unblocked and the queued vehicles have left. When a virtual link is suspended, the upstream Mezzo link stops sending vehicles to it, and a queue starts to build up on that link. After the blockage disappears, the start-up shock wave that traveled in the microscopic model (VisSim) toward the mesoscopic–microscopic boundary continues inside the mesoscopic area, ensuring a correct arrival process at the mesoscopic–microscopic boundary [see thesis by Burghout (14) for details].

In the section on the mesoscopic–microscopic integration framework, the concept of virtual vehicles was described to ensure that vehicles exiting from the microscopic into the mesoscopic model have the appropriate speed (according to the speed in the downstream mesoscopic link). As with vehicle loading, this concept is difficult to implement in VisSim since it requires modification of the car-following model, allowing a vehicle to follow a virtual vehicle projected past the end of the link. Instead, the same idea was implemented by modifying the speeds of the exiting vehicles directly according to the following rules.

For each vehicle on an exiting link in VisSim,

\[ V = V_0 \quad \text{if } X > X_{\text{breakhead}} \quad \text{or if } V_0 \leq V_{\text{Mezzo}} \]  
\[ V = V_{\text{Mezzo}} \quad \text{if } X < X_{\text{critical}} \quad V_0 > V_{\text{Mezzo}} \]  
\[ V = \alpha V_0 + (1 - \alpha) V_{\text{Mezzo}} \]
where

\[ \alpha = \frac{X}{X_{\text{lookahead}}} \quad \text{if} \quad X_{\text{critical}} < X < X_{\text{lookahead}} \quad \text{and} \quad V_0 > V_{\text{Mezzo}}, \]

\[ V = \text{new speed of vehicle}, \]

\[ V_0 = \text{initial speed of vehicle}, \]

\[ V_{\text{Mezzo}} = \text{speed on downstream link in Mezzo (which is a function of density on that link)}, \]

\[ X = \text{distance of vehicle to exit point}, \]

\[ X_{\text{lookahead}} = \text{distance to exit point from which vehicle’s speed is influenced by speed in mesoscopic segment downstream}, \]

\[ X_{\text{critical}} = \text{distance to exit point from which vehicle’s speed is set to speed in mesoscopic segment downstream}, \]

\[ \alpha = \text{smoothing parameter}. \]

\( X_{\text{lookahead}} \) and \( X_{\text{critical}} \) were set to 150 m and 10 m, respectively, but should be calibrated.

Equation 1 implies that if a vehicle is further than the look-ahead distance away from the exit point, its speed remains unchanged. Equation 2 sets the speed of the vehicle to that of the downstream Mezzo segment if it is within the critical distance of the exit point. Equation 3 implies that if the vehicle is at a distance between the lookahead and critical distances from the exit point, its speed is interpolated between its current speed and the speed in the downstream Mezzo segment, depending on its distance from the exit point. The closer it is to the exit point, the closer its speed will be to the speed in the Mezzo segment. All of the foregoing are under the condition that the Mezzo segment speed is lower than the vehicle’s current speed.

If a mesoscopic link downstream of a microscopic-mesoscopic boundary becomes blocked (because of congestion), the foregoing mechanism also ensures that the VisSim vehicle at the end of the microscopic link is stopped. Conversely, when the blockage disappears, the density in Mezzo decreases, allowing the vehicle to exit as soon as a space at the beginning of the Mezzo link becomes available. Since Mezzo represents the start-up shock wave explicitly, a vehicle exiting downstream from a fully congested (Mezzo) link does not immediately result in there being space available upstream (at the end of the queue). Instead the shock-wave speed is calculated and the space becomes available with a delay [see work by Burghout (14) for details].

The speeds of the vehicles queued in VisSim (to enter the mesoscopic link) are modified only in case they are higher than the speed in Mezzo. So in the case of a start-up shock wave following the dissipation of a queue that has blocked back into the microscopic area, the speeds of the vehicles in VisSim remain untouched, allowing the (superior) microscopic mechanisms in VisSim to take care of the propagation of the start-up shock wave when it passes the microscopic-mesoscopic boundary.

Intermodel Communication and Synchronization

Whereas in earlier work (13, 14) the synchronization of the two submodels was managed by MITSimLab, the different implementation of VisSim and its COM interface requires Mezzo to control the synchronization. Each VisSim time step (typically 0.1 s) is an explicit event in the Mezzo event list, and for each VisSim event Mezzo executes exactly one VisSim time step itself while calling VisSim to do the same. At the end of this event Mezzo checks to see if there are any vehicles that exited VisSim to enter Mezzo. In addition it checks the VisSim entry parking lots for space for new vehicles, sends any new vehicles into VisSim where possible, and modifies the speeds of vehicles in VisSim that are approaching exits to Mezzo. Although it is possible to reduce the frequency of communication of traffic conditions and vehicles to decrease the communication overhead, this act may introduce some errors in vehicle loading caused by slightly delayed entry of vehicles. The current setup was deemed sufficiently fast.

EC-1 Signal Control Simulator

One of the most common signal controllers used in Sweden is the Peek Eurocontroller EC-1 (19). Part of the programming toolkit for the EC-1 is a signal controller simulator with which signal plans for the EC-1 can be tested and evaluated on a personal computer before implementation in the field. This simulator uses the exact same signal control logic as the real signal controller in the field. It has been coupled to VisSim (Figure 1) to simulate the control strategy exactly as it is implemented for the three signalized intersections. The EC-1 simulator–VisSim interface translates detector readings from VisSim to the controller inputs and signal status in the controller to red, amber, and green orders in VisSim. Multiple EC-1 simulators are controlled and coordinated via EC-1–Simulator-Control, which is also used to configure the VisSim–EC-1 simulator connections.

The detector status needs to be updated 10 times per second in order to make the EC-1 simulator work as a real signal controller in the street, and therefore VisSim needs to have a 0.1-s time step when the EC-1 simulator is connected. Another challenge with the EC-1 simulators is the absence of time synchronization with VisSim; the EC-1 simulator is set to a fixed speed, which it will keep regardless of VisSim’s simulation speed at the moment. This feature means that VisSim (and likewise the hybrid InterMezzo) needs to keep a pre-defined simulation speed when connected to the EC-1 simulator. A function that keeps a fixed simulation speed, for example, real time or two times real time, is therefore implemented in InterMezzo.

CASE STUDY

The implemented hybrid model, InterMezzo, is applied on a network for which Mezzo simulates the entire Stockholm network (6,000 links) and VisSim simulates a small portion on the southeast border of central Stockholm containing three signalized intersections, with adaptive signal control, bus lanes, and bus-priority functions in the signal control.

Recently improved adaptive signal control was implemented for these intersections, and this new signal control scheme is compared...
against the original scheme by using the hybrid model. When the capacity of a highly congested facility, such as the one in this case study, is improved drastically, it can be expected to draw additional traffic from alternative routes, which in its turn can be expected to partly offset the achieved improvement. To study such redistribution effects a large surrounding area needs to be simulated, larger than is practical to study with a microscopic model because of the extensive network coding and calibration efforts needed. The InterMezzo hybrid model, however, allows such a study of a large surrounding network together with the microscopic precision of the area of specific interest.

**Study Area**

The area of specific interest (Londonviadukten) simulated in VisSim is part of a main arterial from the southeast of Stockholm into central Stockholm (Figure 2). The capacity of this part of the arterial is limited mainly by a number of signal-controlled intersections periodically resulting in long queues, especially for Nacka in the morning peak hours. Furthermore, Londonviadukten is an important arterial for the bus lines connecting the southeast to central Stockholm, especially since this part is not connected to the Stockholm subway network. Buses make up a large portion of the traffic, around 8% in the morning peak hour. Although the buses travel mostly on separate bus lanes in the study area, they have a large impact on the other traffic because of bus-priority rules in the signal control and the fact that the bus lanes end before each intersection (to make room for right-turn pockets). The actual bus lines, stops, and timetables are coded within VisSim. Since Mezzo does not yet have these public transport facilities (bus lines, bus routes, etc.), the buses are simulated within VisSim only. As a result, the traffic flows on some roads in the Mezzo part of the model are somewhat lower than they should be.

The old signal control scheme for Intersections 1 and 2 (Figure 2) was fixed time during the morning peak hour. The new control scheme for these intersections uses vehicle-actuated Swedish signal-group control (LHOVRA). The two intersections are coordinated through fictive signal groups that can obtain green waves by delaying, shortening, or extending respective phases in the other intersection. More information may be found elsewhere (20–22). The original fixed-time control would cause long queues toward Nacka during the morning peak hour.

For Intersection 3, the signal control scheme is the same in both scenarios: LHOVRA with bus priority (PRIBUSS) (22) for buses toward Nacka. The bus priority modifies the standard functioning of the signal control to a large degree and includes green extensions as well as shortenings, phase recalls, and extra bus phases. The bus priority can cause long delays for traffic from Nacka. Although the signal control in this intersection is not coordinated with the other intersections, their operations influence each other because of their proximity.

The large surrounding network simulated by Mezzo, which spans the entire Stockholm area, was imported from the CONTRAM-8 (23) format. In addition, the demand is converted from CONTRAM time-dependent O-D matrices that were calibrated previously to fit traffic measurements (flow counts) for a number of main roads, freeways, and arterials.

The area that is now simulated by VisSim has been cut out of the mesoscopic network and replaced by the virtual links discussed earlier (green lines, Figure 2, Mezzo). The demand was calibrated in addition by selecting a number of O-D pairs that had routes through the VisSim network, to match flow counts from the city of Stockholm for most of the ingoing and outgoing links. The simulation period for both scenarios was 1 h, 7:00–8:00 a.m., of which the first 10 min was used as a warm-up period (to allow traffic to enter the network and reach the VisSim area).
Discussion on Case Study

To reduce the number of replications and iterations needed, Mezzo was run in the deterministic mode, with five iterations for each scenario, to reproduce possible redistribution effects. VisSim was run with the same random seed to avoid deviations due to stochasticity. The authors recognize that preferably Mezzo should be run in the stochastic mode and both models should be replicated with different random seeds to obtain reliable results that reflect the complete range of stochastic variation of the traffic processes modeled. Burghout (24) provides a detailed discussion on the method and number of replications needed to create reliable confidence intervals for the means of a certain set of output measures of performance. Nevertheless, the largest source of stochastic variation in VisSim is probably the vehicle arrival process, which in this case was replaced by the arrival process from Mezzo.

In addition, the Stockholm network used in Mezzo had been only partly calibrated (previously using CONTRAM and here locally to match measured flows through the area of interest). In a parallel project the network is being extensively calibrated with a large set of measurement data including speeds, flows, and travel times.

Moreover, Mezzo is being extended to incorporate public transport such as buses, bus lanes, and timetables. This extension should improve the quality of the hybrid simulation, since in this application the simulation of bus operations was restricted to the VisSim microscopic area.

Results

Table 1 shows that traffic flows from Nacka toward Stockholm in the morning peak hour increase by 23%, or 470 vehicles, with the improved signal control. The improved traffic conditions in the studied area have probably attracted traffic from alternative routes. The other differences in traffic flows are small in absolute numbers and can, at least partly, be a result of variation in the signal control start-up states (the start-up state of the signal controller simulator is not deterministic). In the direction from Slussen there is only a small increase in traffic flow; however, the improvement in traffic conditions in this direction during the morning peak hour is not that big either. The traffic flows from Folkungag decreased by 3%, or 18 vehicles, and from Tegelviksg by 11%, or 13 vehicles. One possible reason for the decreased flows from the minor approaches may be that traffic is attracted to alternative routes by the changes in traffic flows in other parts of the Mezzo network outside the studied area. In total, the traffic going through the VisSim area increased by 453 vehicles, or 12%.

The results in VisSim (Table 2) show that the average speed for vehicles in the area of interest has increased by 29%, from 22 to 28.3 km/h. At the same time the average delay per vehicle decreased by 27%, from 369 s to 268 s. The average stopped delay per vehicle also decreased, by 31%, from 161 to 111 s, and the average number of stops decreased by 27%, from 10.5 to 7.7. This finding means that with the new signal control the local conditions have improved inside the area of interest, in spite of the increased flow from the main approach into the area caused by redistribution of flows from Mezzo in reaction to the improved local conditions.

Table 3 shows the average travel time per vehicle (in seconds) for four selected O-D pairs that have traffic crossing the microscopic area (in VisSim). The first O-D pair, Nacka–Slussen (1), is relatively close to the edges of the microscopic area, meaning that there are few alternative routes. The new signal control drastically improves the travel time for these vehicles (by 52%, or 367 s). Even for the second O-D pair, Nacka–Slussen (2), which has an origin and destination located farther from the entry and exit points to the microscopic area, the travel time per vehicle is improved substantially (by 51%, or 397 s). The Nacka–Folkungag O-D pair shows even larger (relative) improvement, 60%, or 361 s. However, the O-D pair in the opposite direction, Slussen–Nacka, shows no significant improvement in travel time per vehicle.

The congestion that builds up toward Nacka (from both the Nacka–Slussen O-D pairs) with the old signal control crosses the boundary into Mezzo (as can be seen from Figure 2) and causes large queues and heavy delays. This factor accounts for the larger differences in O-D travel times than the differences observed inside VisSim alone.

In a parallel project (MATSIS), VisSim and the EC-1 simulator were run standalone on a similar microscopic network as that used in this study. The preliminary results show larger improvements with the new signal control strategy than those presented here. The average speeds increased by 39% (compared with 29% here), and the average delay per vehicle decreased by 62% (27% here). The

| Major O-D Pairs in Mezzo with Traffic Through VisSim Area for Old and New Control Strategies |
|---|---|---|---|---|
| Old (s) | New (s) | Abs. (s) | Rel. (%) |
| Nacka–Slussen (1) | 705 | 337 | −367 | −52 |
| Nacka–Slussen (2) | 778 | 381 | −397 | −51 |
| Nacka–Folkungag | 602 | 242 | −361 | −60 |
| Slussen–Nacka | 307 | 303 | −3 | −1 |

Table 1 shows the average travel time per vehicle (in seconds) for four selected O-D pairs that have traffic crossing the microscopic area (in VisSim). The first O-D pair, Nacka–Slussen (1), is relatively close to the edges of the microscopic area, meaning that there are few alternative routes. The new signal control drastically improves the travel time for these vehicles (by 52%, or 367 s). Even for the second O-D pair, Nacka–Slussen (2), which has an origin and destination located farther from the entry and exit points to the microscopic area, the travel time per vehicle is improved substantially (by 51%, or 397 s). The Nacka–Folkungag O-D pair shows even larger (relative) improvement, 60%, or 361 s. However, the O-D pair in the opposite direction, Slussen–Nacka, shows no significant improvement in travel time per vehicle.

The congestion that builds up toward Nacka (from both the Nacka–Slussen O-D pairs) with the old signal control crosses the boundary into Mezzo (as can be seen from Figure 2) and causes large queues and heavy delays. This factor accounts for the larger differences in O-D travel times than the differences observed inside VisSim alone.

In a parallel project (MATSIS), VisSim and the EC-1 simulator were run standalone on a similar microscopic network as that used in this study. The preliminary results show larger improvements with the new signal control strategy than those presented here. The average speeds increased by 39% (compared with 29% here), and the average delay per vehicle decreased by 62% (27% here). The

| Average Traffic Times Per Vehicle in Seconds for Four Major O-D Pairs in Mezzo with Traffic Through VisSim Area for Old and New Control Strategies |
|---|---|---|---|
| Old | New | Abs. | Rel. (%) |
| Nacka–Slussen (1) | 705 | 337 | −367 | −52 |
| Nacka–Slussen (2) | 778 | 381 | −397 | −51 |
| Nacka–Folkungag | 602 | 242 | −361 | −60 |
| Slussen–Nacka | 307 | 303 | −3 | −1 |
average stopped delay decreased by 49% (31% here), and the average number of stops per vehicle decreased by 53% (27% here). Some of these differences can probably be attributed to the redistribution effects that are captured by the hybrid model in this study, but not by the standalone application. However, the VisSim network used in MATSIS is extended on the approach from Nacka, meaning that the queues that now spill over in Mezzo are captured in VisSim, and therefore the effects in terms of speed, stops, and delay per vehicle are larger since they take into account more queued vehicles in the old signal control case.

CONCLUSION

The development, implementation, and application were discussed of a hybrid mesoscopic–microscopic model that applies microscopic simulations to areas of specific interest while simulating a large surrounding network in lesser detail with a mesoscopic model. The hybrid framework presented in an earlier paper was adapted to allow for implementation in a commercial model. This hybrid model was interfaced with an adaptive signal controller simulator, which reproduces the exact control behavior as is implemented in the field. With this setup, the hybrid model was applied to a case study in which the hybrid setting allowed the study of redistribution effects of the improved local conditions in the microscopic area due to the improved control scheme. The results show that even though the improved control caused the attraction of more traffic through the microscopic area (higher flows), the net effect was still positive, with fewer delays and stops and shorter travel times inside the microscopic area and shorter travel times for the main O-D pairs with traffic through the microscopic area.

ACKNOWLEDGMENTS

The authors would like to thank the city of Stockholm, especially Tobias Johansson, for funding and supporting the work described in this paper. In addition, the authors thank Fredrik Davidsson (MOVEA) for providing the CONTRAM-8 network for Stockholm, which was used as the basis for the Mezzo network.

REFERENCES


The Traffic Flow Theory and Characteristics Committee sponsored publication of this paper.
Paper II

Impacts of bus priority in coordinated traffic signals

JOHAN WAHLSTEDT

Published in
Impacts of Bus Priority in Coordinated Traffic Signals

Johan Wahlstedt *

Division of Traffic and Logistics, Royal Institute of Technology, SE-10044 Stockholm, Sweden

Abstract

Delay at signalized intersections is a considerable part of the journey time for public transport (PT) in urban areas. However, PT priority in traffic signals can reduce travel time and improve service regularity for buses at a relatively low cost to other traffic. In general it is difficult to apply analytical methods for calculating the effects of dynamic signal timing enhancements or active bus priority at traffic signals. Furthermore, no common methods are available for optimizing PT priority conditions and timings, which are usually conducted on the basis of traffic engineers’ experience and fine-tuned afterwards. This paper presents a simulation-based method for analyzing partial dynamic signal timings as well as fully adaptive signal control systems. Experiments with a microscopic traffic simulation model and a software-in-the-loop signal controller simulator are carried out to study the impacts on travel times for buses and other traffic of conditional active bus priority with the Swedish PRIBUSS method. The results show that PT priority results in shorter travel times for buses, and longer travel times for crossing traffic and traffic following the prioritized buses in one direction. This implies that there is a need for better methods to set the conditions for the bus priority in empirically based systems such as PRIBUSS.

© 2011 Published by Elsevier Ltd.

Keywords: Traffic signals; Bus priority; Transit priority; PT priority; Signal control; Simulation of traffic signals

1. Introduction

Delay at signalized intersections constitutes a large part of Public Transport (PT) journey time in urban areas. PT priority in traffic signals can reduce travel time and improve service regularity for buses at a relatively low cost to other traffic (Bang 1987; Al Mudhaffar & Bang 2006; Zlatkovic 2009). Bus priority is easier to arrange in isolated signalized intersections than in coordinated systems, where the “green waves” may be disrupted and it may be necessary to make green time compensations to other approaches after the bus has passed. However, compensations made locally in the signal controllers may have negative impacts on the performance of the coordinated system if they are not carefully timed. Previous work (Wahlstedt 2005) has shown that bus priority can have negative impacts not only to traffic that crosses the prioritized bus route, but also to other movements following the bus route.

* Corresponding author E-mail address: wahlsted@kth.se
PT signal priority can be Passive or Active. In passive signal priority signal timings can be set favoring approaches with PT, and the speed of the progression can be adapted to typical PT-speed including time at bus stops. Passive priority is often good when the PT frequency is high and the dwell-time at stops is short and predictable. Active PT signal priority triggers priority measures when there is a PT vehicle present, thus requires selective detection of PT vehicles. Active priority measures can give larger benefits to PT vehicles than passive measures since they are only applied when they are needed, and can therefore be allowed to give more negative short term impacts to other road users.

Active signal priority can be Unconditional or Conditional. Unconditional priority is mostly used for trams and will always favor the PT vehicles without consideration of negative impacts for other road users. Conditional methods may apply PT priority so that the overall intersection delay is minimized (Bang 1987). It can also be arranged by 1) setting limits for extension lengths; 2) by restricting the priority to uncongested time periods; 3) only giving priority to late/not to early busses which will also improve PT regularity. There are two main methods for conditional active signal priority: 1) Local signal timing adaptations with restrictions in a fixed timed system, and 2) Self optimizing methods that minimizes an object function (e.g. minimizes total road-user costs).

In general it is difficult to analytically calculate the effects of dynamic signal timing enhancements or active PT priority at traffic signals. Although initial signal timing may be determined based on theoretical analysis, PT priority settings are usually conducted on the basis of traffic engineering experience. Microscopic traffic simulation is often applied to assess the impacts of dynamic signal timing as well as fully adaptive signal control systems.

The aim of this paper is to study strategies and modeling of active bus priority in street networks with coordinated traffic signals. The impact of the priority is estimated based on resulting travel times for buses and other traffic. The Swedish standard PT priority method PRIBUSS (Burghout & Wahlstedt, 2007) is compared to a reference case with no priority in a case study for a network in Stockholm.

The study is based on experiments with a microscopic traffic simulation model and a software-in-the-loop signal controller simulator including PRIBUSS. The simulations are done with a calibrated and validated model of an arterial street (Fleminggatan) in downtown Stockholm based on previous field studies (Al Mudhaffar 2006). Four cases are simulated: 1) No bus priority; 2) Bus priority in both directions of travel; 3) priority for Westbound busses only; and 4) priority for Eastbound busses only.

2. PT priority in traffic signals

Terminology and philosophy behind signal control strategies varies around the world, the terminology is even different between British and American English. Therefore the term “phase” will be avoided in this paper, “stage” and “signal group” are used instead to avoid confusion.

In the U.S. signal control is based on “stages”. A stage is a portion of the cycle time for a set of movements that can, and must, have green at the same time. It is normally divided into green, yellow and all-red intervals. The signal controller changes stage in a specified order. The length of the green interval can be altered but the inter-green times are constant. A stage can be skipped if no movements in the stage have demand for green, but will be activated if there is a demand for one or more signal groups in the stage (Davol 2001).

The signal control strategies used in Sweden and many other European countries are based on “signal groups”. A signal group is a set of signals that always must show the same aspect, and control one or more movements that are given right-of-way simultaneously. The timings for each signal group are defined “periods” as minimum green, green extension time, all-red time etc. Each group advances in time and changes indication
independently according to its given periods. Primary phase-pictures are defined instead of stages, but and the controller is free to form secondary phase-pictures as a combination of the primary phase-pictures.

The advantage of the more flexible signal group technique is most evident in isolated, vehicle actuated signal control strategies, such as the Swedish LHOVRA control which is implemented as standard in almost every signal controller in Scandinavia (Vägverket 2004). The LHOVRA strategy was developed for isolated control at high speed road intersections as a modular toolbox of functions for signal group based control. Some of the functions are also used for coordinated signals (Al Mudhaffar & Cunningham 2001). Since the rigid stages are replaced by more adaptive phase-pictures it is easier to achieve an effective, active bus priority.

2.1. Swedish signal group control in coordinated systems

In Sweden the same type of signal controllers are used for coordinated, fixed timed control as well as for isolated vehicle actuated (VA) control. It is common that signals work isolated VA at night and fixed timed coordinated during day time. All signal controllers are therefore equipped with logic and detectors for making local adaptations within the coordination. Often vehicle actuated past end green (PEG) is used for some local adaptations, where the exact moment for the signal change is set in real time when the last vehicle releases the detector at the stop line. The general approach is to optimize the signal settings in detail locally, rather than optimize the overall settings in the coordinated system, mainly because of traditions and the available hardware.

2.2. Methods for conditional PT priority in coordinated signals

Conditional buss priority strategies grant priority based on certain criteria concerning the detected PT vehicle and/or the traffic signal. Priority criteria for PT vehicles can be lateness with regard to its schedule or selected routes. The number of passengers in the vehicle or time headways to the previous and the next bus are other proposed criteria for future implementation. Criteria for the signal can be that the priority measure is not applied when the intersection is congested. The impacts to other traffic movements and road users can also be limited by setting conditions for how often and how long extensions may be or given, and guaranteed minimum green time to conflicting stages.

There are two main approaches to conditional priority:

1. Prioritizing PT at the cost of other traffic with local adaptations in a normal fixed timed coordination, but with restrictions on how large the negative impacts may be.
2. Minimizing the overall delay with an optimization-based adaptive control strategy.

Adaptive control strategies work by trying to minimize a cost function for the impacts of the signal timings. The function normally includes weighted costs for delay, stops etc. for all links in the system. A bus is normally given higher weight in the optimization process since it carries more passengers, e.g. equivalent to 20-50 cars. The adaptive systems count vehicles and makes predictions of their arrival times at the next signal as a part of the optimization process. Most adaptive systems are stage based and do not attempt to optimize the signal changes exactly after every vehicle, but rather optimize the overall signal settings.

SCOOT is one of the first commercial, wide spread adaptive signal systems for coordinated networks (Arveland 2005). SCOOT has a centralized structure with a central UTC computer that optimizes the whole area communicating directly with every signal controller. The cycle time is optimized every five minutes, offsets every cycle and green split a few seconds in advance of every stage change. SCOOT only makes small changes from the previous cycle in order to make the system stable (Arveland 2005) There are two types of bus priority in SCOOT; central and local. A detected bus can be weighted up and taken into account in the optimization process as a central priority, but the communication will delay the call for priority 3-4 seconds. There is also a possibility to do a local extension in the signal controller to grant priority for busses arriving in the last seconds of green. The local priority can be inhabited by SCOOT if the degree of saturation is too high. Also the central priority can be restricted to degrees of saturation. It is also a possible to restrict priority to late busses. The stage order cannot be changed, but in the latest version of SCOOT (MC3) there is a function for stage skipping. SCOOT does not predict the dwell time at stops, so the bus detector needs to be placed downstream of the nearest bus stop (TRL 2010). The average travel time reduction when introducing SCOOT is typically around 11% (Arveland 2005).
SCATS is an Australian partly adaptive hierarchical signaling system, it could be described as a mix of an adaptive system and a plan selection system (Arveland 2005). The development started in the 1970’s and SCATS is spread outside Europe. The network is divided in sub-systems of one to ten signalized intersections around one critical intersection. Cycle time and green splits are optimized for the critical intersection based on real time vehicle counts. Timings for the other intersections in the sub-system are then selected from a library of preset time plans so that they are compatible with the optimized plan for the critical junction. The local controllers can skip stages if there is no demand, and shorten stages for side roads if there is low demand within limits set by the strategic level. Optimal offsets are calculated every cycle and are changed if three out of five subsequent calculations suggest the new offset. At a strategic level a choice can be made if two or more sub-systems should be joined with a common cycle time at the current traffic situation. PT priority is done locally and includes extensions, recall, extra stages and green split compensations after priority. SCATS is good for making green waves at the main road, and have showed main road travel time reductions of 18%. The effects on side roads and PT is unclear (Arveland 2005, Wood 1993).

SPOT/UTOPIA is distributed adaptive system developed in Italy, spread in a number of European countries and the only adaptive system implemented in Sweden. SPOT sets the signal timings by minimizing a cost function calculated locally in a SPOT unit in each intersection. Neighboring SPOT units share vehicle counts and predicted signal changes for the next minutes. UTOPIA is the higher level that tries to optimize the whole system. SPOT/UTOPIA is developed to prioritize PT, the movements though the system, including stops, are predicted in UTOPIA and the intersections are “prepared” in advance for the arrival of the PT vehicle. At the local level in SPOT the PT vehicle will be given a value of 25 cars in the optimization process.

In Torino SPOT has resulted in reduced travel times of 2-7% for PT and 10% for cars. The average travel time reduction when introducing SPOT has been 10% for PT and 12% for cars (Arveland 2005). Field tests of SPOT in Stockholm have showed 10% travel time reduction for PT and 7% for cars (Al Mudhaffar 2006). The possibility to prioritize PT in SPOT/UTOPIA is one of the reasons for the interest for the system in the Scandinavian countries.

Another approach to conditional priority is to make local priority in a fixed timed coordination, and try to restrict the impacts on other traffic. This approach is wide spread in German-speaking Europe, where the signals often are stage based and with almost unconditional PT priority (Wood 1993).

3. PT priority with PRIBUSS

PRIBUSS (Swedish acronym for Prioritizing of Busses in Coordinated Signal systems) was developed in Stockholm in the early 1990s. PRIBUSS is now included as standard in most signal controllers on the Swedish market, and is the common method for PT priority in Sweden. Another method is used for the more unconditional tram priority in Gothenburg (Wahlstedt 2005).

PRIBUSS could be described as a toolbox of PT priority procedures for the traffic engineer to choose from when designing the traffic signal. It is developed for conditional PT priority on top of the normal primary fixed-time control. The engineer decides which procedures, conditions and limitations that will be applied by parameter programming. It can be used for isolated as well as coordinated control (Bjork & Dahlgren 1991). The priority and compensations are done locally in each signal controller and neighboring intersections are not considered. There is no method for prioritizing between buses with conflicting calls, the bus that first calls for priority will be served. It is up to the engineer to limit the functions so that the negative impacts on the coordination are acceptable.

There are six PRIBUSS procedures:
- Extension: extends ongoing green
- Re-taken start: return to green if conflicting groups have not yet have become green (eg. under inter-green and red-amber)
- Early green: cuts the ongoing stage, with optional limitations
- Extra stage: inserts an extra stage in or between the normal stages
- Double early green: cuts to subsequent stages
- Double extra stage: cuts to subsequent stages and inserts extra stage
The PRIBUSS programming in the signal controller is built up of a number of functions that start a selected procedure given a set of conditions. If all conditions are met, the function is started and all other functions are blocked until the function is ended. A bus arriving during green will use an extension function to block other PRIBUSS functions to for conflicting buses. This is the only way conflicting calls for priority are handled. There could be more than one function with different conditions for the same type of procedure. Each function is conditioned by a “window” in the cycle, the status of some signal groups, and to one or more time plan in order to make sure that the appropriate PRIBUSS function is chosen. There could also be a conditioned, guaranteed green time for conflicting signal groups. There is also a set of similar conditions deciding whether to make compensations or not after the function is ended.

![Flowchart for PRIBUSS (Björk & Dahlgren 1991)](image)

There is a bus-counter dedicated to each approach, the busses are counted in when they call for priority and counted out after passing the stop line. In downtown Stockholm a radio based AVL system is used for direct communication between the bus and signal controller. A bus more than two minutes ahead of its schedule will not be approved for priority. An approved bus will be counted in and start a PRIBUSS function if the conditions are met. A time-out counter is started with the PRIBUSS function and it can be reset by a second bus, or by a bus in the opposing approach if it is allowed to use the same function.

When the function is ended, the “c-pulses” for signal changes that have been blocked and “saved” by the PRIBUSS function will be conducted and compensations to conflicting signal groups are made by changing the timings in the next cycle if the specified conditions for compensation are met. The function can be programmed to block itself and/or another function for a certain time after the function is ended to reduce the impact to other traffic.

4. Case study

The case study aim is to study the travel time impacts of PT priority with PRIBUSS. Simulation experiments are carried out using the VISSIM microscopic traffic simulation model connected to EC1 signal controller simulators with exactly the same programming as in the field. The studied time period is the morning peak hour 08.00 - 09.00.

The area chosen for this study includes a section of the urban arterial Fleminggatan and the crossing street Scheelgatan at Kungsholmen in downtown Stockholm with a system of six coordinated signals. The streets in the area are relatively congested during peak hours and there is heavy bus traffic in general purpose lanes. Fleminggatan is a typical Stockholm main street, 18m wide, accommodating approximately 21 000 vehicles daily and trunk bus route 1. It has one lane for general purpose in each direction, on street parking and short turning pockets at some intersections. The intersection Fleminggatan - Scheelgatan is the critical intersection in the signal coordination, the other side roads to Fleminggatan are only of local importance.

Trunk bus line 1 on Fleminggatan has a five minute headway with articulated busses and signal priority. Bus line 56 with 15 minute headways has the same route through the studied area but no signal priority. Line 40 with 10
minute headways operates on Scheelegatan without signal priority. There are no dedicated bus lanes in the studied area.

The coordinated signal system includes four intersections and one pedestrian crossing at Fleminggatan and one intersection at Scheelegatan that is only coordinated during day time. The signals have a cycle time of 82s in peak hours and are mainly fixed-time with some local traffic adjustments using PEG of c:a 5 seconds in most signal groups.

The area has been studied several times in recent years (e.g., by Al-Mudhafar 2006) and there is much traffic measurement data available.

5. Simulation model

Analytical methods can be used for analysis of fixed-time traffic signal impacts as well as some simple vehicle actuated control, but are normally not applicable for estimation of the impacts of more complex vehicle actuated signals, adaptive signal control and PT priority. It is, however, possible to study such control methods using simulation. A commonly used method in recent years for more detailed studies of signal control strategies is to link a commercially available traffic simulation model (e.g., VISSIM, Aimsun etc.) to a signal controller simulator for the studied type of signal controller (Zlatkovic et al. 2009).

The method used in this case study was “software-in-loop”, a microscopic traffic simulation model (VISSIM) connected to a signal controller simulator (EC1-simulator). The vehicular movements and interactions were modeled in the traffic simulator which also modeled the signal controller detector loops. The detector status was sent to a signal controller simulator which processed the data and sent signal group statuses back to the traffic simulator. The returned signal group status controlled the indications of the traffic signals in the traffic simulation model.
Figure 4 Simulation model setup

5.1. VISSIM model

The traffic simulation software used in this study was VISSIM. The 1000m x 500m VISSIM network contains six signalized intersections, ten incoming and twelve outgoing links for vehicular traffic, pedestrians crossing at crosswalks, one bus line with signal priority and two lines without. The traffic signals are equipped with detector loops for receiving data to the EC1-simulator. The percentages of heavy vehicles except busses are 2-3% in the different incoming links according to counts, but VISSIM default vehicle lengths are used for cars which can be somewhat too short for Stockholm conditions.

This simulation network was originally built and calibrated in 2004 with an older version of VISSIM. It has been modified and used for different purposes in newer versions of VISSIM since then, without being re-calibrated. The network can therefore not be considered as fully calibrated to current conditions, but as a realistic hypothetical network. The traffic flows, turning percentages etc. used in the model where collected in May 2003. As far as possible, data from 08.00-09.00 in the morning of Wednesday 2003-05-14 were used in the model.

5.2. EC1 Signal controller simulator

One of the most common signal controllers used in Sweden is the Peek EC1. Part of the programming toolkit for the EC1 is a signal controller simulator with which signal plans/programming for the EC1 signal controller can be tested on a PC before implementation in the field. This simulator has been coupled to VISSIM to simulate the control, including the bus priority, exactly as it is implemented on the street. The EC1-simulator / VISSIM interface translates detector readings from VISSIM to the controller inputs and signal status in the controller to red, amber and green orders in VISSIM. The detector status needs to be updated 10 times / sec in order to make the EC1-simulator work as a real signal controller in the street, and therefore VISSIM needs to have 0.1s time step when the EC1-simulator is connected (Burghout & Wahlstedt 2007). The signal programming used in this work is that which was used on the street 2003, at the time of data collection for the simulation model, including some programming errors later corrected in the controllers at the street.

6. Simulation results

The results for each scenario presented below are an average of the results from ten simulation runs with different random seeds. The same ten random seeds have been used for all scenarios. Travel time has been used as main indicator of the impacts of the PT priority in this study, the delays caused by the traffic signals can be assumed to vary proportionally since everything else is kept constant. Travel time was measured from passing the stop line at one intersection to passing the stop line at the next intersection. When travel times for one segment are mentioned below, they are they denoted with the latter of the two streets. The number of stops has not been studied for each approach, but only for the whole area.

6.1. Results with and without PT priority
The signal system was simulated with and without “PRIBUSS” PT priority. The traffic signals in the simulation model were in both cases controlled by EC1 signal controller simulators connected to VISSIM as described above. The bus detectors were disconnected in the case without PRIBUSS so that no priority calls could reach the signal controller.

The travel time for busses decreased with bus priority while travel time for traffic that crossed the bus route increased as could be expected, see Figure 5. More interesting is that travel times for eastbound cars, i.e. the direction with bus priority, also increased. Westbound cars experienced a small decrease in travel time. One possible reason for this could be that long green time extensions caused by PT priority shifted the “green wave” so the coordination for opposing traffic was disrupted.

The number of stops for all vehicles (including busses) increased with 4% when the PT priority was used, but it decreased with 23% for buses.

6.2. PT priority only in one direction

Two more scenarios, PT signal priority only for Eastbound buses, and PT signal priority only for Westbound buses, were simulated in order to analyze the impacts of the PT priority on traffic that opposes the prioritized bus direction, see results in Table 1.

<table>
<thead>
<tr>
<th>PT priority</th>
<th>Diff.</th>
<th>Travel times: Car</th>
<th>Travel times: Bus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Westbound</td>
<td>Eastbound</td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>201.43</td>
<td>140.43</td>
</tr>
<tr>
<td>Both directions</td>
<td></td>
<td>197.28</td>
<td>149.39</td>
</tr>
<tr>
<td>Abs.</td>
<td></td>
<td>-4.15</td>
<td>8.96</td>
</tr>
<tr>
<td>Rel.</td>
<td></td>
<td>-2.06%</td>
<td>6.38%</td>
</tr>
<tr>
<td>Eastbound</td>
<td></td>
<td>201.65</td>
<td>139.37</td>
</tr>
<tr>
<td>Abs.</td>
<td></td>
<td>0.22</td>
<td>-1.06</td>
</tr>
<tr>
<td>Rel.</td>
<td></td>
<td>0.11%</td>
<td>-0.75%</td>
</tr>
<tr>
<td>Westbound</td>
<td></td>
<td>181.94</td>
<td>147.6</td>
</tr>
<tr>
<td>Abs.</td>
<td></td>
<td>-19.49</td>
<td>7.17</td>
</tr>
<tr>
<td>Rel.</td>
<td></td>
<td>-9.68%</td>
<td>5.11%</td>
</tr>
</tbody>
</table>

Figure 5 Travel times with and without PT priority

Table 1 Travel times with PT priority in no, both or eater direction
The travel time for buses changes as follows:
• shorter travel times in scenarios with priority only in “own” direction than with priority in both directions
• slightly longer travel time with priority only in “other” direction than without priority.

Other traffic traveling along the whole length of the coordinated system showed the same trend, bus priority in “own” direction only was slightly better and in “opposite” direction slightly worse than no priority. The travel times between intersections differed much more, as can be seen in Figure below.

The results for each segment can in some cases be explained by the implemented priority functions. There are bus stops upstream from the intersection Pipersgatan and CG Lindstedts gata westbound as well as upstream from Wargentinsgatan and Scheelegatan eastbound. The dwell times at the bus stops are included in the travel times in Figure 6.

![Figure 6: Travel times for buses and general purpose traffic at main street (Fleminggatan) between each intersection](image-url)
7. Discussion and conclusions

This paper presents a study using experiments with a microscopic traffic simulation model and a software-in-the-loop signal controller simulator to study the impacts on travel times for buses and other traffic of conditional active bus priority with the PRIBUSS method.

The results of the simulations show that active bus priority resulted in considerable benefits to the bus passengers, 9% and 2% travel time reduction in respectively direction. The travel time reduction in both directions together is 6%. This is in line with data collected from field in previous work (Al Mudhaffar 2006) showing travel time reductions of 7% with PRIBUSS and 20% with SPOT for both directions together, compared with no bus priority.

This study also showed longer travel times for other traffic, up to 13% on the cross street and 6% on the main street. One could expect that traffic that traffic crossing the prioritized bus route would get increased travel times since the PT priority re-locates green time from the crossroad to the main street with the bus route. The PT priority also disturbs the coordination between the main intersection and the side streets (Scheelligatan) intersection with Kungsholmsgatan by changing the offsets.

On the other hand one would also expect that cars on the same street as the prioritized bus route should benefit from the re-located green time, and not get an increased travel times as the Eastbound cars in this case. The travel times between each intersection show both increased and decreased travel times for vehicles in the same direction as the prioritized bus. One reason for this could be that the green wave progression partly will be adjusted to bus speed, including dwell time at stops, and cars benefiting from the bus priority therefore have to wait for the buss (or the next normal green period) at the next signal. This is the case at C.G Lindstedts gata in both directions as well as Scheelligatan Eastbound. Another reason for this could be that compensations made after the bus priority destroys the green wave in the opposite direction as indicated by the increased travel times with bus priority only in opposite direction.

The impacts of bus priority with PRIBUSS seem to be very much dependant on the individual settings in the intersections. Those are set by experienced traffic engineers but since there is no good way to calculate optimal settings, and there is no built in optimization in the method, the result is unpredictable. One possible method could be to use simulations to systematically test different settings in the priority functions before implementing them on the street.

The used simulation method seems to have a good potential for evaluation and improvement of bus priority functions. Further research will try to describe the impacts of different priority functions and finding criteria for bus priority that gives shorter travel times for public transport and/or less negative impacts on other traffic.

8. References

Angus P. Davol (2001), Modeling of Traffic Signal Control and Transit Signal Priority Strategies in a Microscopic Simulation Laboratory, MIT
Arveland, G (2005), Adaptiv kjøretøjstyring med spot i gatenett med skjev trafikkfordeling, Doktoravhandlinger ved NTNU 2005:115
Bang, K-L (1987), Bussprioritering i trafiksignaler, TFK rapport 1987:3,
Björk & Dahlgren (1991), PRIBUSS En metodbeskrivning, Gatukontoret Stockholms Stad
Burghout, W and Wahlstedt, J (2007), Hybrid traffic simulation with adaptive signal control, Proceedings of the 86th Transportation Research Board Annual Meeting, TRR 1999
TRL (2010), Advice Leaflet 4: Bus Priority in SCOOT, TRL
Wahlstedt, J (2005), Simulerings av bustprioritering i traffiksignaler, Master theses, KTH
Wood (1993), Urban traffic control, systems review, TRL
Vägverket (2004), Utdrag ur VGU Trafiksignaler, Publikation 2004:80
Paper III

Evaluation of Different Bus Priority Functions in Coordinated Traffic Signals

JOHAN WAHLSTEDT

Published in
Proceedings of The conference of the TRB 92th Annual Meeting, January 2013, Washington DC, USA
Evaluation of Different Bus Priority Functions in Coordinated Traffic Signals

Johan Wahlstedt
Division of Traffic and Logistics
KTH, Royal Institute of Technology
SE-100 44 STOCKHOLM, Sweden
Phone: +46 70 6275458
Email: wahlstedt@kth.se

Submitted to the 92st Annual Meeting of the Transportation Research Board,
January 2013, Washington, D.C.

5495 words + 8 figures x 250 = 7495 words
ABSTRACT

Delay at signalized intersections is a considerable part of the journey time for public transport (PT) in urban areas, but the delay can be reduced with PT priority in the signals at a relatively low cost to other traffic. The impact on other traffic is dependent on which priority functions are used, but the effects are hard to calculate in a coordinated system and do not always act intuitively. In this paper, experiments with a microscopic traffic simulation model are conducted with a software-in-the-loop signal controller simulator to study the costs and benefits, primarily in terms of travel times for buses and other traffic, of different priority functions in the Swedish PRIBUSS method for conditional active bus priority. Based on the analysis of the different functions, new conditions for priority are proposed and tested with the microscopic simulation model.
INTRODUCTION

In the last decade increasing congestion problems on urban streets combined with environmental concerns and reduced funds for adding road capacity, have evoked an interest in promoting Public Transport (PT) all over Western Europe. There have also been shifts in focus from building metros to high quality surface transport, e.g. Bus Rapid Transit (BRT) and/or tram lines, which share at least the intersections with other road traffic. One of the most important factors for mode share is the travel time ratio between car and PT.

Delay at signalized intersections constitutes a large part of PT journey time in urban areas. PT priority in traffic signals can reduce delay and improve service regularity for PT at a relatively low cost to other traffic (1-3). PT priority is easier to arrange in isolated intersections than in coordinated signal systems, where the “green waves” may be disrupted. It also may be necessary to make green time compensations to other approaches after the PT vehicle has passed without changing the cycle time. However, compensations made locally in the signal controllers may have negative impacts on the performance of the coordinated system if they are not carefully timed.

Background

Traffic signals are traditionally designed to minimize the delay per vehicle. However vehicles carry a very different number of passengers; a bus typically has 10 – 20 times the number of passengers compared to an average car and a tram even more (1). In order to minimize the delay per person PT vehicles need to be treated differently from cars in the traffic signal, i.e. the PT vehicles need to be prioritized in order to minimize person delay.

PT signal priority can be passive or active. In passive signal priority, signal timings can be set favoring approaches with PT, and the speed of the progression can be adapted to typical PT-speed including time at bus stops. Passive priority is often good when the PT frequency is high and the dwell-time at stops is short and predictable. Active PT signal priority triggers priority measures when there is a PT vehicle present, thus requiring selective detection of PT vehicles. Active priority measures can give larger benefits to PT vehicles than passive measures since they are only applied when they are needed, and can therefore be allowed to give more negative short-term impacts to other road users. Adaptive signal control strategies minimize the overall intersection or area delay by minimizing an object function (e.g. minimize total road-user costs) where the PT vehicles are given a higher weight. In conventional signal control strategies, the signal timings are adjusted to the PT vehicle arrivals by local signal timing adaptations done within a fixed timed coordinated system.

Active signal priority can be unconditional or conditional. Unconditional priority will always favor the PT vehicles without considering negative impacts for other road users, while conditional priority in one way or another tries to restrict the impacts on other traffic. In conventional signal control strategies, this is done by setting conditions/restrictions on the priority measures including; only using certain priority functions (green extension, red truncation, etc.) in the specific intersection; setting limits for extension lengths; setting minimum green times on side streets; setting a minimum time between two priority measures; and restricting the priority to uncongested time periods. The conditions can also be based on the buses: only give priority to certain lines; not giving priority to early buses/only giving priority to late buses, which will also improve PT regularity. In this paper the term conditional priority is used both for conditions considering the state of the PT vehicles and restricting the impact on other traffic.

Many case studies of PT priority in coordinated systems can be found in the literature, both field studies (2, 6-7) and simulation experiments (3-5), (8-10). The effects of different priority measures (green extension, red truncation, etc.) have been studied in an isolated intersection as well; and also the total effects of active PT priority in a coordinated system. Green extensions are reported to be less disturbing to other traffic than truncations or phase rotations (6). However, there is few studies of the impacts of different priority measures within a coordinated system. The impacts of PT priority within a coordinated signal system will be different from an isolated signal since the arrivals are in platoons rather than random.

Analytical models for general evaluation of PT priority in coordinated signals exist (11), but it is hard to calculate the impacts of active PT signal priority in coordinated systems in detail since the dynamic changes in signal timings in the different signal controllers interact with each other. The platoon progression will be different in each cycle when the frequency of the prioritized PT lines is high and/or the coordinated system is large. The impacts of a priority measure can also occur at a downstream intersection rather than at the one where the priority measure is done. The impacts are however possible to study with simulations, and microscopic traffic simulation connected to a signal controller simulator is often applied to evaluate dynamic changes in signal timings such as PT priority.
Conditional priority based on the on-time status of buses can help to improve the reliability of bus service. By giving priority only to buses that are on time or late (or only to late buses), late buses can be “pushed” on time and early buses can be held back. The negative impacts on other traffic will also be reduced compared to unconditional priority (6-8).

**Objectives**

The aim of this paper is to study the cost and benefit, in terms of travel times, of different PT signal priority measures in a coordinated system. Furthermore, the aim is to propose and test other conditions for conditional bus priority than the one presently used, based on the cost and benefits of the different priority measures. This is done by simulation experiments with a microscopic simulation model of Fleminggatan (Flemming Street) in Stockholm with Software-In-the-Loop signal controller simulators including the Swedish standard priority method “PRIBUSS” (10).

**PT SIGNAL PRIORITY THEORY**

**Swedish signal control with “PRIBUSS” PT priority method**

Swedish signal control strategies are signal group based. Minimum green times, extensions, red times, etc. are measured individually for each signal group; which is an advantage when signal timings should be changed dynamically as with PT priority. The ring-barrier strategy is not used. Coordinated signals are usually equipped with detectors and operate isolated-vehicle actuated at night and coordinated fixed-time with local adoption using a few seconds vehicle-actuated past-end-green (PEG) during day time.

The priority method PRIBUSS (Swedish acronym for prioritizing of buses in coordinated signal systems) was developed in Stockholm in the early 1990s and is now included as standard in most signal controllers sold in Sweden (4). It is developed for conditional PT priority on top of the normal primary fixed-time control, but can also be used for isolated control (12). The priority and compensations are done locally in each signal controller and neighboring intersections are not considered. There is no method for prioritizing between buses with conflicting calls; the vehicle that calls for priority first will be served (5).

PRIBUSS consists of six priority procedures:
- Extension (BF): extends ongoing green
- Re-taken start (ÅTS): return to green if conflicting groups have not yet become green (i.e. under inter-green and red-amber)
- Red truncation (AK): cuts the ongoing stage, with optional limitations
- Extra stage (EF): inserts an extra stage in or between the normal stages
- Double red truncation (DAK): cuts to subsequent stages
- Double extra stage (DEF): cuts to subsequent stages and inserts extra stage

In the signal controller, one or more PRIBUSS functions are configured; each function has a number of conditions and if all conditions are fulfilled, the corresponding PRIBUSS procedure is started. The conditions are usually: bus counted in, signal group green (one or more), signal group red (one or more) and cycle time slot.

The conditions must be set so only one function starts at a time.

![Flowchart for PRIBUSS functions (12)](Figure 1: Flowchart for PRIBUSS functions (12))
Johan Wahlstedt

**PT priority in Stockholm**

Conditional PT signal priority is implemented in central Stockholm since the early 1990s; the conditions are mainly set to restrict the disturbance of other traffic. Currently, only buses less than two minutes ahead of schedule on the four trunk bus routes ("blue bus" lines) and one tram line receive signal priority, while early vehicles and local bus lines ("red bus" lines) do not. There are a few holding points along the lines, and during daytime few buses run two minutes or more ahead of schedule.

The buses and trams are equipped with an AVL system that keeps track of location and schedule, and a short-distance data radio device for communicating directly with the signal controllers. The bus calls for priority only when running: in service, on a trunk line, and has not arrived at the last stop more than two minutes early. If the calling point is located at a stop, the bus will not call until the door brake is released when the bus starts driving.

There are many conditions/limitations set locally in the signal controllers to restrict the impacts on other traffic. Those are set beforehand by the traffic engineer based on experience, and are not dependent on actual traffic conditions or the state of the PT vehicle. The cycle time slots for the start of priority functions are sometimes limited. Extensions are prevented to be too long by setting a time-out and a last cycle second for the extension, i.e. the extension will be terminated at that point even if the bus has not passed the stop line. Minimum green times for unprorized approaches are often determined by pedestrian crossings, but sometimes a "guarantee time" is given to vehicle signal groups to avoid oversaturation. In Stockholm, another condition is that all signal groups with green request should be served in each cycle. The green time can be shortened or shifted earlier or later but not skipped. This is achieved by setting the cycle time slots and giving compensations after the priority measure.

**Priority Functions in the Studied Intersections**

The five coordinated intersections along Fleminggatan are principally two-phase controlled, but the main intersection Fleminggatan/Scheelegatan (3334) has lagging green in the eastbound direction to assist the large left-turning stream onto Scheelegatan northbound.

Each intersection has different priority functions implemented due to local conditions (see Table 1 below). Green extension (BF) is implemented on all approaches while green extension during past-end-green (BF2) only exists in approaches with PEG, but not all of them. Re-taken start (ÅTS) is only used at nearside bus stops as well as double red truncation (DAK), but DAK is not used at the saturated intersection 3334. Since the intersections are two-phase controlled, extra phase (EF) is only meaningful for westbound buses in 3334 between lagging green in the opposite direction and the side street. Red truncation (AK) is not possible in 3307 since the red time is only determined by the minimum green time and clearance time for the pedestrian crossing. In most intersections, red truncation is prohibited a time period after green extension.

<table>
<thead>
<tr>
<th></th>
<th>Polhemsgatan</th>
<th>Wargentinsgatan</th>
<th>CG Lindstedtg.</th>
<th>Scheelegatan</th>
<th>Pipersgatan</th>
</tr>
</thead>
<tbody>
<tr>
<td>3331</td>
<td>EB X</td>
<td>WB X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3307</td>
<td>EB X</td>
<td>WB X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3338</td>
<td>EB X</td>
<td>WB X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3334</td>
<td>EB X</td>
<td>WB X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3321</td>
<td>EB X</td>
<td>WB X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**BF** (Green extension) | X | X | X | X | X | X | X | X | X | X | X | X
**BF2** (Green extension during past-end-green) | X | X | X | X | X | X | X | X | X | X | X | X
**ÅTS** (Re-taken start) | X | X | X | X | X | X | X | X | X | X | X | X
**EF** (Extra phase) | X | X | X | X | X | X | X | X | X | X | X | X
**AK** (Red truncation) | X | X | X | X | X | X | X | X | X | X | X | X
**DAK** (Double red truncation) | X | X | X | X | X | X | X | X | X | X | X | X

Table 1: Priority functions implemented in the intersections at Fleminggatan
STUDY AREA

A part of the urban arterial Fleminggatan and the crossing street Scheele gatan in central Stockholm is used as a case to study the impacts of different PRIBUSS priority measures with simulation experiments. Fleminggatan is a typical main street in Stockholm, 18m wide, accommodating approximately 21 000 vehicles daily and heavy bus traffic in mixed traffic lanes. It has one lane in each direction, on-street parking and short turning pockets at some intersections. The intersection Fleminggatan - Scheele gatan is the critical intersection in the coordinated signal system; the other side streets are only of local importance. Five intersections and one signalized pedestrian crossing are coordinated with a 82s cycle time during peak hours. The signals along Fleminggatan have “PRIBUSS” bus priority implemented. Fleminggatan is relatively congested during peak hours.

Blue bus line 1 on Fleminggatan has five minute headway with articulated buses and signal priority. Bus line 56 with 15-minute headway has the same route through the studied area but no signal priority. The area has been studied several times in recent years (4, 5, 13) and there is much traffic measurement data available.

Figure 1: The case study area marked with the signalised intersections and blue bus line 1

Simulation Model

The simulation model used for the experiments was previously developed and calibrated (5). A “software-in-the-loop” setup is used with the microscopic traffic simulation model VISSIM connected to six EC1 signal controller simulators.

Traffic flows, turning percentages etc. were collected in May 2003 and the model has been calibrated with this data (4). Data on boarding and alighting passengers per stop per hour was taken from SLs automatic passenger counting system “ATR”. Since only ten percent of the buses are equipped with ATR, the data is only usable at an aggregated level. The number of passengers in May 2003 and April 2011 were similar according to ATR data.

Data on individual buses entering the simulated area was taken from SLs AVL system “BussPC” from 08.00-09.30 on the morning of Thursday April 7, 2011, departure times from the upstream stops have been used as entering times in the VISSIM network. The buses have been classified into nine classes according to their on-time status (compared to their time table in the AVL system at that stop). The 1.5 hour long morning peak hours have been repeated twice to get three hours output data from each simulation run.

The number of passengers on each bus entering the VISSIM model has been calculated according to the headway from the previous bus, adjusted for bunched buses. The number of passengers varies between five and 62 per bus due to the headway variation between 1.09 to 12.05 minutes, with an average occupancy of 26 westbound and 38 eastbound when entering the network and a nominal headway of five minutes.
The dwell time at bus stops in VISSIM has been calculated as $(11.3 + 2 \times \text{boarding passengers} + 0.5 \times \text{alighting passengers})$ according to findings in a previous study (4). In this way, the dwell time at stops in the simulation model will be dependent on the headway from the previous bus on the same line. In general, late buses will have longer headways and thereby longer dwell times, but it does happen that late buses have short headways due to the previous bus being even later.

The six EC1 signal controller simulators use in principle the same programming as the signal controllers on street 2011, some programming errors have been corrected (on street) since previous studies (4, 5). The “PRIBUSS” bus priority has been reprogrammed in such a way that the different priority functions can be triggered separately via separate inputs for each function.

**IMPACTS OF DIFFERENT PRIORITY MEASURES**

Twenty-nine scenarios divided in five groups have been simulated and analyzed with respect to travel times:

1. Different priority functions active in the main intersection, full priority in the other intersections;
2. Different priority functions active in the main intersection, full priority in the other intersections, 10% reduced traffic;
3. Different priority functions active in all intersections;
4. Different priority functions active in all intersections, 10% reduced traffic;
5. New conditions for different priority functions based on the buses on-time status.

The signal controllers have been re-programmed so the different PRIBUSS functions that are in each signal controller can be called separately via separate I/O ports. Due to technical limitations in the signal controller, the priority functions have been grouped in five groups to be called by the buses:

- Short extension; i.e. max time extension (BF);
- Long extensions; i.e. extension during PEG and re-taken start (BF2+ÅTS);
- Extra phase (EF);
- Red truncation (AK);
- Double red truncation (DAK).

**Study of Different Priority Functions in One Intersection in a Coordinated System**

In order to study the impacts of different priority measures within a coordinated system in a more structured way different priority functions are allowed one by one in the main intersection (3334). The intersection is located in the middle of the coordinated system with a coordinated signal on three sides. In the other intersections, all priority functions are allowed so that buses can arrive outside the normal green wave, i.e. “No Priority” means full priority in all intersections but the studied one (3334).

Five scenarios with different priority measures allowed in signal controller 3334 were formed:

1. No priority;
2. BF; short extension;
3. BF+BF2+ÅTS+EF; short and long extension;
4. AK; red truncation;
5. All functions.

Difference in travel time, compared to the “No Priority” scenario, is used as main evaluation parameter. Travel times have been grouped in “blue bus” which is trunk route buses receiving priority when they are not too early, and “GP Traffic” which is cars and delivery trucks. The travel times are calculated from the stop lines (also at the side streets) in the upstream intersections to the stop lines in the downstream intersections to capture the effect that some of the travel time gained or lost by a priority measure in one intersection can be equalized out in the downstream signal. In the westbound direction, the queue is often passing the upstream intersection and the travel time is here counted further back at the end of the queue.
Travel Time per Vehicle

As seen in Figure 2, the results show substantial travel time savings for westbound buses as well as for westbound GP traffic. In the eastbound direction, the travel time savings is more moderate as is the increase in travel time on the side street. This is due to the long queues in the westbound direction and the effect of buses “pushing” the queue ahead when receiving priority. This is the true effect in this specific intersection, but not in intersections in general. In order to get more generally applicable results, a sensitivity analysis with 10% less traffic has been performed. A 10% reduction in traffic flows is of the same magnitude as the impact at the location of the congestion charges introduced in 2007. With the 10% traffic flow reduction, there are still queues in the westbound direction but the oversaturation without priority is relieved.

Figure 2: Travel times around intersection 3334 compared with the “No Priority” scenario

Not allowing one priority function to start implicitly means that another priority function has a larger chance of being started. If green extension is not allowed, the bus will meet red and receive red truncation instead. The red truncation function, which is the most disrupting one, will therefore be started more often in the red truncation only (AK) scenario compared to the “All Functions” scenario which explains the larger negative impact on the NB/SB traffic.

As can be seen in Figure 3, the results with reduced traffic flows are showing the same pattern, but with another magnitude, as with higher flows. The bus travel time savings are relatively large with a small increase in travel time for both short (BF) and long extensions (BF+BF2+ÅTS+EF). The effects for eastbound buses are limited by the short distance from the nearside bus stop to the signal in this direction. Red truncation (AK) shows a reduction in bus travel time, but at a cost to the side street traffic. The short distance from the bus stop in the eastbound direction is of less importance for truncation compared to extension. The combination of extension and truncation gives the largest travel time reductions for buses and GP traffic following the buses, at a cost for side street traffic between the one of only extension and only truncation.
Travel Time per Person

From a socioeconomic point of view, the travel time savings per person is more important than travel time per vehicle. Westbound buses have in average 26 passengers and eastbound buses 38 passengers in the morning peak, according to counts. In Stockholm the average number of passengers per car is 1.2 during peak hours.

With those numbers, the travel time savings by PT priority in the studied intersection in person hours per hour by scenario are; BF 7.0pers h/h; BF+BF2+ÅTA+EF 10.6 pers h/h; AK 5.9 pers h/h and All Functions 12.9 pers h/h. A combination of all priority functions gives the largest travel time savings. The travel time savings in the reduced volume scenario show the same pattern.

Study of Different Priority Functions in all Intersections in a Coordinated Signal System

In order to verify the conclusions from the simulation experiments in the main intersection, the different priority functions have been tested one by one in all intersections in the coordinated system. Only the functions existing on street in each signal controller were tested.

Eight scenarios with different priority measures allowed in the signal controllers were formed:
1. No priority;
2. BF; short extension;
3. BF+BF2+ÅTS; short and long extension;
4. BF+BF2+ÅTS+EF; short and long extension and extra phase;
5. AK; red truncation;
6. DAK; double red truncation (truncate ongoing green and insert side street phase at minimum length);
7. All functions but DAK;
8. All functions.

In Figure 4 and Figure 5 below WB/EB “GP traffic” are cars and trucks on Fleminggatan including inturning traffic from side streets; NB/SB is traffic on Scheelsgatan and “Side Streets” traffic from the minor side streets.

There is a small unavoidable double counting of traffic from the side streets onto Fleminggatan. A sensitivity test scenario with 10% less traffic was simulated in this case as well.

The results generally follow the same pattern as at the main intersection. A large part of the travel time savings in the whole system also originates from the main intersection, but the results around each intersection follows the same pattern generally.
Johan Wahlstedt

Travel time differences between two adjacent intersections also show effects of priority measures in upstream intersections, which can be understood only by studying the time-space diagram. Generally the bus travel time savings are large when entering the coordination and smaller further down when the initial queue is cleared. In some cases, a part of the time gained in one signal is lost in a downstream intersection.

The travel time differences for side streets is dominated by effects on Pipersgatan, the easternmost intersection, which often becomes blocked by the queue from the main intersection in the “No Priority” scenario and benefits from buses clearing the queue when receiving priority. In the full-traffic scenario, those results can therefore not be generalized.

Figure 4: Travel times in the coordinated system compared with the “No Priority” scenario.}

As in the case at the main intersection, the bus travel time reductions with extensions (BF, BF+BF2+ÅTS) are relatively large at a moderate cost for side street traffic; this effect is larger for higher traffic flows. Also in this case WB/EB GP traffic following the prioritized bus route has somewhat increased travel times, especially from long extensions. One can suppose that this is due to moves of the side street green period that disturbs the green wave in the cycle following the priority measure, as indicated in previous work (5). The effect of the extra phase (EF) seems to be important only with high flows. Red truncations (AK) shorten bus travel times, but at a higher cost to side street traffic and are beneficial for traffic following the buses.

An somewhat unexpected result is that double red truncation (DAK) when being the only allowed function increases the travel time for westbound buses. DAK is only used for westbound buses at two nearside stops close to intersections 3321 and 3338 when the buss arrives late at green. In combination with other functions, the outcome of the function is positive. However, it needs to be combined with other functions in order to get the intended results.

Travel time gained in one intersection often seems to be lost in a downstream signal because the priority measures are only done locally in the signal controllers without knowledge of the actual status in the downstream intersection. This effect is strengthened in the studied case by late calls for priority since the intersections are closely spaced.
New Conditions for Conditional PT Priority

The conclusion from the experiments allowing different priority functions in the intersections along Fleminggatan is that all tested priority functions reduce bus travel times, but that the negative impacts on other traffic vary with the respective priority function; short extensions having the smallest impact on other traffic and double red truncation the largest.

A new set of conditions based on these results is proposed; all buses but the ones more than one minute early, also local (red) buses, are allowed to use the less disruptive functions, but only really late buses having the most use of priority may use the more disruptive functions. The different functions are allowed for buses with different status as follows:

- **BF**, blue buses less than one minute early and red buses;
- **BF2+ÅTS**, blue buses less than one minute early;
- **AK**, late blue buses;
- **EF**, late blue buses;
- **DAK**, blue buses more than one minute late.

In order to evaluate the impacts of different priority conditions, five scenarios with different priority conditions have been tested with simulation experiments:

1. No Priority;
2. Existing conditional priority;
3. Existing conditional priority + short extension (BF) for red (local) buses;
4. Conditional priority with tighter on-time threshold + short extension (BF) for red buses;
5. Conditional priority with new on-time-status based conditions + short extension (BF) for red buses.

In scenario 4, the threshold to receive priority is narrowed from 2 minutes ahead of schedule to one minute. In scenario 5, the new conditions described above are used.

If conditional PT priority shall be used to improve the reliability of the bus service, a well-prepared timetable with correct timings at all stops is of great importance (8). In the studied case, almost all buses were late on Fleminggatan, but many of them were on time upstream and/or downstream; obviously the timetable is not correctly timed over the length of the line. In order to test the potential of conditional priority, the timetable in the simulated case is manipulated so that approximately the same numbers of buses are early and late. A way to avoid the problem of making a timetable that is well timed at all points on the line could be to change to an even
Johan Wahlstedt

headway strategy on high frequency lines (14).

The results shown in Figure 6 are the averages of fifty simulations with different random seeds. The travel time includes dwell time at stops. Since late buses are more likely to have many boarding and alighting passengers, late buses will have longer travel times than early ones. In the simulations the buses are classified according to their on-time performance when they enter the simulation model. They will not change status if they catch up or become more delayed on their way through the VISSIM network.

Giving the local red buses soft priority, i.e. allowing them to call for short extension (BF), reduces their travel times, but also increases the travel times somewhat for the blue buses and for GP traffic. If it, in total, saves travel time per person depends on the number of passengers on the local buses compared to the trunk route.

To tighten the threshold for receiving priority to one minute ahead of schedule seems to have the desired effect; the travel time increases for the early buses and decreases for the other buses that are late or on time. The travel time for GP traffic also shows a small decrease.

The proposed new more diverse conditions do not seem to have a much better regulating effect on the buses than tightening the threshold does, but the GP traffic travel time decreases. Slightly early buses have somewhat increased travel time. However the number of buses in the affected lateness span in the simulation is relatively small.

Figure 6: Travel times with different conditions for PT priority, including dwelltime at bus stops both directions added

CONCLUSIONS
PT priority in traffic signals can significantly reduce the delay at signals for public transport with small increases in delay for other traffic. The travel time savings per person are relatively large in the studied case. The negative impacts on other traffic are dependent on which priority functions are used. The impact on other traffic can be reduced by restricting the use of the more disruptive functions, but then the effect of PT priority will not be maximized.

Limiting the priority to buses with a higher need for travel time reductions seems to give greater benefits with less negative impacts. The decision to limit the priority to high demand routes or not is not easy to make.
REFERENCES


4. Wahlstedt, J, Simulering av bussprioritering i trafiksiganer, Master’s theses, KTH Royal Institute of Thecnology, Stockholm, Sweden 2005


Paper IV

Evaluation of the two self-optimising traffic signal systems Utopia/Spot and ImFlow, and comparison with existing signal control in Stockholm, Sweden

JOHAN WAHLSTEDT

Published in
Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems, October 2013, Den Haag, the Netherlands
Evaluation of the two self-optimising traffic signal systems
Utopia/Spot and ImFlow, and comparison with existing signal control in Stockholm, Sweden

Johan Wahlstedt

Abstract— The city of Stockholm has adopted the “Urban mobility strategy” stating that the more surface efficient modes public transport (PT), cycling and walk should be prioritised over cars. This has led to an interest to extend the signal priority for PT beyond its present level. Self-optimising signal control systems can keep the impacts on other traffic at a controlled level when introducing extensive PT priority, and they may reduce the cycle time causing less waiting time for pedestrians and cyclists. In this paper two self-optimising systems are tested and compared with each other and with the existing signal control using microscopic simulation. VISSIM was connected to PC versions of the self-optimising systems Utopia/Spot and ImFlow and to signal controller simulators including PRIUBUS. Both self-optimising systems reduced the average delay per person compared to existing signal control. ImFlow performed better than Utopia/Spot in the evaluation.

I. INTRODUCTION

A. Background

Stockholm is steadily growing and by the year 2030, the population of the City of Stockholm is expected to have increased by around one quarter from today. The trend today is that the city grows “inwards” making the city denser. If everyone continues to travel as we do today, the number of road journeys affected by congestion during the rush hour will increase five-fold to 2030, despite the construction of new roads and railways. It is not possible to widen the streets in the historic centre of the city or to build enough roads so that everyone can travel by car. This means that the streets need to transport more people and goods on the same surface area, i.e. there needs to be a shift to transport modes using the surface area more efficiently [1]. With this background the City of Stockholm has adopted the “Urban mobility strategy” (Swedish: Framkomlighetsstrategin) stating that the more surface-efficient modes of transport; public transport, cycling and walk should be prioritised over cars, and moving cars should be prioritised over parked cars.

One of the most important factors for the mode share is the travel time ratio between car and PT. Delay at signalised intersections constitutes a large part of PT journey time in urban areas. PT priority in traffic signals can reduce delay and improve service regularity for PT at a relatively low cost to other traffic [2-4]. Extended PT priority in traffic signals is therefore a natural part of improving the Urban mobility strategy. PT priority is not trivial to implement in conventional coordinated signal systems, where the “green waves” may be disrupted. It may also be necessary to make green time compensations to other approaches after the PT vehicle has passed which may have further negative impacts on the performance of the coordinated system [5]. PT priority can also result in extended waiting times for pedestrians, who also should be given priority according to the Urban mobility strategy.

Self-optimising signal control has been tested at a few locations in Sweden, but only one installation (Luthagseplanaden, Uppsala) is in operation today. Utopia/Spot was tested at two sites in Stockholm ten years ago. The results showed travel time savings around 5-10% compared to updated conventional control [6]. Despite the travel time savings the City of Stockholm chose not to implement self-optimising signal control at that time due to several reasons: the extra investment and maintenance costs, concerns about adaption of the system to Swedish control strategies and that there was only one supplier.

B. The “Adapt” project

With the background mentioned above, and the fact that two competing self-optimising signal systems now were available from the suppliers on the Swedish market, the City of Stockholm wanted to look closer at self-optimising traffic signals. A project called “Adapt” was started by the City of Stockholm. The objectives of the project were to make a beta test of the self-optimising traffic signal control systems available on the Swedish market in a simulation environment, compare them against each other and with the existing signal control scheme.

The main goal was to improve the PT-priority without too large impacts on other road users, and to improve accessibility for pedestrians and cyclists, in order to implement the Urban Mobility Strategy and minimise the overall delay per person. The main traffic engineering findings of the Adapt project are described in this paper.

C. Method of the Adapt project

The suppliers of self-optimising signal control systems on the Swedish market (e.g. Peak Traffic and Swarco) were invited to a “competition” where their systems would be tested in a simulation environment against each other, and against the existing fixed-time signal control with local traffic adaptation and PT priority. The suppliers of signal control equipment were invited by the City of Stockholm to participate in the project at a fixed remuneration. They were to provide a simulator version of their system connectable to the microscopic traffic simulation software VISSIM, and programming of their system for the test area including five signal controllers. They would also specify the location of...
additional loop detectors and/or PT priority calling points needed. One slightly modified version of the VISSIM model later used to evaluate the systems was provided to the suppliers, as well as the signal timing plans for the existing signal control. The final VISSIM models with different traffic flows used to evaluate the signal control systems, and the exact evaluation method were not revealed to the suppliers, only the goal to minimise the overall delay per person and that traffic flows from different time periods would be used. Based on these instructions the suppliers were free to set up and program their systems in the way they believed to be optimal.

The evaluation was done with “software-in-the-loop” simulations, i.e. the microscopic traffic simulation model VISSIM was connected to signal control simulators for each signal control strategy tested. The self-optimising signal control simulators were treated as a “black box” set up by the suppliers.

II. SIGNAL CONTROL

A. Conventional signal control

Conventional traffic signal technique is either fixed timed or vehicle actuated. Fixed timed signals have a fixed cycle-time and the timings are pre-set based on historical data on traffic flows etc. There can be some local signal timing adjustments based on vehicle actuation within the fixed cycle. Several fixed timed signals can be coordinated to achieve green waves in the main directions.

Conventional vehicle actuated signals in principle work with green-demand based on detection and green-extension, within pre-set limits, as long as the headways in the traffic stream are short enough on the detector. The cycle time is variable, and adjacent signalised intersections can therefore only be coordinated in special cases.

Swedish signal control strategies are signal-group based and do not use the ring-barrier concept common in U.S. In central Stockholm most signals are equipped with detectors and operate isolated-vehicle actuated at night. At day time they operate coordinated-fixed timed with local traffic adaptation, using a few seconds vehicle-actuated post-end-green (PEG) in order to react on small fluctuations in traffic flow locally [7].

B. PRIBUSS and PT signal priority in Stockholm

The PT priority method PRIBUSS was developed in Stockholm in the early 1990s and is now included as standard in most signal controllers sold in Sweden [15]. It is developed for conditional PT priority on top of the normal primary fixed-time control, but can also be used for isolated traffic actuated signal control [9]. The priority actions and compensations are done locally in each signal controller without consideration to neighbouring intersections. In the signal controller program, one or more PRIBUSS functions are configured; each function has a number of conditions and if all conditions are fulfilled, the corresponding PRIBUSS procedure is started. The conditions can be: bus counted in, (one or more) signal group green, (one or more) signal group red and a specified cycle time slot. The conditions are set in a way that only one function starts at a time. The six PRIBUSS procedures are: Extension, Re-taken start, Red truncation, Extra stage, Double red truncation, and Double extra stage.

There is no method for prioritising between PT vehicles with conflicting calls; the vehicle that calls for priority first will be served [7]. The buses and trams in Stockholm are equipped with an AVL (Automatic Vehicle Location) system that keeps track of location and schedule, and a short-distance data radio device for communicating directly with the signal controllers [8]. Only trams and buses on trunk routes running less than two minutes ahead of schedule receive signal priority, while early vehicles and local buses do not. Due to the large variation in dwell times at stops, calls for priority are not done upstream of a stop. If the stop is close to a traffic signal, the bus will call for priority when the door-brake is released and the bus starts driving.

C. Self-optimising signal control

Self-optimising control strategies try to adapt the traffic signals to the traffic flow in real time. They normally work by minimising a cost function for the calculated impacts of the signal timings. The cost function usually includes weighted costs for delay, stops etc. for all links in the system. The self-optimising system count vehicles with detectors and make predictions of their arrival times at the next signal as a part of the optimisation process. Most self-optimising systems are stage based and do not attempt to optimise the signal changes exactly after every vehicle, but rather optimise the overall signal settings. PT priority is often included directly in the optimisation process; a bus or a tram can be given higher weight in the cost function since it carries more passengers, e.g. equivalent to 20-50 cars. Since the PT priority is included in the optimisation, the impacts can be balanced out before or after the PT vehicle has passed.

A number of self-optimising signal control systems are available on the market, some of the most widespread systems are described below. The results mentioned are acquired with different methods and are not fully comparable with each other.

1) SCOOT

SCOOT is one of the first commercial, wide spread self-optimising signal systems for coordinated networks [10]. SCOOT has a centralised structure with a central computer that optimises the whole area communicating directly with every signal controller. The cycle time is optimised every five minutes, offsets every cycle and green split a few seconds in advance of every stage change. SCOOT only makes small changes from the previous cycle in order to make the system stable [10]. Newer versions including BUS SCOOT however allows larger occasional changes for PT priority. The emphasis in SCOOT is put on minimising the overall delay, according to British tradition [10]. The average travel time reduction when introducing SCOOT is typically around 11% [10].

2) SCATS

SCATS is an Australian partly self-optimising hierarchical signalling system, it could be described as a mix of a self-optimising system and a plan selection system [10]. The network is divided in to sub-systems with one critical intersection, for which cycle time and green splits are optimised in real time. Compatible timings for the
surrounding intersections in the sub-system are selected from a library of pre-set time plans, and some traffic adaption is done locally in the controllers. At a strategic level a choice can be made if two or more sub-systems should be joined with a common cycle time at the current traffic situation. SCATS is good for making green waves at the main road, and has showed main road travel time reductions of 18%. The effects on side roads and PT are unclear [10] [11].

3) **MOTION**

MOTION is a German partly self-optimising, hierarchical traffic signal system. It could be described as a plan selection and automatic updating system for fixed-time plans, combined with some local traffic adaption. The optimisation is done in a central computer with some optional, non-essential intelligence for traffic adaption distributed to the local signal controllers. Traffic flows at all links are predicted by a traffic model based on available detector data. The predicted traffic flows are used to select from pre-defined plans, or to adaptively calculate new plans. Optimal cycle time, green split, offset and basic stage sequence (from predetermined, allowed stage sequences), are calculated every 5 to 15 minutes for the whole coordinated system. If the new time plan is estimated to be significantly better than the currently used one, plans for each intersection is created and sent to the local controllers. Locally in the controllers some traffic-actuated changes can be done to react immediately to individual vehicles, e.g. to give priority to PT vehicles. The basic stage sequence can be modified by a traffic-actuated local control method for the current cycle, e.g. by inserting special stages etc. [12]. At central level only passive PT priority can be given, and active PT priority made locally in the controllers are not recognised by the central logic. The emphasis in MOTION is put on making good green waves for the main road according to German tradition.

4) **Utopia/Spot**

Utopia/Spot is a distributed self-optimising system developed in Italy, spread in a number of European countries, and is the only self-optimising system implemented in Sweden so far. Utopia/Spot optimises the signal timings by minimising a cost function locally in a Spot unit in each intersection. Spot units in neighbouring intersections communicate counted/predicted traffic flows and predicted signal changes for the next minutes. The neighbouring intersections are thereby adapting to each other’s predictions and thus creating coordinated green waves. In contradiction to the systems mentioned above a common cycle time is not necessary, the coordination is decided in each controller and each cycle. In Torino Utopia/Spot has resulted in reduced travel times of 2-7% for PT and 10% for cars. Field tests of Utopia/Spot in Stockholm have showed 10% travel time reduction for PT and 7% for cars [3].

5) **ImFlow**

ImFlow is a self-optimising signal system with distributed intelligence and a structure similar to Utopia/Spot. A multi-criteria cost function is minimised. The optimisation is done in two steps; stage based optimisation at network/route level based on a cost function with user defined weights, and signal group based optimisation at intersection level based on logical rules [13]. Both optimisation levels are distributed at the controllers. The actual signal controller can be allowed to override the optimisation, and the controllers decisions are sent back to the ImFlow optimiser. Different number and placement of detectors can be used. As in Utopia/Spot no common cycle time is calculated in ImFlow. In contradiction to Utopia/Spot the two step optimisation allows ImFlow to make signal changes exactly when the last vehicle in a platoon passes the stop line using a detector at the stop line, as with conventional coordinated control with local adaption. ImFlow is a new system, launched 2012, and there is only one installation implemented so far (Helmond, 2011). According to the manufacturer ImFlow reduced delay 10-20% compared to Utopia/Spot in this installation. A few systems will be implemented in different European countries in the near future.

III. **CASE STUDY**

**A. Method**

Analytical methods can be used for analysis of the impacts of fixed-time traffic signals as well as of some simple vehicle actuated control systems, but are normally not applicable for estimation of the impacts of more complex signal control such as self-optimising signal control and PT priority. It is, however, possible to study such control methods using simulation. A commonly used method in recent years for more detailed studies of signal control strategies is to link a commercially available traffic simulation model (VISSIM, Aimsun...) to a signal controller simulator for the studied type of signal controller [4].

From a scientific view, one problem with simulation experiments like this is that it is hard to generalise the conclusions from one case study. The results are valid for the studied area only. Many case studies need to be conducted in order to draw general conclusions. On the other hand, it is in general not possible to evaluate the effects of self-optimising signal control with analytical methods.

**B. Study area**

The studied area is situated in central Stockholm, Sweden and consists of five intersections in an X-shape. Due to geometrical reasons the central intersection (5515) needs an all-pedestrian phase, which limits the overall capacity of the coordinated system. Left turns are prohibited in the central intersection in order to keep the cycle time down, and the turning movements moved to the northern-most intersections (5516 & 5518). Those two intersections have complicated patterns with protected left-turns while the two southern-most intersections (5512 & 5514) have simple patterns with unprotected left turns and are less saturated.

The traffic flows have moderate AM and PM peaks and fairly high mid-day but drops at night, which is typical at Stockholm downtown main roads. The flow at Torsgatan (location marked in Figure 1) and its variation over time one average day is shown in Figure 2 below.

The two highly loaded trunk bus routes 3 and 4 traversing the study area operate with five minute headway. Six local bus lines also passing through the area operate with ten minute headway or less. The trunk bus routes are given signal priority and traverse the area in another route than the main traffic streams. The local buses are not given signal priority.
C. Existing signal control

The existing signal timing plan used on street today was developed in the Matsis project 2006 with the same simulation set up as used in this study. One single signal time plan with 100s cycle operates 06.30 – 20.00, covering all studied time periods. The signal plans were made by hand by an experienced traffic signal engineer, fine-tuned and evaluated with VISSIM simulations in the “Matsis” project 2006 [14]. Bus priority with PRIBUSS was introduced in 1999, but not updated with the new signal plans in the “Matsis” project. PRIBUSS functions for Extension, Red truncation and Extra stage are used. The bus priority is working, but the PRIBUSS settings are not optimal with the new signal timings. The bus priority in the central intersection (5515) is disabled due to major impact on green waves and capacity in the north-south main direction.

The goal of the Matsis project was to minimise the emissions by reducing the stops within the coordinated signal system, as opposed to minimising delay as in this project. The traffic flows and patterns have also changed since the signal plans were optimised. Traffic flows in central Stockholm have decreased approximately 10% due to the introduction of congestion charges. The introduction of congestion charges has affected both modal split and departure time of car trips [16]. Large construction works have also affected the routes used in the surrounding area.

D. Simulated scenarios

Three scenarios with different signal control strategies were simulated and evaluated;

- Existing fixed timed signals with local traffic adaption and PRIBUSS bus priority
- Utopia/Spot
- ImFlow

Each signal control scenario was simulated with five traffic conditions:

- AM peak (07.00 - 09.00)
- Mid-day (12.00 - 14.00)
- PM peak (16.00 - 18.00)
- PM peak with 5% increased traffic flows
- Evening (18.00 – 20.00)

The combination of the different signal control and traffic condition scenarios gives totally fifteen two-hour scenarios that were simulated with ten random seeds each. The results for different scenarios presented below are all based on an average of ten simulation runs.

E. Simulation setup

The method used in this case study was “software-in-loop”, a microscopic traffic simulation model connected to different signal controller simulators for respective scenario. The traffic simulation software used was VISSIM.

1) VISSIM model

The VISSIM network was originally developed in the “Matsis” project [14] and includes the five studied intersections and surrounding streets. The traffic flows and turning percentages in the area were counted in May 2012, an OD matrix developed from the counts and put into VISSIM. Pedestrians were counted at every pedestrian crossing and at one crossing at a time. Cyclists were counted in 2005 and are modelled throughout the network in order to be able to study the coordination effect on bicycles.

For the trunk bus routes actual bus arrival times and dwell times at stops were collected from the regional PT authority (SL) AVL system and inserted in the model. Separate dwell time distributions per stop, line and direction was calculated; the dwell time shows a large variation in the range 0 – 120 seconds with an average around 40 s, and most observations between 20 and 60 seconds. Local buses are modelled to run according to timetable with small random disturbance. Saturation flows at stop lines have been calibrated against field observations for some important movements.
2) EC1-simulator

The existing signal control strategy, including some vehicle actuated local adaption and PRIBUSS bus priority, was simulated with EC1 signal controller simulators. EC1 is one of the most used signal controllers in Sweden, and the simulator version of the signal controller was already used in the Matisis project in 2006. The programming of the signal controllers were done by the City of Stockholm. The five EC1-simulators and EC1-Simulator-Control controlling them were run on a separate computer connected to VISSIM via tcp-ip.

3) Utopia/Spot

The Utopia/Spot system consisting of a number of programs was installed on a separate laptop provided by the signal control supplier (Swarco). All programming and set up of the Utopia/Spot system was done by the supplier, and the system was used as a “black box” connected to the VISSIM model used for evaluation of the system. The Utopia/Spot and VISSIM computers were connected via tcp-ip communication.

4) ImFlow

The ImFlow system simulator was provided by the other signal control supplier (Peek Traffic) and installed on the same computer as VISSIM. The ImFlow simulator is an integrated program which calls and runs VISSIM via the COM interface. Also in this scenario the ImFlow system was programmed and set up by the supplier and was used as a “black box” connected to VISSIM.

IV. RESULTS

The aim of the Adapt project was to minimise the average delay per person. Both Utopia/Spot and ImFlow reduced the overall delay, as required by the project, however, the delay and number of stops for cars increased. ImFlow performed better than Utopia/Spot in most time periods.

Delay and number of stops in the whole network for: trunk route- and local buses, cars, cyclists and pedestrians were evaluated using VISSIM. The total delay time for all road users was calculated. Cars where multiplied by 1.2, which is the average number of passengers per vehicle in peak hours in Stockholm. Buses were multiplied with the average number of passengers indicated in Table 1 below. Since the load per bus is significantly higher in peak hours, different numbers of passengers were assumed for different time periods.

<table>
<thead>
<tr>
<th>Number of passengers</th>
<th>Bus on trunk route</th>
<th>Local bus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AM Mid</td>
<td>Day</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>40</td>
</tr>
</tbody>
</table>

As can be seen in Figure 3 below, both Utopia/Spot and ImFlow showed substantial decrease of total delay time compared to the existing signal control. ImFlow showed the best results in all but one time period.

Buses, pedestrians and cyclists gained from the tested self-optimising signal control strategies, partly at the cost of the cars. In the different time periods the average delay per bus was reduced 14-21% (Utopia/Spot) 21-24% (ImFlow); delay per pedestrian reduced 8-18% (Utopia/Spot) 26-35% (ImFlow); delay per cyclist 12-23% (Utopia/Spot) 28-33% (ImFlow). The largest decreases occurred off-peak when the cycle time could be reduced from the fixed 100 s cycle used in the existing signal control.

The change in average delay for cars was both positive and negative; in the increased traffic PM scenario the existing control was temporarily locally overloaded causing extra delay for cars. Both self-optimising signal systems managed to handle this situation and the delay per car was decreased 2% with Utopia/Spot, and 8% with ImFlow in this scenario. In the other time periods the average car delay was increased 0-23% with Utopia/Spot, and 5-37% with ImFlow. The larger increases occurred off-peak, and the smaller in the morning peak. The total travel time savings were larger when the number of road users was high, but the changes in travel time per vehicle, both up and down, was larger off-peak when larger changes could be made.

Both self-optimising systems showed a substantial increase in number of stops, compared to the existing control. The existing fixed timed control created good “green waves” which the self-optimising systems could not match.
V. DISCUSSION

After the evaluation was done it was found that calling points for PT priority upstream of stops where used in the tested Utopia/Spot programming. The large spread in dwell times thereby destroys the optimised timings. Later tests with changed programming show better results. The timings of the existing control have been optimised for the simulated traffic flows with TRANSYT connected to VISSIM, showing large improvements compared to fixed-time control with existing timing. On-going work looks at possible improvements on all tested control strategies.

In this project the main criteria for optimising the signal control was total delay for all travellers, but with the tested programming both Utopia/Spot and ImFlow caused a large increase in number of stops for cars. An increase in number of stops results in increased fuel consumption, emissions and noise. Another balance between the costs on stop and delay in the programming of the self-optimising signal control systems would perhaps give a better performing signal control strategy in total, but less delay time savings. Future work should look at the optimal trade-off between stop and delay.

One approach could be to use the road users’ perceived value on delay time and stops respectively in order to achieve a signal control regime causing the least negative experience by the travellers. The perceived value of delay/stops is probably regionally dependent and not well known; more research in the area is needed before these criteria could be used.

A cost-benefit analysis could also be done to evaluate the effect of self-optimising signal control systems. Standardised monetary values on stops and delay can be used to balance the control strategy in an “optimal” way. Using the standardised values on delay for different modes of road users do not necessarily give the same optimum as asked for in the Adapt project: prioritising surface efficient modes of transport in order to implement the Urban mobility strategy adopted by the City of Stockholm.

The controlled simulation environment ensured the same traffic flows, as well as other conditions, for each signal control strategy in all tested time periods. All but the signal control is exactly the same in the different scenarios in the simulation environment. To do the same evaluation of three control strategies in field tests would require much more field data due to day-to-day variations of traffic flows etc. Neither is travel time data easy to collect for all groups of road users in the real world.

Both Utopia/Spot and ImFlow need additional loop-detectors, placed at different locations in both systems, and at least ImFlow also requires other signal controller hardware than currently existing in the studied signal system. This would have required substantial costs in a field test, costs which were saved with the simulation approach.

However, a simulation environment can never capture the whole of reality and the impacts of irregularities as parking cars, irregular vehicle movement at detector loops, detector errors and communication problems etc. which are not captured in the simulations. Field tests are necessary to confirm the positive simulation results.

VI. CONCLUSION

Both Utopia/Spot and ImFlow managed to extend the bus priority and improve the accessibility for pedestrians and cyclists in this test case. A negative impact was poorer “green waves” causing somewhat longer delays and more stops for cars. The overall delay per person was reduced with both tested self-optimising signal control systems. ImFlow showed better results than Utopia/Spot in all but one studied time period. To what extent this was due to the performance of the system and/or due to the actual programming of the systems in the studied case is not shown. The statistical significance of the results also remains to be tested.

Both tested self-optimising signal control systems seem to be useful tools to allocate the accessibility between different travel modes in a fairer way compared to conventional coordinated signals, in order to minimise the overall delay on busy city streets.

REFERENCES

[8] Wahstedt, J, Simulerings av bussprioritering i trafiksignaler, Master’s theses, KTH Royal Institute of Thechnology, Stockholm, Sweden 2005
Handbok för kapacitetsanalys med hjälp av simulering, kapitel 6
Signalreglerade korsningar

JOHAN WAHLSTEDT

Published by
The Swedish Transport Administration, Trafikverket
publication TRV 2013:79994, Borlänge, Sweden
Innehåll

Kapitel 6 Signalreglerade korsningar

6 Signalreglerade korsningar.................................................................2
6.1 Introduktion .......................................................................................2
6.1.1 Trafiksimulering ..............................................................................2
6.1.2 Trafiksignaler ...................................................................................2
6.1.3 Trafiksignalterminologi ....................................................................2
6.1.4 Säkerhetstidsberäkning ..................................................................3
6.1.5 Regleringsformer .............................................................................3
6.1.6 Signallyktor .....................................................................................3
6.1.7 Signalreglering i simulering .............................................................3
6.2 Genomförandeprocessen ....................................................................4
6.3 Syfte och avgränsningar .....................................................................4
6.4 Databehov ..........................................................................................6
6.5 Uppbyggnad av basmodell .................................................................7
6.5.1 Korsningsgeometri, trafikmängder och rutter ...................................7
6.5.2 Påverkan på mättnadsflöde ...............................................................8
6.5.3 Styrapparat och signalväxling ..........................................................10
6.5.4 Signallyktor och detektorer ............................................................ 15
6.6 Verifiering, kontroll och felsökning .................................................. 16
6.7 Kalibrering ....................................................................................... 17
6.8 Validering ....................................................................................... 17
6.9 Analys av olika alternativ/scenarier .................................................. 17
6.10 Dokumentation.................................................................................. 18
6.11 Litteraturreferenser.......................................................................... 18
6 Signalregerade korsningar
Huvudförfattare: Johan Wahlstedt
Granskare: Jeffery Archer, Johan Olstam, Benny Bergstrand

6.1 Introduktion

6.1.1 Trafiksimulering

6.1.2 Trafiksiganer

6.1.3 Trafiksiganerterminologi
Trafiksiganer och deras funktion beskrivs närmare i VGU (Trafikverket och SKL 2012). I TRV2013/64343 kapitel 6 Signalregerade korsningar beskrivs styrätt (avsnitt 6.1.1); signalbegrepp (6.1.2) samt termer och beteckningar (6.15) som också är tillämpbara i detta kapitel.
6.1.4 Säkerhetstidsberäkning


<table>
<thead>
<tr>
<th>Sparrmatriser för ANL: 226-14, Länsvägen</th>
<th>Datum:</th>
</tr>
</thead>
<tbody>
<tr>
<td>TILL FRAMRYCKAende</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Figur 1 Exempel på konfliktmatris för beräkning av säkerhetstider.

6.1.5 Regleringsformer

Se TRV2013/64343 avsnitt 6.5.3.

6.1.6 Signallyktor

Se TRV2013/64343 avsnitt 6.1.2

6.1.7 Signalreglering i simulering

En fördel med simulering är det möjliggör ”kontrollerade experiment” i datormiljö för att testa alternativa styrmetoder, signalanordningar, geometrisk utformning mm. Dessa alternativ behöver inte vara beskrivna i varje detalj men vara möjliga att implementera. Värt att tänka på i dessa sammanhang är följande:

- Deltillfarter reglerade med pilsignaler får aldrig ha sekundärkonflikt
- Dubbla svängande körfält får av säkerhetsskäl inte ha sekundärkonflikt
- Om en vänstersväng är separatreglerad (pilsignal) måste även mötande vänstersväng vara separatreglerad (av säkerhetsskäl)
- Ökat antal körfält ger ökade säkerhetstider samt mintider för gående
För en överslagsmässig uppskattning av signaltidssättning kan följande approximationer användas:

- Gröntidsbehovet vid trafikavveckling utan sekundärkonflikter är c:a 2 sekunder per bil och körfält.
- Allrödtiden mellan konflikterande grupper i en mindre korsning är normalt 1-2 sekunder, och ökar för större korsningar pga längre utrymningssträckor.
- Mintider för övergångsställen bör beräknas; \( L/v_{gång} \) där \( L \) = hela övergångsställets längd i meter och \( v_{gång} \) gånghastighet 1,2-1,4 m/sek
- Utrymningstider för övergångsställen bör beräknas: \( Lu/1,4 \) där \( Lu \) = längden mellan kantsten och refug/kantsten, se vidare TRV2013/64343 avsnitt 6.1.4.

Signaltidssättningen kan sedan justeras utifrån simuleringen. CAPCAL kan även användas för att få en gröntidsfördelning och omloppstid att utgå ifrån. För samordnade signaler kan TRANSYT eller liknande program användas.

**6.2 Genomförandeprocessen**

Se kapitel 2.2

**6.3 Syfte och avgränsningar**

Simulering är en lämplig metod för kapacitetsanalys om överslagsberäkning eller analytiska beräkningar inte är tillräckliga, samt om en visualisering av resultaten önskas för presentation för allmänhet, beslutsfattare m.fl.

Syftet med simuleringen och frågeställningen som ska besvaras avgör hur detaljerat trafiksinalerna behöver modelleras. En simuleringsmodell är naturligtvis en förenkling av verkligheten, och beroende på syftet kan förenklingen drivas olika långt. Valet av styrsätt i simuleringsmodellen beror även på hur signalen styrs i verkligheten. Samordnade signaler är huvudsakligen tidstyrda och kan modelleras som fast tidstyrda (men hänsyn kan behöva tas till ev. bussprioritering eller specialfunktioner) medan oberoende trafikstyrda liksom länkade signaler i vissa fall kan modelleras förenklat som fast tidstyrda, men i andra fall bör modelleras som trafikstyrda. I de fall trafiksinalen i sig inte är intressant, men ligger så nära det studerade objektet att den påverkar ankomstfördelningen betydligt, kan signalen modelleras förenklat.

Metoderna att styra signalerna i en trafiksimuleringsmodell på mikronivå kan med stigande detaljeringsgrad delas in i:

- Förenklade signaler
- Tidstyrda signaler
- Trafikstyrda signaler med VAP styrning eller motsvarande
- Extern styraapparatssimulator

**Förenklade signaler**


**Tidstyrda signaler**

Fast tidstyrda signaler är enklast att koda in i modellen och ger i många fall tillräcklig detaljeringsgrad. Kapaciteten underskattas i vissa fall något medan fördröjning och kölängd överskattas om signalen i verkligheten är trafikstyrd. Simulering med fast tidstyrda signaler är normalt tillräckligt detaljerad för att avgöra om en föreslagen lösning har tillräcklig kapacitet eftersom en trafikstyrk signal beter sig som en tidstyrd vid hög belastning. Om signalerna i verkligheten är samordnade går det normalt bra att simulera dem som fast tidstyrda.

**Trafikstyrda signaler med VAP styrning eller motsvarande**

Många trafiksimuleringsprogram ger möjlighet att modellera trafiksignaler med en förenklad trafikstyrning. I VISSIM går det till exempel att programmera önskade signalfunktioner med det grafiska programmeringsverktyget ViSVAP, i andra programvaror finns liknande möjligheter. Trafikstyrda signaler går härigenom att programmera något förenklat liksom kollektivtrafikprioritering, bomfällning vid järnvägsövergångar mm. Detta ger normalt tillräcklig detaljeringsgrad för att studera fördröjning, kölängd mm, och för att jämföra en trafikstyrk signal med andra regleringsformer (väjning, cirkulationsplats osv).

**Externt styraapparatssimulator**

6.4 Databehov

Platsbesök bör göras vid korsningen eller området som simuleras. De indata som huvudsakligen behövs för simulering av signalreglerade korsningar är:

- Trafikflöden
- Korsningsgeometri
- Funktionsbeskrivning och kabel/detektorplan för signalen
- Hastighet i svängande rörelser
- Hastighetsspridning (samordnade signaler)
- Väjningsbeteende
- Körfältsutnyttjande

**Trafikflöden** på minst timnivå med andel svängande trafik samt tung trafik behövs för alla trafikströmmar. Ungefärligt antal gående per övergångsställe uppskattas om räkningar inte finns tillgängliga. En kortvarig räkning i samband med platsbesök kan ge betydande insikt.

**Geometrin** behöver främst visa antal körfält och deras längd (t.ex. längd på vänstersvängficka). Därutöver behövs information om stopplinjernas placering och möjlighet att magasinera svängande fordon som väger mot mötande, fotgängare mm) i korsningen utan att blockera bakomvarande fordon.

Busshållplatser nära korsningen har också betydelse om inte antalet bussar är försumbart.

**Funktionsbeskrivningen** för en befintlig signalanläggning visar hur signalen växlar och hur detektorerna fungerar. Stopplinjer och detektorers placering (trafikstyrd signal) fås från kabel/detektorplan. Finns ingen dokumentation för signalanläggningen (t.ex. inte byggd ännu) får en rimlig funktion skissas upp för simuleringen, se ovan. Noggrannheten vid bestämning av kapacitet, fördröjning mm. blir naturligtvis beroende av kvaliteten på dessa antaganden.

**Hastigheten i svängande rörelser** har stor betydelse för mättandsflödet och därmed kapaciteten i korsningen. Om möjligt bör hastigheten i de svängande rörelserna mätas upp stickprovsmässigt, annars uppskattas de utifrån radierna/körgeometrin.

**Hastighetsspridningen** har betydelse för kolonnspridningen mellan samordnade signaler. Om signalsamordningens funktion studeras i simuleringen bör hastighetsfördelningen på sträckan mellan signalerna mätas upp och läggas in i simuleringsmodellen.

**Väjningsbeteendet** vid sekundärkonflikter (mot mötande fordon, gående och cyklister etc.) är platsberoende och bör om möjligt studeras genom platsbesök. Om den simulerade korsningen inte går att studera (ännu inte byggd etc.) kan beteendet i en annan korsning i området studeras. Väjningsbeteendet har en avgörande betydelse för mättnadsflödet och därmed kapaciteten i korsningen.

**Körfältsutnyttjandet** för parallella körfält har betydelse för kapaciteten och beror på närliggande korsningar, ändrat antal körfält mm och varierar därför från plats till plats.
6.5 Uppbyggnad av basmodell

Följande tillämpningsexempel är utfört med simuleringsprogramvaran VISSIM. Exemplet visar främst hur en modell med trafiksignaler kan byggas upp, inte hur man använder VISSIM i allmänhet.

6.5.1 Korsningsgeometri, trafikmängder och rutter

Korsningens byggs upp efter grundkarta eller signalanläggningsplan, en karta/plan i .dwg format är att föredra. I brist på karta kan flygfoto användas, men dessa är sällan fullt skalenliga och viss information saknas. Längden på eventuella svängfält fästs normalt från grundkarta men den längd som i praktiken används, mht. busshållplatser, parkerade bilar mm. kan behöva stämmas av vid platsbesök. Geometrin utformas så att magasinsutrymme för svängande bilar stämmer med observerat beteende. Övergångsställen och separata cykelbanor bör i förekommande fall läggas in. I den mån det går bör överlappande länkar/konnektorer undvikas.

![Figur 2 Exempel på parametrar för konnektor](image)

För konnektorerna kan ”Emergency stop distance” anges, detta är avståndet till den punkt där ett fordon som ska byta till ett körfält som leder till den aktuella konnektorn slutar försöka byta körfält och stannar för att invänta en lucka. Genom att ange en punkt uppströms signalens stopplinje motsvarande den heldragna körfältsmarkeringen undviks orealistiska körfältsbyten nära stopplinjen, samt säkerställs att fordonen passerar ”signal head”/stopplinjen för det aktuella körfältet om länken uppströms har flera körfält.

Parametern ”Lane change” anger hur långt uppströms fordonen ”får reda på” att de ska till den aktuella konnektorn för att följa sin rutt och därmed var körfältbyte i riktning mot konnektorn påbörjas. För att detta ska ha effekt
måste ruttvalet ligga uppströms om den angivna punkten om ”static routing” används. Dessa båda parametrar påverkar körfältsbytena och bör kalibreras mot observerat trafikantbeteende.

I VISSIM kan ruttvalet göras statiskt eller dynamiskt. Med ”static routing” skapas trafiken i ”vehicle inputs” och leds genom modellen via statiska rutter som anges av användaren. Med ”dynamic assignment” genereras trafiken med en OD matris och ruttvalet genom nätet väljs med en iterativ process av VISSIM. ”Dynamic assignment” kräver mer arbete och är endast meningssfullt vid större modeller där det finns flera möjliga vägar genom nätverket. Nedan beskrivs tillvägagångssätt med ”static routing”.

Trafikmängder läggs in för alla tillfarter, i förekommande fall med olika andel tung trafik mm. Beroende på hur väl trafikmängderna är kända kan de läggas in som timtrafik under hela simuleringsperioden eller varierande över t.ex. 15 minuters intervall för att kunna ta hänsyn till effekterna av korttidsflödesvariationer.

Rutter läggs in genom korsningarna. Om det finns flera körfält i samma riktning mellan närliggande korsningar och korsningsavståndet är kort kan rutterna behöva göras genomgående för flera korsningar. Detta rekommenderas för att körfältsbyten ska kunna ske på ett korrekt sätt, se parametern ”Lane change” för konnektorer ovan.

### 6.5.2 Påverkan på mättnadsflöde

Ett av de mest betydelsefulla effektmåten för trafiksignaler är mättnadsflöde, det största antal fordon per gröntimme som kan passera stopplinjen vid mättade förhållanden. I en trafiksimuleringsmodell på mikronivå är mättnadsflöde inte indata som kan anges explicit, utan ett resultat. Mättnadsflödet beror främst på indata avseende car-followingparametrar, hastighet genom korsningen samt väjning vid sekundärkonflikt. VISSIM har med standardparametrar en tendens att ge något för höga mättnadsflöden.

![Figur 3 Exempel på hastighetsreduktion för svängande fordon](image-url)
Hastigheten i svängande rörelser minskas med "reduced speed areas", en "desired speed distribution" per fordonsslag anges för det angivna området. Fordonen anpassar hastigheten innan de når en "reduced speed area" och accelererar åter till "desired speed" efter att ha lämnat den. Den valda hastigheten har stor påverkan på mättnadsflödet, om mättnadsflödet blir för högt kan en "reduced speed area" läggas in även i genomgående körfält.

Primärkonflikter regleras alltid av trafiksignalen medan sekundärkonflikter i de flesta fall regleras av väjningsregler i trafiklagstiftningen, se kapitel 8.1. Väjning kodas i VISSIM antingen med "priority rules" eller "conflict areas". Att koda väjningar med "priority rules" kräver mer arbete men ger större flexibilitet än med "conflict areas". Båda typerna kan kombineras efter behov.

Väjningsregler i alla sekundärkonflikter behöver läggas in i VISSIM modellen, och om tillbakablockering ska kunna hanteras även primärkonflikter. Var fordonen stannar för att väja har betydelse för korsningens kapacitet, det påverkar om bakomvarande fordon blockeras när ett eller flera svängande fordon väjer mot gående eller mötande trafik och om fordon magasineras i korsningen. Detta bör om möjligt studeras vid platsbesök.

**Figur 4  Exempel på väjning med "priority rules"**


6.5.3 **Styrapparat och signalväxling**

Trafiksinalernas styrs såväl i verkligheten som i VISSIM av en styrapparat, i VISSIM definieras en ”signal controller” och till den kopplas ”signal heads”/stopplinjer, detektorer mm. I styrapparaten definieras ett antal signalgrupper som styr tillhörande ”signal heads”. Signalgrupperna bör om möjligt följa de verkliga, deras indelning fås från signalväxlingsschema, se fFigur 6.8 och signalplan fFigur 3. För respektive signalgrupp anges i VISSIM växlingsschema samt rödgul och gultid (gröntblink). OBS att VISSIMs standardvärden för gult och rödgult avviker från svenska regler, se TRV2013/64343 kapitel 6.

Styrapparaten kan styras på olika sätt, de vanligaste är fast tidsstyr, ”fixed time”, trafikstyr med VISSIMs VAP/VisVAP logik eller med extern styrapparatssimulator se även avsnitt 6.3. ovan. Här behandlas främst fast tidsstyrda signaler vilket är det enklaste fallet.
Figur 1 Exempel på signaltidsättning i VISSIM

I VISSIMs ”signal controller” skapas ett eller flera ”program”, för dessa anges omloppstid samt gröntid i form av rödslut och grönslut för respektive signalgrupp. Dessa tider fås från signalväxlingsschemat, se exempel i fFigur . Finns inte signalen i verkligheten, eller om signaldokumentationen inte är tillgänglig får en rimlig signaltidsättning och tillhörande körfältsindelning som uppfyller gällande regler ansättas, se även avsnitt TRV2013/64343 kapitel 6.1. En signaltidsättning och/eller körfältsindelning (samt annan geometrisk utformning) som inte uppfyller gällande regler ger ett ogiltigt simuleringsresultat.

Tidsstyrda samordnade signaler

Som beskrivits i TRV2013/64343 kapitel 6.1. kan signalens styrform vara (beroende) trafikstyrd eller (samordnad) tidsstyrd, i de flesta fall finns minst en tidplan för oberoende styrning och i den finns grunddata såsom mintider mm. År signalen samordnad finns ofta flera tidstyrda program för olika trafikfall/tidsperioder på dagen. Vid simuleringen väljs den tidplan som motsvarar det studerade trafikfallet. Simuleras flera tidsperioder, t.ex. förmiddagens och eftermiddagens maxtimme, behövs normalt olika signaltidssättning för dessa tidsperioder. Om framtida fall med ändrade trafikflöden och/eller geometri studeras kan signaltidsättning och körfältsindelning (geometri) behöva ändras motsvarande vad som kan antas ske i verkligheten vid ändrade trafikförhållanden.
Figur 7 Exempel på signalväxlingsschema för samordnad styrning

Ett samordnat program är normalt huvudsakligen fast tidstyrt, men kan ha viss lokal anpassning i form av signalgrupper som endast går till grönt efter anmälan, trafikstyrda fräntid mm. Signalväxlingsschemat visar signalväxlingen när alla signalgrupper går in och tar ut full förlängning av trafikstyrda tider, se exempel i Figur 7 ovan. Detta ger normalt högst kapacitet, men kan resultera i något för hög fördröjning, och är normalt en lämplig utgångspunkt vid simuleringen. I exemplet ovan visas ett startomlopp (överkryssat) och ett normalt omlopp där det senare ska användas.

För en samordnad (fast tidstyrd) signal kan omloppstid och gröntider från signalväxlingsschemat i de flesta fall mer eller mindre direkt sättas in i VISSIM om fräntider mm. förutsätts tas ut helt.

Trafikstyrda oberoende signaler
Om en (oberoende) trafikstyrad signal ska modelleras som fast tidsstyrd behöver vissa antaganden och förenkningar göras. Effekterna av signalen blir annorlunda än med korrekt signalstyrning, åtminstone om det inte är kö i alla tillfarter. Beroende på simuleringens syfte kan detta i alla fall vara en fullt tillräcklig approximation, se kapitel 6.3 ovan.
Figur 8 Exempel på signalväxlingsschema för oberoende styrning

Signalväxlingsschemat visar hur signalen växlar om den startar från allrött om alla detektorer är belagda och därmed alla signaler anmäler grönbehov och förlänger alla trafikstyrda tider fullt ut, se exempel i figur 8 ovan. Observera att signalväxlingsschemat inte alltid är skalenligt och att den verkliga växlingen enligt de i signalväxlingsschemat angivna funktionerna kan se annorlunda ut. Signalen startar normalt inte heller från allrött under de förhållanden som simuleras, så det går inte att rakt av sätta in de tider som visas i signalväxlingsschemat i VISSIM.

Vanligtvis är det rimligt att anta att alla grupper alltid har grönbehov och att all trafikstyrd grön- och gultid tas ut. Om max- och/eller fråntiderna är långa i förhållande till den simulerade trafikbelastningen kan en mindre del användas, t.ex. 75% av maxtid och 50% av fråntid, detta måste bedömas från fall till fall. Det är också troligt att sidotillfarterna tar ut en mindre del av max gröntid än huvudvägen. Variabel gultid kan antas tas ut fullt. Variabel röd tid kan användas för rödkörningskontroll (R-funktion i LHOVRA) eller för att fordon som magaserats i korsningen ska hinna lämna denna innan nästa grupp får grönt. Om den variabla rödtiden används för rödkörningskontroll kan det antas att den inte tas ut, medan en lämplig del av den variabla rödtiden för att utrymma magaserade fordon bör tas med.
För att förenkla den trafikstyrda signalstyrningen till fast tidsstyrd görs en delvis iterativ genomgång av signalväxlingen enligt följande:

1. Välj en lämplig fasbild att börja med som har en enkel inväxling
2. Bestäm gröntider genom att kontrollera mintid och trafikstyrda maxtid för de grupper som har gemensamma fientligheter mot grupp(er) i nästa fasbild. Grupper med gemensamma fientligheter har samma frånväxlingspunkt, den längsta tiden väljs.
3. Addera trafikstyrd fråntid och grönblink efter växlingspunkten för de grupper som har det, individuellt för varje grupp.
4. Addera eventuell trafikstyrd rödtid.
5. Bestäm grönstart för grupper i nästa fasbild genom att kontrollera säkerhetstider mot frånväxlande grupper i spärrmatris, se exempel i Figur 2 nedan, och eventuell trafikstyrd rödtid om denna är längre än säkerhetstiderna. Grupper med samma fientligheter kan ha olika inväxlingspunkt beroende på olika säkerhetstider.
8. Beräkna omloppstid för signalväxlingen

När signalväxling bestämts enligt ovan kan den sättas in i VISSIM som fast tidsstyrd. Om signaltidsättningen inte stämmer med de simulerade trafikförhållandena kan de uttagna max- och fråntiderna justeras utifrån simuleringens animering.
6.5.4 Signallyktor och detektorer

Körflätsindelning, stopplinjer (målning) och signalernas placering fås från signalanläggningsplanen (ibland kombinerad med kabelplan), se figur 10. Stopplinjernas placering/målning bör stämmas av vid platsbesök då denna mer eller mindre avsiktligt kan ha ändrats vid asfalterings/målningsarbeten. Om den simulerade korsningen inte finns i verkligheten eller om dess utformning ska ändras ansätts en rimlig körflätsindelning, denna måste stämma med den regleringsform och signaltidsättning som används för att uppfylla gällande regler (se vidare TRV2013/64343 kapitel 6.1) för att simuleringsresultatet ska vara användbart.

I VISSIM modelleras signallyktor och stopplinjer med en ”signal head” per körfläkt placerad vid stopplinjen, se Figur 11. Denna kopplas till en ”signal
controller” och en signalgrupp i denna, om ”Or Sig. Group” används blir signallyktan grön om någon av de angivna grupperna är grön. Detta används om ett körfält ska regleras med flera signalgrupper, t.ex. om tvåskens undantagssignal för högersväng finns. I detta fall anges både signalgruppen för cirkulärt sken och för högersvängspilen. För gående läggs ”signal heads” ut vid kantstenen samt vid refug om gående i verkligheten stannar på refugen när signalen slår om till rött, annars inte.

I VISSIM kan detektorer läggas in, läge och numrering för dessa fås från signalanläggningsplanen, se Figur 4. Även tryckknappar för gående och cyklister modelleras som detektorer liksom radiodetekteringspunkter för bussprioritering. Detektorer behövs inte för tidstyrda signaler utan endast i kombination med VAP styrdning eller extern styraparatssimulator.

**6.6 Verifiering, kontroll och felsökning**

Efter att simuleringsmodellen byggts upp och trafiksignalen kodats in bör en visuell kontroll av signalens funktion göras genom animering:

- Är stopplinjer korrekt placerade, stannar alla fordon där de ska?
- Har konflikterande grupper grön samtidigt?
- Har någon signal rött tillsynes omotiverat?
- Är undantagssignaler korrekt kodade?
- Verkar säkerhetsstiderna rimliga eller tenderar fordonen att ”krocka”?
- Uppstår köer i någon tillfart som kan antas ha tillräcklig kapacitet?
- Uppstår köer i tillfarter som antas vara högt belastade?

Om något verkar konstigt, kontrollera mot funktionsbeskrivningen.
Kontrollera även att:
- hastighetsbegränsningar finns i alla svängande rörelser,
- värningar är kodade för alla sekundärkonflikter och vid behov även för primärkonflikter om tillbakablockering riskeras
- Vid trafikstyrdasignaler; använd ”test mode” för att kontrollera att detektorerna amäleri och förlänger rätt signalgrupp

6.7 Kalibrering
Kalibrering kan göras mot uppmätta värden i den mån de finns tillgängliga och/eller är mätbara samt visuellt så att modellbeteendet stämmer överens med hur trafikanterna beter sig i verkligheten. Vid platsbesök studeras de beteenden som ska kalibreras, den som utför den visuella kalibreringen behöver skaffa sig en ”känsla” av trafikanntävandelet på den aktuella platsen.

De viktigaste kalibreringsmåtten på signalreglerade korsningar är
- Genomströmning (uppmätta trafikflöden som går genom korsningen)
- Fördröjning/restid
- Kölängd
- Mättnadsflöde (vilket i en mikroskopisk simuleringsmodell ges indirekt av ett antal parametrar)

Parametrar som bör kalibreras är:
- Magasinering i korsningen
- Hastigheter i svängande rörelser
- Väkningsbeteende vid sekundärkonflikter
- Körflätsanvändning och körförlustenbyten
- (Acceleration)
- (Tidsavstånd vid köavveckling)

Kontrollera hur den valda programvaran definierar kölängd innan kölängder mäts upp i fält för att om möjligt anpassa mätningarna så att jämförbarhet mellan data uppnås.

Se i övrigt kapitel 2.7 om kalibrering.

6.8 Validering
Se kapitel 2.8 om validering.

6.9 Analys av olika alternativ/scenarier
Obs! För att resultaten ska vara giltiga kan tidsättningen av signalerna behöva ändras om trafikflödena förändras i alternativa utformningar av vägnätet eller med förändrad trafikefterfrågan.

Frågeställningen och syftet med simuleringen styr vilka utdata som väljs för analys. Ofta är trafikflöde (genomflöde) kölängd och restid eller fördröjning lämpliga utdata.
6.10 Dokumentation

Innan simuleringsprojektet avslutas bör det dokumenteras för att man i efterhand ska kunna värdera resultaten och de slutsatser som dragits, samt för att den framtagna modellen i ett senare skede ska kunna användas vidare på ett adekvat sätt. Det som åtminstone bör dokumenteras är använda indata i form av trafikmängder och vilka signalplaner (eller motsvarande) som använts, hur signalerna styrts samt vika förenklingar som gjorts. Kända brister och osäkerheter i modellen bör även dokumenteras.

Se i övrigt kapitel 2.10 om dokumentation.

6.11 Litteraturreferenser

Se avsnitt 1.4 samt TRV2013/64343 avsnitt 1.6.