Working Paper 346

December 1991

QUEUE MANAGEMENT PROJECT: Strategies for the Management of Queues at Upstream Junctions

A D May

ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

This work was sponsored by the Science and Engineering Research Council

CONTENTS **Page**

1. INTRODUCTION

This Working Paper is one of a series representing work under a SERC grant on queue management strategies for urban traffic control systems, whose objectives are:-

(i)to generalise the strategies developed in an earlier study of queue management;
(ii)to develop a computer graphics based representation of queue propagation and management;
(iii)to test the strategies' applicability and performance in UK networks;
(iv)to investigate their incorporation into standard signal optimisation programs.

The study is based on earlier work in Bangkok (May et al, 1988) also funded by SERC, in which queue management measures were developed by trial and error to allow for the fact that queues from downstream junctions frequently blocked upstream junctions and could, as a result, unnecessarily obstruct crossing movements thereby reducing junction capacity and spreading congestion to other areas.

The methods developed in that study involved:-

(i)predicting the growth and decline of queues from downstream junctions;
(ii)estimating the speeds of the starting and stopping waves which determine (i);
(iii)identifying the most appropriate stage at the upstream junction in which queues should arrive;
(iv)identifying the most appropriate time in that stage during which queues should be present;
(v)adjusting the split between stages to allow for loss of throughput in blocked stages;
(vi)adjusting the cycle time as necessary in the light of (v).

It was steps (iii) and (iv) which involved the greatest element of trial and error, and it was accepted that further work was needed to identify alternative strategies for determining:-

(i)during which stage to allow queues to block an upstream junction;
(ii)how to adjust the stage timings at that junction to allow for the loss of capacity.

This paper reviews these approaches and makes recommendations. It considers first of all, in Section 2, the options at an individual junction. It then considers in Section 3, the application of these options in a range of increasingly complex networks.

2. STRATEGIES FOR AN INDIVIDUAL JUNCTION

2.1 Underlying Assumptions

It is assumed that it has been predicted that a queue from Junction A will extend to Junction B as shown in the time-space diagram Figure 2.1. The timing and duration of that queue in the junction will be determined by:-

(i)the start of red at A on the approach from B to A, t_r
(ii)the speed of the stopping wave from A to B, u_r
(iii)the start of green at A on the approach from B to A, t_g
(iv)the speed of the starting wave from A to B, u_g
(v)the length of the link from A to B (strictly from the stopline at A to the junction exit at B, L).

On this basis, the time of arrival of the queue at B, t_a, is given by

$$t_a = t_r + L/u_r \qquad (1)$$

and the time of clearance of the queue at B, t_c, is given by

$$t_c = t_g + L/u_g \qquad (2)$$

The speeds of the starting and stopping waves can be determined from observation or calculated by applying the principles of traffic flow theory. The first approach was adopted in Bangkok, where it was found that the starting wave speed, u_g, could be predicted far more reliably than the stopping wave speed, u_r (May et al, 1988).

The second approach has been adopted in the current study and in a related research student project (Al Madani, 1992), using the conventional Greenshields parabolic speed-flow relationship (Greenshields, 1934). That analysis has also shown that u_r is less reliably predicted than u_g, since it depends on the input flow, while u_g is determined by the saturation flow. However, the research has also shown that predictions based on Greenshields are themselves not very accurate; wave speeds are often disrupted by incidents on the link, such as pedestrians crossing, and they may well be lower than expected if further congestion, downstream of A, reduces the speed at which traffic crosses the stopline at A. This uncertainty is greater on the shorter links which are common in the UK than on the long (300-900m) links in Bangkok. It seems reasonable, therefore, to use equations (1) and (2), the Greenshields relationship, and reasonable estimates for free flow speed, u_f, and saturation flow, q_m to provide estimates of the earliest time, t_a and the latest time t_c, at which the queue will arrive at and clear B.

This analysis assumes that there is only one stage at A which produces queues on AB. Where left or right turns at A are signalled separately, this will not be the case. In the extreme, different queues may move upstream in different lanes, but usually they will combine to form one queue. In these situations the values for t_a, t_c can be approximated by adjusting t_r, t_g appropriately.

Similarly it assumes a uniform link in which the starting wave can move upstream at a uniform rate. Where the number of lanes varies, this will not be the case. However, values of t_a, t_c can be calculated by making appropriate adjustments to the value of L.

Finally, it is worth noting that, in closely spaced junctions or conditions of high saturation, queues from junctions downstream of A may arrive at B. This would occur during a green stage at A for the approach from B, but the starting wave speed would be determined by the saturation flow of the junction downstream of A.

Queues will be generated most rapidly when the input flow is highest. This implies that there is a case for arranging the offset at B such that queues propagated from A are fed by the minor input flows (usually the turning flows) as they approach B. However, this may well not be sustainable in practice. Unless it is possible fully to clear the queue before green is given to the main input flow, that input flow may well itself generate a queue which will arrive during that stage. Section 2.2 assumes that it is possible to choose when the queue arrives at B, and indicates those stages which are most appropriate. It needs to be borne in mind, however, that with certain distributions of input flow such choices will not be possible.

2.2 Appropriate Stages to be Blocked

2.2.1 General Principles

Earlier work in Bangkok demonstrated clearly that a queue should not exist in a junction during a change of stage. If it does, it is likely to encourage violation of the red signal, and aggravate blocking of the junction.

The choice of stage in which to accept a queue from downstream will depend on:-

(i)the movements which would as a result be blocked;
(ii)whether these movements are critical (ie. have the highest degree of saturation for the stage) or not;
(iii)the disruption which those movements could cause while blocked.

As general principles:-

(a)where possible, queues should be introduced when they only block non-critical movements;
(b)where critical movements have to be blocked, it is generally preferable to block those movements which contribute least to the generation of the queue;
(c)however, it is also preferable to avoid blocking movements which will in turn trap other movements in the junction.

These principles are illustrated for the general case of a crossroads between two-way roads NS and EW in Figure 2.2. Junction A is taken as being in direction N from Junction B.

2.2.2 Two Stage Cycle (Figure 2.2(a))

Blocked Movements. If the queue arrived during Stage I, it would block movements SN and NW. It could also block SE and SW if separate lanes are not provided for those movements. If the queue arrived during Stage II, it would block those movements which directly fed it (WN and EN) and this in turn would almost certainly block WE. Movements EW, ES and WS could be blocked if they had no separate approach lanes.

Critical Movements. If SN is not the critical movement (although this is unlikely), it is better to accommodate the queue during Stage I. If none of WE, WN and EN is critical, it is better to block Stage II. If both SN and one of WE, WN and EN are critical, the choice is less clear-cut, since the queue blocks the critical movement in both cases. It will generally be preferable to block those movements which contribute least to the generation of the queue; in most cases these will be WN and EN.

Disruption. Blocking of Stage I will leave movement NW in the junction, partially blocking other movements. Blocking of Stage II will leave movements EN and EW in the junction; this will in turn block movement WE. The relative impact of these will depend on the time taken for the disrupted movements to clear, and hence on their magnitude.

In this situation, therefore:-

(a)if NS is the critical movement in Stage I, the junction should be blocked in Stage I
(b)if SN is the critical movement in Stage I, the junction should be blocked in Stage II unless:-
  (i)movements WN and EN generate the queue more rapidly than SN (which will rarely be the case); or
  (ii)movement EN is substantial enough to cause clearance problems for WE.

Simpler cases of junction layout can be treated by applying these same principles:

(i)if NS does not exist (ie. for a one way road) then it will be better to block Stage II, except in cases (b)(i), (ii) above;
(ii)if WE does not exist (ie. for a one way road) then it is better to block Stage II, unless NS is the critical movement;
(iii)if arm S does not exist, it is better to block Stage I, in which the queue is not fed (indeed, blocking is unlikely to arise in such circumstances);
(iv)if arm E or arm W does not exist, it is better to block Stage II, unless condition (b)(i) above applies.

2.2.3 Separate Right Turn Stages (Figure 2.2(b))

Blocked Movements. If the queue arrives in Stage I, movement SN is blocked, and movement SW may be, if it does not have a separate lane. If it arrives in Stage II, movement WN is blocked; movement WE may be. No movements in Stage III would be blocked. In Stage IV, movement EN would be blocked. On this basis, it appears preferable to time the arrival of the queue in Stage III. However, this is unlikely to be sustainable since arrival of the queue will be determined by the stages which primarily feed it, which will be Stages I, II and IV.

Critical Movements. It is again best to accommodate the queue in Stage I if NS is the critical movement. Similarly, it will be better to accommodate it in Stage II if EW is the critical movement and in Stage IV if WS is the critical movement. If SN and EN are both critical movements, then the best stage will be Stage II, unless WN is also critical.

Disruption. In terms of disruption to the junction, only blocking of Stage IV will cause the junction itself to be blocked, by the queue from EN; this should be avoided if possible.

In this situation, therefore:-

(i)if NS is the critical movement in Stage I, the junction should be blocked in Stage I;
(ii)if SN is the critical movement in Stage I and WS is the critical movement in Stage IV, the junction should be blocked in Stage IV, but care should be taken to avoid blocking by EN;
(iii)in other cases Stage II should be blocked, except where WN generates the queue more rapidly than SN; in this case Stage I should be blocked.

2.2.4 Separate Stages for Each Movement (Figure 2.2(c))

Blocked Movements. If the queue arrives in Stage I, movement SN is blocked, and movements SE and SW may be affected if they have not got separate lanes. If it arrives in Stage II, no movements are blocked, but this is not a sustainable situation. If it arrives in Stage III, movement WN is blocked and movement WE may potentially be affected. Similarly, if it arrives in Stage IV movement EN will be affected, and movement EW may be.

Critical Movements. It is preferable to block Stage III, unless WN is its critical movement. Where all movements WN, SN and EN are critical, it is best to block that movement which generates the queue least rapidly.

Disruption. In terms of disruption only movement EN will cause problems if blocked, by blocking other movements through the junction until it has cleared.

In this situation, therefore:-

(a)if WE (or WS) is the critical movement in Stage III, Stage III should be blocked;
(b)if WN is the critical movement in Stage III and EW (or ES) is the critical movement in Stage IV, Stage IV should be blocked, but care should be taken to ensure that movement EN does not later block the junction;
(c)in other cases (where WN, SN and EN are all critical movements), that stage which generates the queue least rapidly should ideally be blocked.

2.3 Timing of the Arrival of the Queue in the Blocked Stage

The main consideration in determining when the queue should arrive during a given stage must be the need to avoid inadvertent blocking during the stage change. This may occur if the queue is timed to arrive early in the stage, and the stopping wave arrives earlier than expected. Equally, it may occur if the queue is timed to arrive late in the stage, and the starting wave arrives later than expected.

As noted earlier, experience in Bangkok indicated that the speed of the stopping wave was much more variable than that of the starting wave. A sudden surge of input traffic or poor packing in the queue, could lead to rapid propagation of the stopping wave, which would thus arrive earlier than expected. Conversely, late arrival of starting waves will only occur when vehicles in the queue fail to maintain saturation flow. Thus generally it is better to allow the queue to arrive as late as possible in the stage, while allowing a margin of error for the prediction of starting wave speeds. In the Bangkok experiments, this was based on two standard deviations in excess of the expected travel time for the wave from A to B.

Such an arrangement is also likely to avoid undue disruption in situations in which saturation flow falls later in the cycle, or where vehicles waiting to make turning movements in the upstream link cause a reduction in saturation flow.

2.4Adjustment of Split and Cycle Time

The process for adjustment of the split between stages is in principle relatively straightforward. The time for which the stage is blocked is first estimated and this is considered as adding to the lost time in the cycle. The remaining time is then allocated optimally between the stages, and the blocked time added to the length of the blocked stage. However, there are several variants of this process which need to be borne in mind.

In some cases in Section 2.2, blocking can be arranged so that it is not the critical movement in that stage which is blocked. In these situations, an initial calculation is required to assess whether the blocked movement, with its movement time reduced by blocking, becomes critical. If it does not, then no adjustment to the split is needed. If it does, then it is necessary to calculate the amount by which that stage would need to be extended to make the blocked movement no more critical than other movements. This is then the equivalent blocked period.

In a limited number of cases, as noted in Section 2.1, more than one queue will arrive at B per cycle. Where this happens, two responses are possible. One is straightforwardly to add both blocked periods to the lost time, reoptimise the split and add the blocked periods to the length of the two blocked stages. However, it may be that only one stage is suitable for blocking. In this case, it may be possible to double cycle the signal at B, so that its cycle length is half that at A, and both queues per cycle at A arrive in the same stage at B. This technique was applied successfully in Bangkok. It does, however, have implications for capacity at B, as discussed below, and for the performance of junctions further upstream, if they are affected by queues from B.

In certain circumstances, it may be appropriate to modify the stages included in the cycle or their sequence. As Section 2.2 indicates, certain stages in multi-stage cycles may be less affected by blocking. However, care needs to be taken, if adding to the number of stages, that the additional lost time does not critically reduce the capacity of that cycle.

The final situation arises where the loss of time from blocking during the cycle renders the cycle over-saturated. It should be noted that shorter cycles are generally to be preferred since they generate shorter queues. However, they also have a higher proportion of lost time; this may be a particular problem with double-cycling, as suggested above. In these situations, it will be necessary either:-

(i)to reduce the cycle time at A (and other downstream junctions which are generating blocking queues), to reduce queue length or, where this is not feasible on capacity grounds:-

(ii)to lengthen the cycle time of upstream junctions.

It will, of course, be necessary to have a common cycle time throughout, or common multiples of that cycle time. Where such adjustments result in the cycle time of A being extended, this will increase the queue length from A, and require further adjustments at B.

3. APPLICATION TO NETWORKS

3.1 The Selected Networks

In this section, the principles outlined in Section 2 are applied to three networks:-

(i)the linear network used in the Leeds experiment on Wellington Street;
(ii)the ten junction two dimensional network used in the original experiments in Bangkok;
(iii)a six junction two dimensional network, which represents a simplification of the Park Row/East Parade network in Leeds.

3.1.1 Wellington Street, Leeds

This network is shown in Figure 3.1, together with the flows and turning proportions observed in the September 1990 survey. Queues regularly form from junction A through junction B and occasionally junction C. Output from junction A is rarely disrupted, thus this can be considered the critical junction. Junction B is a simple crossroads, but with no crossing movements, and operates on a simple two stage cycle. Junction C also operates on a two stage cycle, but all movements are permitted. The critical movements at both junctions B and C are those in the direction of A.

3.1.2 Bamrang Muang Road, Bangkok

This network is shown in Figure 3.2. Queues regularly form from junction 1 through junctions 2 and 3; these occasionally reach junction 4, while side street queues occasionally reach junctions 9 and 8. Output flows at junction 1 are rarely disrupted, thus this can be considered the critical junction. Junction 2 simply splits the feed to junction 1, and input flows are similar in magnitude from the two approaches. Junction 3 provides a further split in the input flow, and also the potential for the WS movement to be disrupted by the NE one. Junction 4 is the first crossroads, with two way traffic NS. Junctions 5, 6, 7 and 10 are rarely affected by serious queues in the pm peak which was being studied in Bangkok. However, some problems arise at junction 9 as a result of queues blocking back on the short link to junction 8.

Junctions 2 and 3 operate on two stages. Junctions 4 and 5 have three stages, for traffic from W, N+S and S. Junction 6 has two stages. Junctions 7 and 8 have three stages for traffic from E, N+S and N. Junction 9 has three stages for traffic from E, SE, S. Junction 10 has three stages for traffic from N, SW, E.

3.1.3 Park Row/East Parade, Leeds

This network is shown in Figure 3.3. It differs from reality in one important respect. Junction 2 is in practice a priority junction, with priority for the WS bus movement. It was this which rendered it impossible to test strategies on the real network. Queues regularly form from junction 1 through junction 2 and on to junctions 3, 4 and sometimes 5, exacerbated by the loss of priority at junction 2. Queues also form from junction 5, but rarely affect junction 6. Output flows from junction 1 are sometimes restricted by queues from the east and south. Even so, this can be considered to be the critical junction.

Junction 2, as signalised, would be a simple two stage junction, each with one critical movement. Junction 3 operates similarly. Junction 4 operates as a two stage cycle, separately for movements from the west and east. WS and ES are the critical movements. Junction 5 operates on a two stage cycle, with SN movements one way only. WE and SN are the critical movements. Junction 6 also operates on a two stage cycle, for a crossing of two one-way movements. The critical movements are SN and EN. Junction 7 operates on two stages, one of which is simply for the WE bus movements. WE and ENW are the critical movements. Junction 8 operates on two stages, for movements from the west and east separately. WN and EW are the critical movements.

3.2 Wellington Street, Leeds

3.2.1 Junction 2

The principles in Section 2.2.2 indicate that the queue is best introduced into the junction during the stage for movements from S and N, since:-

(a)EW is the critical movement in Stage I;
(b)SW and NW contribute far less flow than EW;
(c)SN is not possible, and thus cannot be blocked by NW.

3.2.2 Junction 3

Similar arguments apply to junction 3; while SN is a possible movement, NW is a relatively minor movement and unlikely to disrupt SN, which is itself also minor. Thus the principles in Section 2.2.2 suggest that the queue is, once again, best introduced into the junction during the stage for movements from S and N.

3.2.3 Summary

The above principles suggest that, for Wellington Street, the queues are best stored in the side streets, provided that these have sufficient demand to sustain them. This is, in practice, the basis of Strategy 3, initially developed by Quinn and Topp (1991). Model tests of this strategy (Clark and Montgomery, 1991) have since demonstrated that it performs most successfully.

3.3 Bamrang Muang, Bangkok

3.3.1 Junction 2

The principles in Section 2 do not apply to this simple case. Since each stage contains only one movement, the queue can be accommodated in either; the sole consideration is queue storage. In practice queues are likely to extend to either junction 3 or junction 9, and are less likely to be disruptive in junction 3, where there are fewer movements. This also avoids further disruption to junction 9, which is also affected by junction 8.

3.3.2 Junction 3

The principles in section 2.2.2 suggest that it is better to introduce the queue into the stage when the movement from S is running, since:-

(a)EW does not exist;
(b)WE generates the queue more rapidly than SE;
(c)NS does not exist to be blocked by SE.

3.3.3 Junction 4

If the queue from junction 1 is stored to the South of junction 3, it may not be necessary to provide for queues in junction 4. If they do arise, the principles of Section 2.2.2 suggest that the queue is probably better accommodated in the stage for NS movements, or for movements from S, since:-

(a)EW does not exist;
(b)WE is the main contributor to the queue.

However, the fact that movements into E are split into three stages may mean that neither SE nor NE develops the queue sufficiently. On this basis, it may be better to introduce the queue into the WE stage.

3.3.4 Junction 9

It is not necessary to consider queues at junctions 5, 6, 7, 8 or 10, but queues from junction 8 do occasionally disrupt junction 9. Applying the principles of section 2.2.2 again suggests that it is better to introduce the queue in the movement from S, since:-

(a) WE does not exist;
(b) NW does not exist.

However, the movement from S is in practice split between S and SE, and neither may be sufficient on its own to generate a queue effectively. In this case, the queue would need to be introduced when traffic from E was moving.

3.3.5 Summary

The principles evolved above are somewhat different from those tested in Bangkok. Arrangements at junction 2 were similar, but it was found at junction 3 that queues arrived both from junction 1 and from junction 2. This was accommodated by double cycling junction 3, and accepting the two queues during the movements from W, rather than from S as recommended above. This in turn required junction 4 to be double cycled, with the queue accommodated in the movement from W.

3.4 Park Row/East Parade, Leeds

3.4.1 Junction 2

Because the stage from W only contains buses, it is only really feasible to accommodate the queue from junction 1 when traffic from N is running.

3.4.2 Junction 3

Junction 3 simply splits the feed, and the principles in Section 2 cannot readily be applied. However, the link from junction 4 is very short, and queues are therefore probably best accommodated when the stage from W is running.

3.4.3 Junction 4

Queues from junction 3 will almost certainly spill back to junction 4. This junction can best be treated using the principles of section 2.3. On this basis, the stage for traffic from E should be blocked. Both ES and WS are critical, but ES generates the queue less rapidly.

3.4.4 Junction 5

Following the principles of Section 2.2, the stage from S should be blocked, since:-

(a) WE is critical in its stage;
(b) NS does not exist.

However, this implies that two sets of queues will be generated along the links from junction 6 (from junctions 2 and 3 and from junctions 4 and 5). This may introduce some problems north of junction 6.

3.4.5 Summary

The principles outlined above are markedly different from current practice in their treatment of queues at junctions 3, 4 and 5, where timings encourage the queue to develop from junction 2 through junctions 3, 4 and 5 and on to the west. It would be interesting to test their treatment as suggested above, although as noted earlier the existence of give way markings at junction 2 makes this difficult.

4. REFERENCES

Al-Madani, H.M.N. (1992) the prediction of the location of head and tail of queues between two oversaturated signals. Proc 24th UTSG Conference, Newcastle

Clark, S.D. and F.O. Montgomery (1991) Queue management project: calibration of the traffic model. WP347, Institute for Transport Studies

Greenshields, B.D. (1934) A study of traffic capacity. Proc. Highways Research Board, Vol.14. HRB, Washington.

May, A.D., F.O. Montgomery and D.J. Quinn (1988). Control of congestion in highly saturated networks. Proc. IVth Conference in Urban Transport in Developing Countries. Jakarta, June 1988. CODATU, Paris

Quinn, D.J. and C.F.E Topp (1991) Queue management project: testing strategies using the model. WP332, Institute for Transport Studies

ADM\WP346.LIZ (15/01/92)