Transit priority performance at signalised junctions

J. TAN, ITRA Ingénierie du trafic SA

Conference paper STRC 2005

STRC | 5^{th} Swiss Transport Research Conference
Monte Verità / Ascona, March 9-11, 2005

Transit priority performance at signalised junctions

Jaubert Jianan TAN
ITRA Ingénierie du trafic SA
1253 Vandoeuves / Geneva

Phone: ++41-22-860-0981
Fax: ++41-22-860-0982

email: jaubert.tan@itra.ch

March 2005

Abstract

In order to reduce transit (bus or tram) delays at intersections and protect them from the effects of growing traffic congestion and hence make them more attractive and more reliable, the transit priority strategy (such as transit lane concept and transit priority treatment measures) is widely applied at signalized junctions.

As each signalised junction is different, the same bus priority method may give different results. It is quite difficult to quantify the transit priority performance and its consequences in a general way. Based on experiences of junction geometry design, signal design with transit priority treatment and controller programming for more than fifty signalised junctions in Geneva, the author tries to develop a method to evaluate quantitatively transit priority performance and to estimate its impact on non-transit traffic in junction geometry design and junction signal design.

In the junction geometry design step, two new definitions, “transit greenable length” and “transit priority performance ratio”, are introduced. It is shown that the average transit delay depends on them and formulas to estimate the average transit delays (without transit priority treatment) are developed. Then the transit priority performance of “transit lane concept” at a signalised junction is analysed.

In the junction signal design step, different transit priority treatment measures can be applied so that the signal green interval for a transit movement matches the transit arriving interval and the transit delay will be minimised. The choice of transit priority treatment measures in different signal designs is discussed.

In summary, a good junction geometry design taking account of transit priority performance will limit transit average delay. A good signal design combined with a good junction geometry design and suitable transit priority treatments may considerately decrease transit delay. Therefore a high-transit-priority-performance signalised junction should have a high transit priority performance ratio, suitable transit priority treatments and acceptable impact to non-transit traffic. Finally, some further researches are suggested.

Keywords

Transit priority – Intelligent traffic control – Signalised junction – Bus priority – Tram priority - Public transport – Junction design – Signal design

1. Introduction

Bus and tram (named as transit here) are significantly more efficient than private cars in their use of road space, carrying many more passengers for the same amount of space. However, they suffer from the adverse effects of traffic congestion caused by private vehicles. The congestion can make them slow, unreliable and unattractive to use. In order to reduce transit delays at intersections and protect transit from the effects of growing traffic congestion and hence make transit more attractive and more reliable, the various transit priority methods are used.

There are following main types of transit priority methods:

a. Transit (bus or tram) lane concept (lane only reserved for transit).

b. Transit priority treatment (signal timing or signal control logic with priority to transit).

c. Transit lane combined with transit priority treatment

Evidently the measurement of transit priority performance is transit delay. The highest performance is that transit passes the junction without delay. The practice in Geneva confirms that transit delays are reduced more or less at signalised junctions with some consequences on non-transit traffic. As each junction is different, the same transit priority method may give different results. It is quite difficult to quantify the transit priority performance and its consequences in a general way. It was said that "the success story of bus priority is the customizing of the priority functions according the signal phasing and the junction geometry. Each junction is different and no permanent rules can be given for designing the best alternative" (1).

Based on experiences of junction geometry design with / without transit lanes, junction signal design with transit priority treatment and controller programming for more than fifty signalised junctions in Geneva, the author tries to develop a method to evaluate quantitatively transit priority performance in a signalised junction design.

A signalised junction design is usually consists of the junction geometric design and then the signal design which is based on the junction geometric design chosen. In the junction geometric design step, the use of transit lane may be analysed. And in the junction signal design step, different transit priority treatment measures may be studied.

2. Junction geometry design and transit priority performance

2.1 Junction geometry design and junction capacity used

In a junction geometry design step, one or more design alternatives may be developed. Each alternative is consists of following data:

- A junction geometry plan with a clear definition of lane utilisation (goes to / comes from).

- Signal definition (SG_i, i=1, 2, …., n)

- The definition of conflict matrix between signals (SG_i, SG_j), where i=1, …, n; j=1,2,…n

- Saturation flow for each signal (Qs_i, i=1, …, n, uv/h)

Furthermore, designing traffic data (usually peak hours) Q_i (uv/h) for each SG_i (i=1, 2, …, n) should be provided.

From the above junction geometry data and traffic data, the key movements (which determine the junction capacity) can be determined [(2) (5)]:

SGC_i (i=1, 2, …, k)

Then the junction capacity used (%) may be estimated

$$CU = (T_1+T_2+\ldots+T_k) /(CL- k*MIntG)$$

Where

CL is signal cycle length for the evaluation purpose (sec)

T_j is green needed (sec/cycle) for the key movement SGCj

$T_j=Q_j*CL/Qs_j$

MintG is mean intergreen between the key movements for the evaluation purpose (sec).

2.2 Transit greenable length and transit priority performance

A transit greenable length (GTmax in sec/cycle) is defined as the maximum green length may be distributed to a transit movement in one cycle without leading to traffic saturation at the junction. The transit performance ratio is then defined as

$$GTp = GTmax / CL$$

Case: the transit movement is a key movement: SGCm

Tm is green needed (sec/cycle) for the key movement SGCm

Because CU = $(T_1+T_2+\ldots+T_k)$ /(CL- k*MIntG)

Supposing Gc= $(T_1+T_2+\ldots+T_{m-1}+T_{m+1}\ldots+T_k)$

Then CU = (Gc+Tm) / (CL- k*MIntG)

In this case, the junction capacity used may be decreased if the green for the transit is not used in an actuated signalised junction control.

If CU <= 1 (junction is not saturated), the reserve of junction capacity can be attributed to the transit movement, i.e.

1 = (Gc+GTmax) /(CL- k*MIntG)

Then

$$GTmax = (CL- k*MIntG)-Gc = CL- (k*MIntG+Gc)$$

And

$$GTp = GTmax / CL = 1- (k*MIntG+Gc)/CL$$

Evidently, 0<GTp<=1, and a higher transit greenable length leads to a higher transit performance ratio

Case: the transit movement is not a key movement:

Supposing that

Its compatible key movements are SGd_j (j=1….p)

Its conflict key movements are SGC'$_n$ (n=1….q) , p+q=k

GC'= Total green needed (sec/cycle) for SGC'$_n$ (n=1….q)

GC"= Total green needed (sec/cycle) for SGd$_j$ (j=1….p)

Then

CU = (GC'+ GC") /(CL- k*MIntG)

In this case, the junction capacity used will not to be decreased if no transit arrives.

Supposing that CU <= 1 and the reserve of junction capacity is attributed to the transit movement

(GC'+GTmax) /(CL- (k-p+1)*MIntG) =1

Then

GTmax =(CL- (k-p+1)*MIntG)-GC'

GTp = GTmax / CL =1-((k-p+1)*MIntG+GC')/CL

Evidently, 0<GTp<=1, and a higher transit greenable length leads to a higher transit performance ratio.

2.3 Transit performance: mean transit delay

Mean transit delay is defined as TCd, in second. If the junction is saturated and the transit movement is saturated too (usually the case of transit mixed with private traffic), then the transit delay increases rapidly. This is the worst thing to the transit and it should be avoided. Only the situation where junction is not saturated, the transit mixed with few non-transit traffic or the transit has its own lane is discussed below.

Case: only single green for the transit movement in each cycle

Supposing that the transit arrives at junction randomly and the red length (including the transition period from green to red) is RL in sec, as the transit movement is not saturated, then

TCd$_s$ = (Probability of the transit arriving in red period) ∗ (0.5∗ red length)

= (RL / CL)∗ (0.5∗ RL)

$=0.5* RL^2/CL$

Or

$TCd_S = (RL / CL)* (0.5* RL)$

$= ((CL-GTmax) / CL)*(0.5*(CL-GTmax))$

$=0.5* ((CL-GTmax)^2 / CL)$

$=0.5* CL*((CL-GTmax)/CL))^2$

$=0.5* CL*(1- GTp)^2$

Therefore a higher transit performance ratio leads to a less mean transit delay.

If the transit arrives at junction regularly in each cycle, i.e. in the interval (t_0,t_1) $(0<= t_0< t_1<=CL)$, mean transit delay depends on the transit signal green interval (to be determined in the junction signal design).

Suppose that the green interval of the transit signal is (g_0, g_1) $(0<= g_0< g_1<=CL)$, Green interval length is $GL = g_1- g_0$.

Then, the minimum mean transit delay is

$$TCd_{smin} =0$$

if arriving interval is in its green interval, i.e.

$g_0<= t_0< t_1<= g_1$

The maximum mean delay is

$$TCd_{smax} = CL – GL – 0.5RL$$

if arriving interval begins at the end of its green interval and the two intervals have no a common interval.

Case: double greens for the transit movement in each cycle

Supposing that the transit arrives at junction randomly and the red lengths of two red intervals are RL_1 and RL_2 (in sec), then similarly

$$TCd_d = ((RL_1+RL_2) / CL)* ((RL_1/(RL_1+ RL_2)) *0.5* RL_1+(RL_2/(RL_1+ RL_2)) *0.5* RL_2)$$

$$= ((RL_1+ RL_2)/CL)* ((0.5*RL_1^2/(RL_1+ RL_2)) +(0.5*RL_2^2/(RL_1+ RL_2)))$$

$$= 0.5* (RL_1^2+ RL_2^2)/CL$$

As $RL= RL_1+ RL_2$, then

$$TCd_d =(0.5* RL^2 -RL*RL_2 + RL_2^2)/CL$$

And

$$d(TCd_d)/d(RL_2) = (-RL+2*RL_2)/CL$$

Suppose $d(TCd_d)/d(RL_2) =0$, then $RL_2= 0.5*RL$

Because $d^2(TCd_d)/d(RL_2) = 2/CL >0$, therefore

TCd_d arrives its minimum value in case $RL_1= RL_2$

As same as the case of single green, it can be shown

$$TCd_{dmin} =0$$

and

$$TCd_{dmax} = CL – GL -RLmin – 0.5*RLmax$$

where GL is transit total green length, RLmin= minimum (RL1, RL2) and RLmax = maximum (RL1, RL2)

As $RL= RL_1+ RL_2$ and RL= RLmin+ RLmax, then it can be shown

$$TCd_d = TCd_s-(RL_1+ RL_2)/CL$$

And

$$TCd_{dmax} = TCd_{smax}– 0.5*RLmin$$

Therefore, there is always less transit delays in the double greens situation than in the single green situation.

2.4 Evaluation of the transit lane concept

Transit lane concept is that at least one traffic lane is reserved for transit (bus and/or tram) only.

2.4.1 Transit lane concept, junction capacity used and transit delay

Here a comparison is made between the situation without transit lane and the situation with a transit lane.

Suppose there is a signalised junction, and two alternatives of the junction geometrical design are proposed: one is without transit lane; the other is with a transit lane.

A. Situation without transit lane:

-there are two vehicle lanes for a traffic movement (SG_j) and the transit movement is mixed in this traffic movement and

-the junction capacity used is CU_1.

B. Situation with a transit lane:

-One lane is reserved only for the transit movement only,

-the other lane is used by the traffic movement (SG_j), and then its saturation flow is decreased (50%).

-the junction capacity used is CU_2.

Case 1: If the traffic movement (SG_j)was a key movement in the situation A

Then it will still be a key movement in the situation B and the junction capacity used [(2)]

$$CU_2 > CU_1,$$

because T_j is increased as Qs_j is decreased in the situation B.

Case 2: If this traffic movement (SG_j) is not a key movement in the situation A

Then there are two possibilities:

1. SG_j is now a key movement in the situation B and then CU_2 will be increased [(2)].

$$CU_2 > CU_1$$

2. SG_j is not a key movement in the situation B and then CU_2 will be not changed (2).

$$CU_2 = CU_1$$

Therefore, a transit lane may decrease junction capacity and increase the junction capacity used.

In order to limit the impact of transit lane on junction capacity, following type of ‘partial transit lane’ may be used.

-transit lane disappeared at X meter before junction;

-transit lane starts at Y meter before junction;

-transit lane mixed with a traffic movement having a low traffic volume.

Transit lane is often used where a traffic entry control is needed to protect transit from traffic queues. As at this type of junction, decrease junction capacity is needed and make the entry control in a more nature way.

2.4.2 Transit lane concept and junction delay

If there is a transit lane, then the transit using this lane always passes the junction within one cycle. So the maximum mean delay for transit is limited to half of signal cycle length. While in the case without transit lane, the maximum mean delay for transit may be more than a signal cycle length (especially at oversaturated situation).

Therefore, if transit mixed with private traffic with high delay, the transit lane concept is possibly a good solution to protect transit from traffic queues, and hence to decrease maximum transit delay.

By comparing with no transit lane alternative, junction capacity used in a junction with a transit lane may be increased so the greenable length for transit may be decreased so the mean delay for transit may be increased.

As junction capacity used in a signalised junction with a transit lane alternative may be increased so the junction delay for non-transit movement who is conflict with the transit movement may be increased. While the delays for non-transit movements who are compatible with the transit movement may be decreased (3).

3. Junction signal design and transit priority treatment

3.1 The junction signal design

In a junction signal design step, based on traffic flow in peak hours, some junction signal design alternatives for a certain junction geometry design alternative may be developed. Each junction signal design alternative is consists of a matrix of intergreen, control strategies and one or more (fixed time and / or actuated) signal timing plans and / or actuated signal operation logics (with or without taking account of signal coordination with adjacent signalised junctions).

Concretely for each signal, a junction signal design gives each signal's green starting point and green end point, i.e. its green interval (g_0,g_1).

Traffic arrives at junction may be periodically (such as traffic arrives from adjacent fixed-time signalised junction), and its average arriving interval (t_0,t_1) ($0<= t_0< t_1<=CL$) covers only a part of signal cycle. Traffic arrives at junction may be randomly, it may be said that the average arriving interval (t_0,t_1), in this case, it covers entire signal cycle.

The idea signal design is that each signal green interval is nearly as same as its corresponding traffic arriving interval, so the junction traffic delay is minimised.

There are following types of junction signal design in the practice:

Cyclic signal design

There are still two subtypes:

-Fixed signal timing design: green distribution (start / end point of green in the cycle) is same for each cycle. It gives green to each signal periodically. This type of signal design is one of simplest signal designs.

-Non-fixed signal timing design: green distribution (start / end point of green in the cycle) isn't same for each cycle. It needs advanced knowledge about signal control and equipments to detect traffic and an intelligent controller.

Non cyclic signal design

This signal design is not used as often as other case because it may need advanced knowledge about signal control and equipments to detect traffic and an intelligent controller.

Cyclic signal design is usually suitable for the case where traffic arrives at junction periodically and regularly. Non cyclic signal design is usually suitable for the case where traffic arrives at junction randomly.

3.2 Transit priority treatment measures

There are following main transit priority treatment measures to decrease transit delay.

1. Distribution of reserved junction capacity to transit movement permanently (for fixed time signal design) or only if transit arrives at junction (semi / fully actuated signals) so

- Transit movement can have longer green length or

- Transit movement can have double greens in one cycle

2. Modification of the green start point based on arriving info of transit permanently (fixed time signal with signal coordination consideration) or temporarily only if transit arrives at junction (semi / fully actuated signals).

3. Distribution of reserved junction capacity to transit movement and modification of the green start point based on transit arriving info (1 & 2).

It is evident that if a transit movement has a higher transit performance ratio, then

-the transit can have more green or it may have possibility of double greens in one cycle,

-the green interval of transit signals is more easily arranged to coordinate with its corresponding arriving interval.

Therefore a higher transit priority performance ratio in a junction geometry design leads to a better transit priority performance in the signal design.

3.3 Transit priority treatment in a junction signal design

The consideration of transit priority treatment measures is different based on different type of signal design. In the following, the choice of transit priority treatment measures in different signal designs is discussed.

First of all, if a signalised junction is saturated, the sufficient green length to the transit movement should be guarantied to prevent the transit movement from traffic queue. It may produce more queues to no-transit traffic movements who have no enough green.

A. Case of cyclic signal design with fixed signal timing plan

If a signalised junction is not saturated, then the reserved junction capacity may be distributed to a transit movement permanently, so the transit signals can have more green length or double green and mean transit delays may be decreased.

Furthermore, if transit movements arrive at junction regularly, a signal timing plan developed by signal coordination design can minimise transit delays.

Evidently, a transit movement has a high transit performance ratio in the junction design leads to a better result for transit priority performance.

B. Case of cyclic signal design with non-fixed signal timing plan

Similarly, if junction is not saturated, the reserved junction capacity may be distributed to transit movements permanently, so the transit signals can have more green length or double green based on transit arriving info and mean transit delays may be decreased.

If a transit movement arrives at junction regularly, a signal timing plan developed by signal coordination design can decrease the transit delays. Furthermore, based on transit arriving info, the signal’s green start point may be advanced or the green end point may be extended so that transit delays may be minimised. The cost of the transit priority measures is that green length for some no-transit traffic may be shortened.

If only transit arriving randomly, a signal timing plan minimising non-transit delay by signal coordination should be found. Based on transit arriving info, the transit signal’s green open point may be any point of the cycle so that the transit delay may be minimised. In this case, some the delay for certain no-transit traffic may be increased and traffic queues may be produced.

If it is difficult to arrange the green to fit the transit arriving period, the double transit green is needed to improve transit priority performance.

Evidently, a transit movement has a high transit performance ratio in the junction design leads to a better result for transit priority performance in the signal design.

C. Case of non-cyclic signal design

In this type of signal design, junction signals operation is based on signal control logics. To minimise transit delay, the transit movement should be treated in this signal control logics as a movement with priority. For example, if a transit is arriving, all the conflicting movements to this transit movement should be closed (set signal to red) if possible and the next green should be given first to the announced transit movement.

In this case, a higher transit performance ratio leads to a better result for transit. But it can produce the best result (zero delay) even it has a low transit performance ratio under certain conditions.

In order to have zero transit delay, the arriving (at junction) info of transit movement should be known in advance, about 20 to 35 seconds depending on minimum green length and intergreen data. If the announced time length (arriving to junction) is not enough, the performance of transit priority will be limited.

The effectiveness of signal priority treatment measures may be tested by a micro-simulation model, such as VISSIM. Also, it may be verified by analysing the real signal recording data of the signalised junction operation (4).

3.4 Impact of transit priority treatment measures on non-transit movements

The impact of transit priority treatment measures on non-transit traffic may be qualitatively determined as follows.

As transit priority treatment measures aims to decrease the transit delay, hence delays may be increased to some non-transit movements who are conflict with the transit movement, and delays may be decreased to some non-transit movements who are compatible with the transit movement.

The level of impact depends on the frequency of using transit priority treatment measures. Evidently the frequency in turn depends on the number of bus / tram in the transit movement. Therefore, more bus / tram, the more impact on non-transit traffic.

For a junction with transit priority treatment measures, if each signal green’s length is known and the junction is not saturated, their impact on non-transit movements may be quantified. Take the advantage of micro-simulation model, such as VISSIM, the impact of signal priority treatment measures to non-transit traffic may be visualised and estimated by the simulation method.

In all the discussion above, all the transit priority treatment measures will not lead to junction saturation. In practice, this may be not satisfactory and a high level of transit priority measures (real green length for the transit > its greenable length) may be applied and it leads to junction saturation. It must pay attention to the impact in this case. Usually it is acceptable if oversaturated traffic does not produce traffic congestion to other neighbouring junctions.

4. Conclusions and further studies suggested

The author has proposed a method to evaluate quantitatively transit priority performance and its impact in a junction geometry design step and in a junction signal design step.

The transit greenable length and the transit priority performance ratio are the important measures to evaluate quantitatively transit priority performance in a junction geometry design step. A high transit greenable length leads to a higher transit priority performance ratio. Different formulas for estimation average transit delays are developed. It shows that average transit delays can be decreased if double greens can be given to the transit movement in place of a long single green in each cycle.

The analysis of transit priority performance in a junction geometry design with a transit lane concept shows that:

- Transit lane concept limits the maximum mean transit delays to the half of cycle length. So transit lane is a solution for the case where transit who has a high delay.

- Transit lane concept will usually decrease junction capacity and consequently increase delays to non-transit traffic and produces traffic queues especially if junction capacity used is more than 100%.

- Transit lane concept may decrease average transit delay but it does not always means to decrease average transit delays.

In the junction signal design, the sufficient green length to the transit movement should be guarantied to prevent the transit movement from traffic queue. Furthermore, the transit green intervals should be coordinated with its arriving intervals so that transit delay may be minimised. A transit movement has a high transit performance ratio in the junction geometry design leads to a better result in the junction signal design. Transit priority treatment may increase delays to some non-transit movements who are conflict with the transit movement, and decrease delays to some non-transit movements who are compatible with the transit movement.

In summary, a good junction geometry design taking account of transit priority performance will limit transit average delay. A good signal design combined with a good junction geometry design may considerately decrease transit delay and could limit its impact on non-transit traffic.

This paper mainly discusses the case of a signalised junction with only compatible transit movements and the impact of transit priority should not lead to the saturation of the junction. The other cases are needed to be studied in the future.

5. References

1. Kari J. Sane, City of Helsinki, Traffic Planning Division (1998) “Short Introduction to the Bus Priority Functions used in Helsinki”

2. J.Tan (2004) “Key movements and signalised junction’s capacity”

3. J.Tan (2005) “SJDesigner, A program for Signalised Junction’s Geometry Design and Signal Design”

4. J.Tan (2005) “SigRecAnalyser, A program to evaluate an intelligent signalised junction’s operation by analysing its real signal recording data”

5. R.Golay, Ph. Glayre and Ph. Bovy (2001) “La régulation des carrefours à feux”, Cahier n°11, TEA-ITEP, Swiss Federal Institute of Technology Lausanne

6. Chada, Shireen and Newland, Robert (2002) “Effectiveness of Bus Signal Priority”, Final Report, National Center For Transit Research (NCTR), University of South Florida

7. Chang, Collura, Dion, and Rakha (2003) “Evaluation of Service Reliability Impacts of Traffic Signal Priority Strategies for Bus Transit” Final report of ITS Center project: Signal Prioritization for the Center for ITS Implementation Research

8. Andrew Butler Nash,Ronald Sylvia (2000) “Implementation of Zuich’s Transit Priority Program”, Mineta Transportation Institute report-00-4