Bus priority in Greater London

6. BUS BUNCHING AND REGULARITY OF SERVICE

by K. W. Huddart, MA, CEng, FIEE, MInstP
Assistant Chief Engineer, Traffic Operations
Department of Planning and Transportation
Greater London Council

The traffic congestion, which is being tackled by bus priority measures, is alleged to produce not only delay to bus services, but also irregularity in their running. Intuitively this is an obvious enough link, but more careful thought shows that an ideally controlled system could recover, in regularity terms, from the effects of traffic congestion. On the other hand, irregularity may be caused by a shortage of crews and vehicles or by lack of flexibility in their allocation to different services. Similarly small perturbations in the service, not necessarily caused by traffic congestion, may be magnified alarmingly when the effects of passenger loading time are taken into account, and this article is restricted to examination of this instability.

Although bus bunching is very much a matter for bus operators, and hence this article is irrelevant to the author's duties, it is clearly of interest to traffic engineers to understand something of how the operational side relates to their efforts on the highway. It is, therefore, a worthwhile topic for the final article in this series on bus priorities in Greater London.

PASSENGER WAITING TIME

The basic relationship. The time that passengers have to wait for buses depends not only on the average service interval (i) between buses, but also on the evenness of the spacing of buses and on the evenness, or randomness, of passenger arrivals. For many London services with close headways there is no published timetable, so it is assumed that passengers arrive steadily or at random. In this case the average waiting time (Average WT) is $\frac{1}{2}i$ because passengers on average have to wait for half the service interval before a bus arrives. This value of $\frac{1}{2}i$ is taken as a basis for comparison of other situations, and called Expected WT.

Excess waiting time. If the bus service is also random[1, 2], or if it is uniform with all buses bunched into pairs, it can be shown that the Average WT rises to i, and there is said to be an 'Excess WT' of $\frac{1}{2}i$. The only ways of achieving an even greater Average WT, such as has been observed in London[2], are for the bunches of buses themselves to move randomly, or to include more than two buses in each bunch. London Transport have examined the more complex situation with deliberately scheduled variations in service interval[3].

Published schedules. If a service is sufficiently infrequent but reliable for its schedule to be published, passengers can take advantage of this information to arrive at the loading point just before the bus is due to depart. This situation is rare for urban bus services because of their unreliability, but Fig 1 shows a typical arrival pattern at a suburban railway station, from which it is seen that an Average WT of much less than $\frac{1}{2}i$ can be achieved.

Effect on bus delay. In assessing a practical situation it is often useful to relate delay of an individual bus to the Average WT. Figure 2 shows the situation if the service is running with alternate buses late by a time d. In this case it can be shown that:

$$\text{Excess WT} = \tfrac{1}{2}d^2/i \qquad \text{......(1)}$$

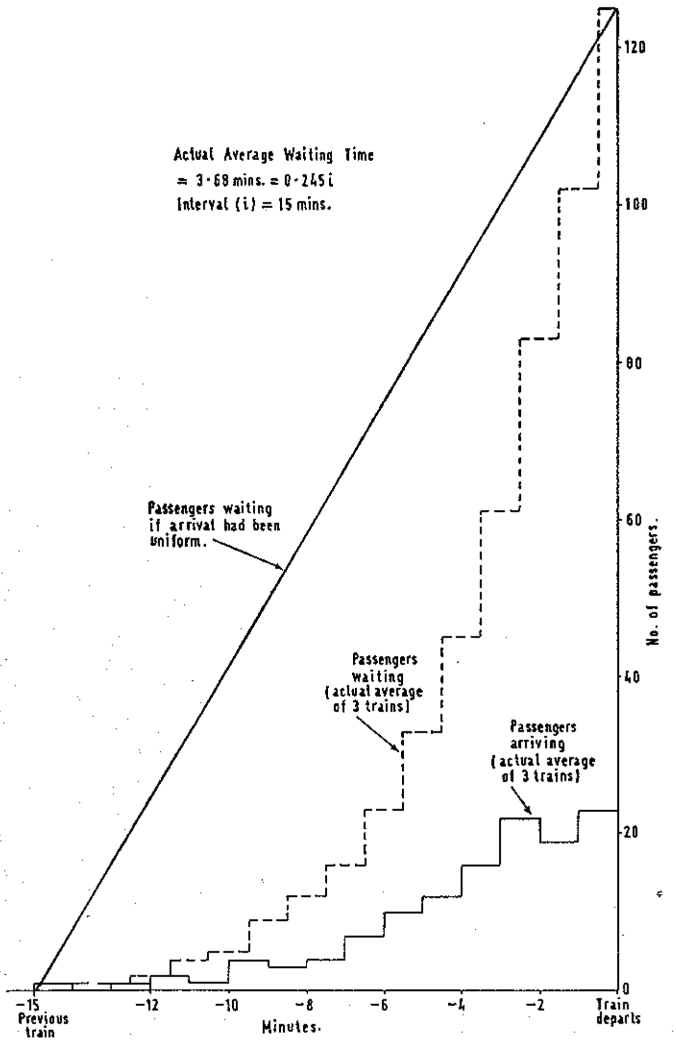
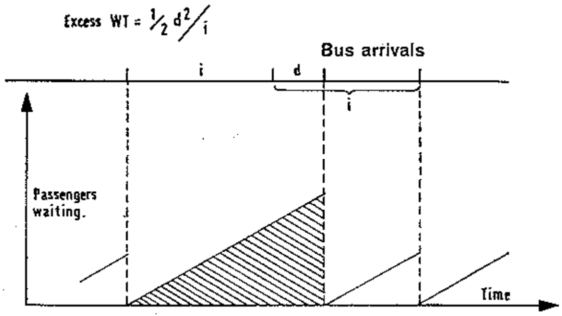


Fig 1 (left). Passengers waiting for a suburban train.

Fig 2 (below). More passengers (shaded) have an extra wait.



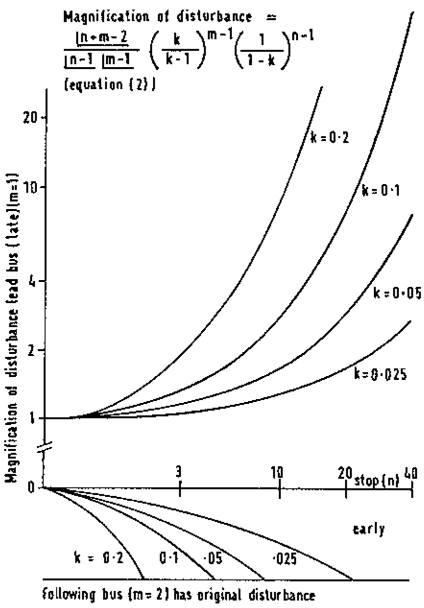


Fig 3. Growth of bunching disturbance.

THE DYNAMICS OF BUNCHING

Fortunately the basic mathematics have been done[4] and the result is shown in Equation (2) in Fig 3. The analysis supposes that passengers arrive regularly at a series of n stops which are serviced by a series of m buses. If for any reason one bus is slightly late at the first stop, it finds that more passengers have arrived than usual and the bus is further delayed by the time taken for these passengers to board. The initial perturbation (delay) is therefore magnified as the bus travels from stop to stop and gets further and further delayed. At the first stop, the next bus arriving on schedule finds that there are less passengers than usual, so it takes less time than usual for the passengers to board and this bus then starts to run early. The analysis underlines the importance of the quantity k which is (passenger arrival rate)/(passenger boarding rate). Typically k is less than 0·05, but an example has been measured in London of a value of 0·5 which, as can be expected, had a dramatic impact on the service. It is sometimes necessary to include an effect for passenger alighting time, particularly for buses with only one door each, when one is trying to study the total delay accumulated along a considerable length (above 3 km) of route. This effect is likely to be less than a doubling of k because passengers alight more quickly than they board, and because they alight later in the journey of the bus and therefore there is less scope for the effect to be multiplied at subsequent stops.

CONSEQUENCES OF BUNCHING

Instability at a single stop. Normally the term in k/(k-l) is small, so that any perturbation gets less with subsequent buses. If, however, k exceeds 0·5 this term exceeds l, so that the initial perturbation is magnified with subsequent buses even at a single stop. Such a situation occurs if, at a stop, the buses of one route (or group of routes) are boarding or alighting passengers at that stop for more than half the time. The situation is particularly likely to occur on routes like the Red Arrow in which large buses on a frequent service are completely filled at a rail interchange. It should be noted that the term k/(k-l) is negative so that alternate buses will become late and early, giving rise to bunching.

The effect on the initial bus. This effect is described by the term in l/(l-k) which is normally slightly greater than l. Eventually, when n becomes large, this term causes the perturbation to increase enormously. Figure 3 shows the magnification factor for different values of k—firstly, it is worth noting how many stops it takes to double the initial perturbation; and secondly, the magnification of the purturbation after 40 stops may easily be by a factor of more than 10.

A typical route. When considering the cumulative effect of a perturbation it is clearly necessary to know whether there are sufficient stops in the route for the effect to continue. In this context a route must be regarded as continuous if there is no opportunity to readjust the timing of the buses, even though there may be tea intervals of fixed duration interposed. A typical inner London route may be 15 km long having 50 stops (only 40 being used much), 2·5 passengers boarding at each stop, and k of 0·03. At the end of this route section taking about one hour, the bus turns without waiting making a double section of two hours; the day comprises four such double sections interspersed with two tea breaks and a meal break. Thus, depending on the flexibility of timing at the turns and breaks, a route may contain effectively between 40 and 320 stops. From Fig 3 it is clear that the earliest possible opportunity should be taken to introduce some flexibility, otherwise any variation of travel time can be magnified and grow out of control.

Effect on following buses. Figure 3 also shows, for different values of k, how many stops it takes for the perturbation to be reproduced in the same magnitude in the following bus. In this case it is important to notice that the following bus tends to run early which is an effect which may be easier to prevent. However, if the service as a whole is running late and hence out of control, it is seen that, even in a route section of only 40 stops, several buses will be seriously affected.

A practical measurement. Excess WT has been measured[5] for the contra-flow bus lane in Tottenham High Road. On three routes measured at the beginning of the bus lane Excess WT ranged from 0·88 to 1·62 mins. This can be converted to individual bus delay, d, by Equation (1) and then, assuming k of 0·05 applies at three stops, a new delay and Excess WT can be calculated. The calculation shows that Excess WT should increase along the bus lane by some 0·32-0·58 mins., compared with values actually measured of 0·20-0·28 mins. There is thus some evidence that the analysis gives answers of the right order.

Effect of one-man operation. The typical k of 0·03 applies to double-deck crew-operated buses without doors. Passenger loading rates can be improved, and k reduced by providing two openings to a bus, but k is greatly increased for a bus with one-man operation, particularly if this man has to deal with all the fares. In the extreme with such a bus having a capacity of 90 people, k could well rise to 0·13, which is clearly unacceptable. Therefore if such poor loading rates are to be used, it is essential that the passenger arrival rate should be kept down; whether this is by the use of smaller buses or longer service intervals, some other route or service must be provided for the residual passengers. The situation is helped if the average distance which passengers travel is greater.

SOLUTIONS

The analysis above could well be considered as a description of complete chaos. But the system does, to an extent, work and must therefore have some effective controls or crew reactions incorporated in it. It is worth examining some of these 'solutions'.

Scheduling. The traditional method of controlling buses is to provide a timetable, even if this timetable is not publicised. Provided that generally the buses are running to the timetable, it is possible to prevent any effects transferring to subsequent buses, because any bus can be prevented from running early. One is still left with the large magnification of delays on the initial bus which is affected. This also can be overcome if there is sufficient slack in the schedule for the bus to be able to keep to schedule even though it is delayed by having to pick up additional passengers. The least possible slack is required if the delay is corrected immediately. The slack must then be sufficient to eliminate the amount by which the term l/(l-k) of Equation (2) exceeds unity. This allowance has to be built in for every stop, and the total for a route section of 40 stops is shown in Fig 4. The total slack schedule shown for the route section is the number of additional minutes that have to be allowed if one is to deal with an initial perturbation of one minute, e.g. missing the phase at a traffic signal.

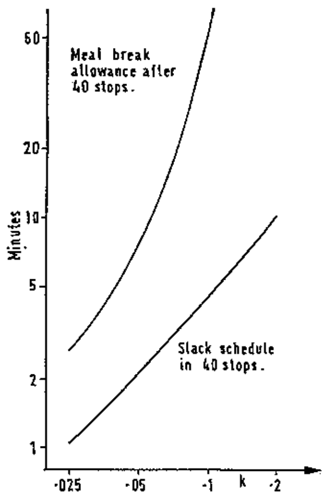


Fig 4. Alternative ways of scheduling for one minute disturbance.

Traffic congestion is neutral in its effect on the dynamics of bus bunching, but destroys the ability of scheduling to control operation. The tendency to reduce the slack in scheduling therefore makes services more susceptible to traffic congestion.

Flexible meal breaks. If the duration of meal breaks is increased by an amount which can be varied to suit operational needs, it is possible to correct any bunching before commencing the next route section. The extra time that has to be allowed at the end of each route section of 40 stops is also shown in Fig 4. It is noted that this time is always greater than the allowance within the schedule, because remedial action has been deferred to the end of the route section and there has been some opportunity for the bunching to increase. However, time inserted at a turning point or during a meal break affects only the operating costs and does not adversely affect passengers or produce annoying slow-running. As the cost of time to passengers is many times the operating cost (including crew time) of the vehicle, it seems that, for values of k below 0·05 the most efficient control would be by having a tight schedule, and providing the flexibility in turning time and meal breaks.

Buses filled to capacity. Once a bus is completely full the time it takes for further passengers to board is determined entirely by the number of passengers that alight, and not by the headway between it and the preceding bus. Therefore the dynamic deterioration of the service is arrested, but if the full bus has to set down many passengers it may be incapable of running at the normal scheduled speed so that there is some further delay. Furthermore, the price of this type of control is that a number of passengers are left at bus stops and have to wait for a later bus.

Turning short. It is well-known that bunching can be controlled by turning buses short of their normal destinations and sending them back on the return route. If this is done by taking a bus out of a bunch and arranging for it to occur during a large gap on the return route it is clear that a considerable improvement in reliability can be achieved. The difficulties are that:

(*a*) the turning short must be anticipated long enough ahead to change the destination board on the bus;

(*b*) it may be difficult to provide the crews with appropriate duty shifts which end at the correct garage for them to go home; and

(*c*) the likely behaviour of the service in the future may be adversely affected by turning short.

In practice the advantageous effect of turning short is more likely to be given by those turns which are already included in the schedule. The effect of this technique is clearly visible on a Cathode Ray Tube display of the simulation of a bus route available at the Marconi Company[6]. The mechanics of a simple situation are shown in Fig 5 in which alternate buses are turned from a route in which all the buses were exactly paired. Turning one of the buses from A from each pair leaves the route at B with a completely uniform service at twice the service interval. The turned bus will arrive in the reverse stream a time e after one of the buses from D, which is also assumed to be a uniform service at interval 2i.

Depending on the value of e, the resultant at C may be anything between a perfectly uniform service and a completely paired service. On average, the Average WT at C is $\frac{2}{3}$i, with an Excess WT of $\frac{1}{6}$i. The turning operation thereby reduces the Average WT of sections A and D from 2i with an Excess WT of $\frac{1}{2}$i to an Average WT of sections B and C of $1\frac{2}{3}$i with an Excess WT of $\frac{1}{6}$i. There is thus a saving of WT of $\frac{1}{3}$i.

Paired working. When buses are so tightly bunched that they are within sight of one another, the lead bus can deliberately leave collection of passengers at a stop to the following bus. Similarly, buses may overtake each other; if a bus calls at alternate stops, the value of k is effectively halved.

FURTHER WORK

Better control systems. It is clearly unsatisfactory to rely on scheduling to control buses, when such scheduling can easily be destroyed by relatively minor delays in traffic and for other reasons. To control bunching of buses a suggested alternative is 'headway control'. In the absence of a central communication facility such control can be done only by providing clocks at stops so that each driver can see his own headway and attempt to adjust it to a specified value. The difficulty is that the only action he can take is to go slower and this action is applicable not to the bus which suffered the initial delay but to the following bus, which is shown in Fig 3 to have the lesser problem. Clearly such headway control could be of some assistance in preventing a perturbation spreading throughout the entire system, but what is really needed is a means of delaying the bus preceding that which has suffered the unexpected delay. This is something which should now be possible with radio communication to every bus on an experimental route, complete with a location system, and is certainly something which should be simulated and then attempted in practice.

Further analysis. On account of the complexity of the problem, digital simulation would seem to be a helpful method of analysis provided it is directed to relatively simple and easily understood situations such as those that have been discussed in this article. It is necessary to build in more economic features so that it is possible to compare the value of time spent by passengers travelling and waiting, and the savings in operating costs which could be achieved by providing a better service with fewer vehicles, although these vehicles had appropriate flexibility built into their schedules. It is also necessary to establish whether a new method of control such as 'headway control' would enable crews to have a more predictable working day than that arising at present with the substantial instabilities in bus running.

Provision of spare vehicles. The perturbations which have been considered in this article have been small ones such as might have been incurred by missing a single cycle at a traffic signal. In practice much larger perturbations can be introduced by the total omission of a bus from a route. This is particularly severe on infrequent services where passengers hope to use their knowledge of the published timetable to reduce waiting time. It could well be that a higher level of service could be provided to passengers, and hence the service would be more attractive and attract more passengers, if less buses were introduced into the schedule, and some were kept deliberately available as spares to make good any deficiencies that might occur.

CONCLUSION

This series of articles has shown the progress and value of an approach of seeking out situations where bus priorities are needed, and then implementing these measures. This approach depends on meticulous examination of individual schemes, rather than arbitrary implementation of a total network, many parts of which would be unnecessary and cause hardship to other users. Bus priorities cannot solve all problems, so the system of supply and control of services must be developed to improve its stability against traffic delays, crew shortage and other effects.

The views expressed in this article are those of the author and are not necessarily those of the GLC.

REFERENCES

[1]O'Flaherty, C. A. and D. O. Mangan. Bus passenger waiting times in central areas. *Traff. Engng Control,* 11 (9), January, 1970, 419-421.

[2]Holroyd, E. M. and D. A. Scraggs. Waiting time for buses in central London. *Traff. Engng Control,* 8 (3), July 1966, 158-160.

[3]Oliver, A. M. The assessment of Marconi location equipment on route 11 and radio control on routes 30 and 76: methodology. Department of Operational Research, London Transport, TN 49, 1972.

[4]Newell, G. F. and R. B. Potts. Maintaining a bus schedule. *Proc. 2nd Conf. Aust. Rd Res. Bd,* 1964, 388-393.

[5]Department of the Environment, Bus Demonstration Project, Summary Report No. 2 Tottenham, 1972.

[6]Winfield, R. C. Omnibus operation and control simulator. The Marconi Co. Ltd, Report Y/248/5676, April 1972.

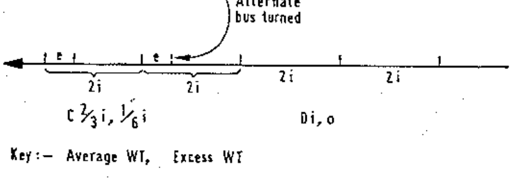


Fig 5. Turning buses decreases passenger waiting time.