TRANSPORT RESEARCH LABORATORY

TRL REPORT 240

UTC STRATEGIES FOR CONGESTED NETWORKS

by K Wood, JP Palmer, RD Bretherton and S Shearn

Prepared for: **Driver Information and Traffic Management Division, DOT**
Project: **UTC Strategies in Congested Networks (NMAM112)**

Transport Research Foundation Group of Companies
Transport Research Foundation (a company limited by guarantee) trading as Transport Research Laboratory. Registered in England, Number 3011746.
TRL Limited. Registered in England, Number 3142272. Registered Offices: Old Wokingham Road, Crowthorne, Berkshire, RG45 6AU.

First Published 1997
ISSN 0968-4107

CONTENTS

EXECUTIVE SUMMARY

The Transport Research Laboratory has undertaken a research project on Urban Traffic Control (UTC) strategies for congested conditions for the Driver Information and Traffic Management Division of the Department of Transport. The aim of the project is to produce strategies to alleviate the effects of congestion in large urban areas.

SCOOT is widely used in the UK, and elsewhere in the world. From Version 2.4 onwards, SCOOT has incorporated extra facilities for the control of congested traffic. The most radical of which is the gating facility to relocate traffic queues by restraining traffic on user specified roads. TRL has successfully operated gating in this project in Southampton and Kingston-upon-Thames to restrain traffic as it approaches a critical part of the road network. In Southampton, gating operated as part of the Bitterne Road bus priority scheme. The practical use of Gating has demonstrated that it is a valuable traffic management tool and the experience gained has been used to write a user guide to help local authorities make the best use of SCOOT gating.

Another strand of work in the project has developed new techniques for alleviating the effects of urban congestion. Although SCOOT has proved to be extremely effective in achieving significant savings in delay when compared to Fixed Time systems, there is the possibility that SCOOT's delay minimisation objective may not be the most effective technique for controlling networks when there is a high level of congestion. Simulation studies were designed to gain an understanding of the mechanism by which congestion spreads, and to investigate those extreme situations and circumstances where SCOOT in its normal mode of operation loses effectiveness. The main problem areas are where the main road link at a junction is short and the saturation flow of the main road is considerably greater than that of the side road, or where blocking back of one link can lead to blocking of a number of other links causing a rapid spread of congestion.

Based on the results of the simulation a new measure of congestion was developed called 'Wasted Capacity' derived from SCOOT data. A computer program, MONACO, has been produced to use the Wasted Capacity measure to diagnose automatically the cause of localised chains of congested links. The program first identifies the critical link causing Wasted Capacity to build up on a route in the controlled area; the critical link is the first link downstream of a chain of links that each have Wasted Capacity. A decision tree is then used to analyse the likely cause of the problem on the critical link.

There are several possible causes of the problem on a critical link. In a large SCOOT system the database contains data on many links. It is possible with such large databases that errors may be overlooked. MONACO will rapidly identify the likely data error where the mistake causes congestion. In addition the program can recommend the use of the various congestion facilities within SCOOT to improve the control on problem links. It will also recognise junctions that are overloaded and causing congestion to build up. In such cases there is no simple solution, either the junction must be physically rebuilt with greater capacity, traffic restrained away from the junction or the congestion accepted.

The third approach, within the project, to try to alleviate the effects of congestion has been to investigate the benefit of a global change in optimisation criterion. The MONACO approach of identifying individual problems and attacking them is appropriate for specific problems, but when a road network is globally over-loaded then a different approach is required. SCOOT's objective function is to minimise a weighted sum of delay and stops. In congested conditions it is theoretically possible to increase the throughput of many networks by changing the objective function to one of maximising capacity. This is achieved by increasing the use of high capacity roads. However, the theoretical increase depends on individual drivers re-routing to take advantage of the extra green time on high capacity roads. The actual increase in throughput depends on the driver response as well as on the change in control strategy. Within this project CONTRAM and CONTRAMI were used to investigate the sensitivity of the benefits of a maximum capacity strategy to the level of driver re-routing. It was found that the benefits were dependent on the level of driver response, but that the modelling programs could give only limited guidance on the level of the benefits in a particular network. Proposals for possible on-street experiments have been developed.

NMAM112 UTC STRATEGIES FOR CONGESTED NETWORKS

ABSTRACT

TRL has undertaken a wide ranging study for the Department of Transport of how to use Urban Traffic Control systems in congested road networks to improve traffic conditions in urban areas of the UK. The first approach was to investigate how best to use the traffic gating facility in SCOOT V2.4 to restrict vehicles approaching particularly sensitive parts of the road network and to provide guidance to users. Simulation studies were used to investigate the extreme conditions of congestion and link geometry under which the normal operation of SCOOT might not be optimum. A quantitative measure of the adverse effects of congestion, Wasted Capacity, was developed. This measure formed the basis of a computer program, MONACO, to automatically analyse and diagnose the causes of congestion. The final approach was to investigate the potential of changing the objective function of SCOOT, under congested conditions, from minimum delay and stops to capacity maximisation. The sensitivity of the benefits to driver re-routing was investigated.

1. INTRODUCTION

TRL has undertaken a wide ranging study of congestion in urban areas for the Department of Transport with the aim of improving traffic conditions in UK towns and cities. The work has concentrated on how to use Urban Traffic Control (UTC) systems to improve the operation of urban traffic signals in congested conditions. SCOOT is widely used in the UK, and elsewhere in the world. From Version 2.4 onwards, SCOOT has incorporated extra facilities for the control of congested traffic. The most radical of which is the gating facility to relocate traffic queues by restraining traffic on user specified roads. TRL has operated gating in this project and used the experience to write a user guide to help local authorities make the best use of SCOOT gating.

Another strand of work in the project has developed new techniques for alleviating the effects of urban congestion. Although SCOOT has proved to be extremely effective in achieving significant savings in delay when compared to Fixed Time systems, There is the possibility that SCOOT's delay minimisation objective may not be the most effective technique for controlling networks when there is a high level of congestion. Simulation studies were designed to gain an understanding of the mechanism by which congestion spreads, and to investigate those situations and circumstances where SCOOT in its normal mode of operation loses effectiveness.

Based on the results of the simulation, a new measure of congestion was developed called 'Wasted Capacity' derived from SCOOT data. A computer program, MONACO, has been produced to use the Wasted Capacity measure to diagnose automatically the cause of localised chains of congested links. In a large SCOOT system the database contains data on many links. It is possible with such large databases that errors may be overlooked. MONACO will rapidly identify the likely data error where the mistake causes congestion. In addition the program can recommend the use of congestion facilities within SCOOT to improve the control on problem links. It will also recognise junctions that are overloaded and causing congestion to build up. In such cases there is no simple solution, either the junction must be physically rebuilt with greater capacity, traffic restrained away from the junction or the congestion accepted.

The third approach has been to investigate the benefit of a global change in optimisation criterion, rather than the MONACO approach of identifying individual problems and attacking them. SCOOT's objective function is to minimise a weighted sum of delay and stops. In congested conditions it is theoretically possible to increase the throughput of many networks by changing the objective function to one of maximising capacity. However, the theoretical increase depends on individual drivers re-routing to take advantage of the extra green time on high capacity roads. The actual increase in throughput depends on the driver response as well as on the change in control strategy.

2. USE OF SCOOT GATING

Part of the research work in this project was to examine how best to use the congestion controls, particularly the SCOOT gating facility, that are available in Version 2.4 of SCOOT. To gain experience in the use of gating, TRL used it to replace the custom fixed-time restraint plans in the Bitterne Road bus priority system in Southampton and to restrain traffic entering Kingston-upon-Thames in the morning peak period. Other examples of the use of gating are on the signalised roundabout at junction 10 of the M25 and to limit traffic approaching Nijmegen in the Netherlands from two motorways. A user guide to gating has been written (Wood and Baker 1994).

2.1 GATING ON THE BITTERNE ROAD IN SOUTHAMPTOS,,

TRL has installed Gating in an area of Southampton known as the Bitterne Road corridor. Gating is working in conjunction with an existing bus priority scheme which has been in use for over 20 years, Department of the Environment (1975). The bus priority scheme is based on fixed-time plans which restrain private traffic in side roads and at gating points on the two main approaches to Southampton that run through the area, see Figure 1. Buses are given priority by "buses-only" facilities on entry to the corridor, which by-pass the queues of restrained vehicles that occur regularly in the area during the morning peak.

TRL has replaced the fixed-time plans by the SCOOT Gating facility to provide the traffic restraint within SCOOT, but no changes are being made to the bus priority features. Gating of the side roads and the Bitterne Road is controlled by conditions at junction A in Figure 1 and the main road traffic on Bursledon Road is controlled by Junction B on the figure.

The TRL ASTRID database system (Hounsell and McLeod) was used to monitor the morning peak traffic conditions to compare Gating and the old fixed-time plans. The database automatically collects, stores and processes traffic information for display or analysis. ASTRID is currently being developed to run on-line and it has been installed in Southampton as part of TRL's work in the DRIVE 2 project, ROMANSE. An on-street survey was undertaken of the effects of Gating on bus journey times.

2.1.1 Assessment of Bitterne Road Gating

Table 1 shows the results of the survey of bus journey times; results are shown separately for buses approaching the city on Bursledon Road from those approaching on Bitterne Road.

No direct changes were made to the bus priority part of the scheme and, therefore, it is not surprising that little change was measured to the overall bus journey times. Similar numbers of buses use the two roads and, as can be seen from the last column of Table 1, overall, no difference was measured in the bus journey times with the two methods of control. However, buses could have been affected *indirectly* by changes in conditions on the main road caused by the different method of controlling non-bus access to the main road. The overall impression from the results in Table 1 is that SCOOT Gating slightly improved conditions for buses approaching Southampton on Bitterne Road and slightly worsened conditions for those using Bursledon Road. The differences recorded in Table 1 are small and not statistically significant at the 95 per cent level.

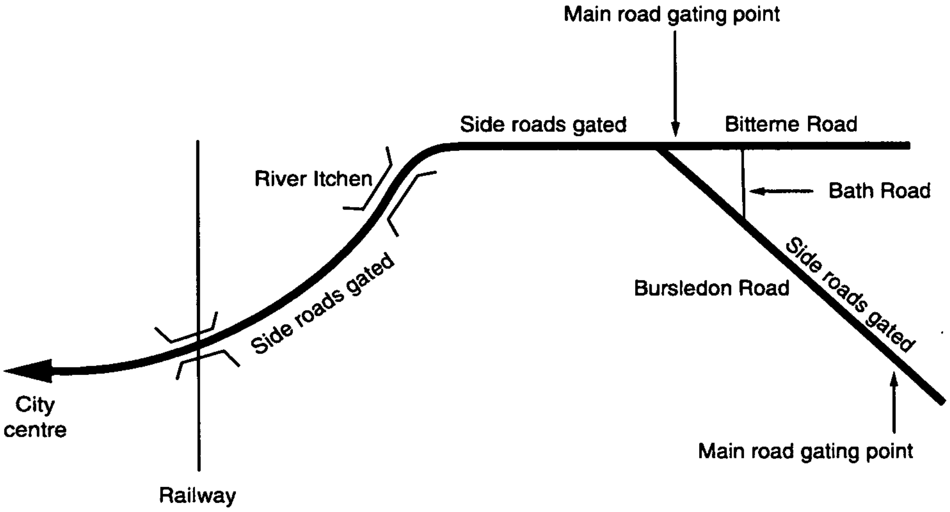

Fig 1. Schematic plan of Bitterne Road scheme

TABLE 1

Results of the bus survey, journey times between timing points (min:s)
Standard errors of the means are shown in brackets

Bus route	Bursledon Road	Bitterne Road	Combined
Fixed time	15:20 (0:32)	13:07 (0:19)	14:14 (0:20)
SCOOT Gating	16:19 (0:42)	12:15 (0:21)	14:22 (0:34)

A possible explanation for the different changes measured for buses approaching the city on Bursledon Road from those approaching on Bath Road is that the SCOOT Gating may have altered the balance of delay at the junction of Bath Road and Bursledon Road. In the morning peak buses approaching along Bitterne Road are routed down Bath Road to join Bursledon Road shortly before its junction with Bitterne Road. The reason for this route is to avoid the queue of gated traffic on Bitterne Road; other vehicles cannot use the route as only buses are allowed to turn right from Bath Road at its junction with Bursledon Road. Therefore, the different effect on journey time on buses on the two roads may be due to the Gating reducing the delay for traffic exiting Bath Road at the expense of traffic on Bursledon Road.

The effect of the SCOOT Gating on non-bus traffic was derived from the SCOOT internal model. The data were collected by the ASTRID database for convenience and analysed to show the effects on all the gated links and on all the main road links that should be free-flowing and benefit from the scheme. Flow-weighting was used to produce the average delay per link. Table 2 summarises the results.

There was no overall difference in the flows in the network between strategies, although there were some differences on individual links. Delay normally increases with the flow in a network and survey results need to be corrected for this effect. However, in this survey no flow corrections were needed to the delays because the overall flows were the same during the periods when the two strategies were being tested. The delays on each link were recorded every quarter of an hour by ASTRID. The differences between the two control strategies were tested for statistical significance using a paired 't' test of the mean values in each quarter hour of the two hour survey period.

All the changes in delay listed in Table 2 are downstream of the relevant SCOOT detector. Delays upstream of the detectors are indicated by congestion on the link. Several gated links were considerably more congested with fixed-time control than with SCOOT Gating, particularly Bitterne Road at its junction with Bursledon Road, the upper main road Gating point in Figure 1. Therefore, the delays on the gated links in Table 2 are underestimates, particularly for fixed-time control. The extra delay, upstream of the detector, was estimated from the degree of congestion and flow and the results are given in Table 3.

When the results for individual links were examined, it could be seen that the SCOOT Gating had redistributed some of the delays on the gated links. With fixed-time there were large delays on Bitterne Road at its junction with Bursledon Road plus the congestion described above. With SCOOT Gating there was less delay here but at the other main road entry into the system, on Bursledon Road, there was more delay than with fixed-time control. With some effort, the fixed-time plans could have been re-tuned to

TABLE 2

Non-bus traffic, flows per link (vehicles/h) and delays per link (veh-h/h).

	Flow per link, Fixed-Time	Flow per link, SCOOT Gating	Delay per link, Fixed-Time	Delay per link, SCOOT Gating
"Free-flow" Main road links	929	925	5.5	1.4*
Gated links	239	244	9.9	12.1
All links	570	571	6.5	3.8*

* Statistically significantly different from the Fixed-time value at the 95% level.

TABLE 3

Estimated delays allowing for congested links (veh-h/h)

	Estimated delay per link, Fixed-Time	Estimated delay per link, SCOOT Gating
Gated links	23.6	15.3*
All links	9.3	4.5*

* Statistically significantly different from the Fixed-time value at the 95% level.

produce more equal delays over the entry links as the SCOOT Gating has done. It was postulated above, when discussing the bus results, that the journey time along Bath Road was probably less with SCOOT Gating than with the fixed time system. The SCOOT model results supported this hypothesis as it estimated a small increase to the main road traffic at the junction of Bursledon and Bath Roads, implying that extra green time was given to the side road.

There were 25 links in the survey and, therefore, the best estimate of the journey time saving for non-bus traffic is (9.3-4.5) x 25 = 120 vehicle-hours of delay per hour during the two hour survey period. This amounts to a saving of some 60 000 vehicle-hours over a year.

2.1.2 Conclusions from the Bitterne Road

SCOOT Gating has successfully replaced the Bitterne Road fixed-time control scheme. The bus survey showed little change; a small reduction in journey times for buses approaching the city on Bitterne Road and a similar small increase for those approaching on Bursledon Road. The small overall change is not surprising as no direct change was made to the operation of the bus priority scheme. For non-bus traffic, the journey times have reduced and the most severe delay has been reduced at the expense of increasing delays at other entries to the system. The delays at the major entries now appear to be more equal than they were.

The time savings quoted in section 2.1 are equivalent to a saving of approximately £240 000 per annum using the standard value of time for non-working cars.

2.2 KINGSTON-UPON-THAMES

There is a problem in Kingston-upon-Thames caused by queues obstructing the operation of a critical junction in the main one-way system. If the merge area at this junction becomes obstructed by queues of stationary traffic, queues rapidly build up in the rest of the system and can cause gridlock if no action is taken. The normal cause is traffic passing through the town blocking back from the roundabout on the West side of Kingston bridge. A Gating experiment was undertaken to restrict traffic on three main entries to the Kingston one-way system. The police have clearance plans available that restrict these entries, but are not usually aware of the problem until it has become very serious. Gating should be able to take action automatically, as soon as necessary, to store the restrained traffic on the same roads that are used by the police clearance procedures. The effectiveness of Gating was assessed by analysis of video records of traffic queues backing up to the merge area at the critical junction.

Figure 2 shows the length of individual blocking periods. Without Gating there were several cycles where the merge area was continuously blocked for over 30 seconds, a considerable period. The results with gating are quite impressive, the conditions at the merge area were greatly improved. The total period, during which any significant congestion was observed, was reduced from about an hour, without Gating, to between a quarter and half an hour on different days with Gating. With Gating, the longer periods of congestion, of over 30 seconds, were virtually eliminated and the total time that the merge area was blocked by stationary traffic was reduced by nearly an order of magnitude. Without Gating the merge area was blocked for about 20 minutes in each morning peak, but with Gating, it was blocked for only about 2 to 3 minutes.

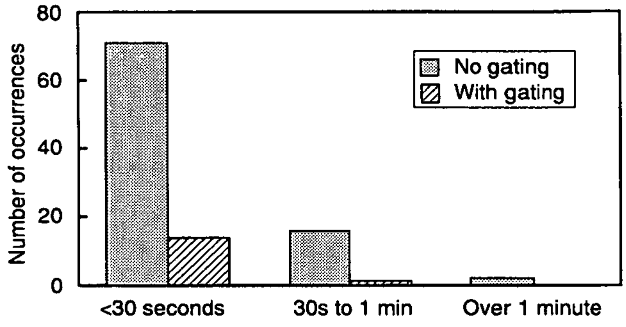

Fig 2. Length of individual blocking incidents.

The Gating experiment was very successful in reducing the level of congestion within the Kingston-upon-Thames one-way system.

2.3 USERS' EXPERIENCE

2.3.1 Kingston-upon-Thames

Since the TRL experiment in Kingston, Gating has been implemented for the morning peaks on weekdays. Because of the severity of Gating, the Traffic Control Systems Unit, in cooperation with the London Borough of Kingston-upon-Thames, have increased the minimum greens on the gated links, during the Gating hours, to prevent excessive queues.

2.3.2 M25 junction 10

Surrey County Council are very pleased with the operation of Gating at the roundabout on the A3 at junction 10 of the M25. The motorway is being widened to 4 lanes west of the junction. During the morning peak, westbound traffic wishing to join the motorway at junction 10 backs up onto the roundabout because of the road works. Gating has successfully prevented severe lock-up problems and has kept the circulatory carriageway free without creating severe queues on the entries to the roundabout.

2.3.3 Nijmegen

The SCOOT system in Nijmegen in the Netherlands has been installed by Dutch consultants for the Dutch Ministry of Transport. Gating is an integral part of the scheme. The two main links into the city from the motorway system are gated to prevent overloading the city network. Previously, restricted fixed signal timings were used, particularly in the morning peak. Now, with the SCOOT system, Gating operates throughout the day. In the morning peak periods there are normally large queues at the gating points because of the amount of traffic approaching the city. At quiet times of day there are no queues, and no restraint, but if there is a sudden influx of traffic, then the Gating logic provides the necessary restraint and holds the excess vehicles at the gating points. Restraint is available throughout the day, but it is only used when justified. The Dutch engineers are happy with the operation of Gating; as in Southampton and Kingston they used minimum stage lengths to prevent excessive restraint and tuned the Gating saturation for their particular circumstances. The Dutch Ministry of Transport are assessing the whole system.

2.4 CONCLUSIONS ON SCOOT GATING

SCOOT Gating has been shown to be a valuable traffic management tool. Traffic engineers can use it to control where queues form in a congested road network and can relocate queues to less critical areas of the network. As part of a bus priority scheme, Gating can move queues of delayed vehicles to roads where it is possible to give buses priority, by bus lanes, or other means, to by-pass the queues of congested traffic. Therefore, Gating enables the traffic engineer to give priority to buses in situations where, otherwise, there would not be sufficient road space to reserve a separate lane for buses. Details of how to use SCOOT Gating are given in the user guide (Wood and Baker 1994).

3. SIMULATION STUDIES OF CONGESTED TRAFFIC CONTROL

Studies of how the details of traffic control affect congested urban traffic have been undertaken on a number of different networks using the STEP simulation. STEP simulates individual vehicles on links in a network of traffic signals and produces information on queues and delays in the network. STEP can be used to simulate the effect of changes to the SCOOT logic or other traffic signal system in a consistent and repeatable manner, under normal, oversaturated and congested traffic conditions. Flows and turning movements, and any SCOOT parameter, can be varied throughout the simulation.

3.1 SIMULATION STUDY

To gain an understanding of the relationship between signal timings and the build up of congestion, an initial study was made using a simple network of two nodes, connected by a one way link. A study has been made of the effect on the capacity of the network of different green splits and offsets under fixed signal settings. The flow was set so that under all the signal settings the network was oversaturated. The parameters studied were: link length, turning proportions and cycle time.

The main results of the study have been to demonstrate the circumstances in which the offset is critical to the capacity of the network, and to assist in understanding the mechanism by which capacity of a network can be lost.

Figure 3 shows the number of vehicles queuing for each offset for four different link lengths. The queues waiting to join the main and minor roads are shown separately. This figure illustrates the different shape of the capacity/offset relationship when the adjoining link is short (LL=8) and when it is longer (LL=12,16 and 20). For this series of runs a cycle time of 60 seconds was used and the main road/side road stage lengths were 40/20 and 30/30 at the upstream and downstream junctions respectively; the traffic demand on the main road at the upstream junction was twice that on the side road and the demands were balanced at the downstream node.

The main conclusions from the simulation runs presented in figure 3 are:

In oversaturated conditions the effect of offset on the capacity of a network is very dependent on the length of the critical link.

A link can be considered as 'long' if the time taken to discharge the queue of vehicles waiting on the link when it is full is greater than or equal to the length of effective green at the downstream junction. Otherwise a link can be considered as 'short'.

If the link is 'long' then the capacity of the *downstream* or bottleneck junction is unaffected by the offset. This is because, irrespective of the offset, there are always sufficient vehicles queuing on the link to fully utilise the green time, i.e. there is no wasted green at the bottleneck junction.

If the link is 'short' then the effective capacity of the downstream junction is dependent on the offset. Initially when the lights go green the queue on the link discharges and the flow of vehicles through the junction is maintained at the saturation flow rate. Once this queue has disappeared, however, the flow of vehicles will be dependent on the stage running at the upstream node. If, for instance, the side road at the upstream junction is on green, the flow through will be only those vehicles that have turned. This will lead to a reduction in the throughput and, hence in the effective

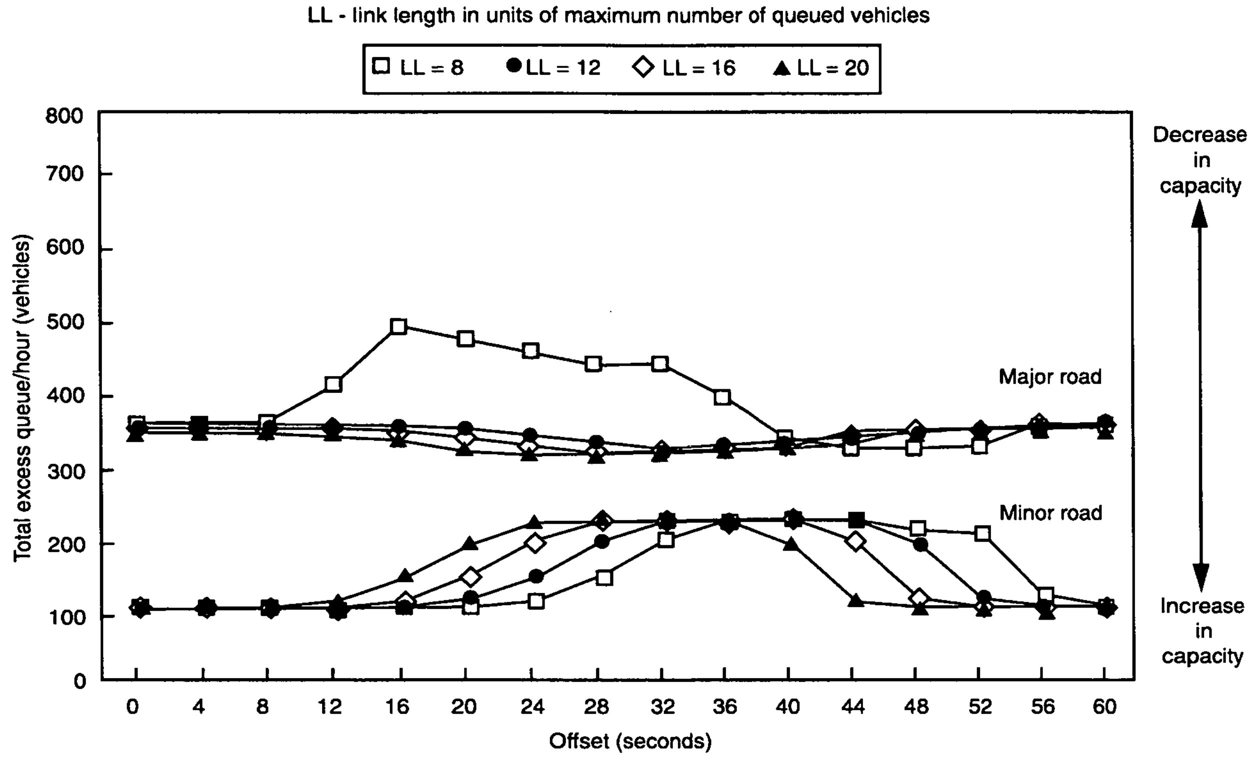

Fig 3. Vehicle queuing for different links under fixed time control.

capacity at the downstream junction, as can be seen for the short link (case LL=8) in Figure 3.

For both long and short links the capacity of the *upstream* junction can be dependent on the offset. The extent to which the offset is critical depends on the effect of a blocked main road link on traffic moving straight ahead from the side road. In the simulation a blocked turning vehicle will itself block any vehicles behind it, irrespective of their destination.

For 'long' links there is a range of offsets which are optimal and produce maximum capacity. For 'short' links the offset is more critical and in many cases, when the combined effect on both main and side roads is considered, only a small range or even just one value of offset will be optimal.

The effect of different turning movements is illustrated in figure 4, which shows the capacity/offset relationship for a range of turning movements. The turning proportions from the main and side roads have been chosen to maintain a constant demand for the critical link. Cycle time and splits were the same as in the previous example. A link length of 12 has been used which could be considered as 'long'. When there is no turning traffic, the offset has no effect on the capacity of the network. With moderate turning movements of cases TM2,1 and TM4,2 maximum capacity is obtained with offsets in the range 0-16 seconds approximately. However, with large turning movements as in the case of TM8,4, maximum capacity is obtained with offsets in the range 32-36 seconds. The outcome depends on whether there is a bigger *percentage* of vehicles from the main or side road wishing to use the blocked link. With turning movements of TM8,4 there is a smaller percentage of vehicles wishing to use the blocked link from the main road than from the side road. Thus a larger percentage (40%) can turn away from the congestion if the offset favours the main road compared to the 20% of side road traffic that avoid the congestion if the offset favours the side road. With TM6,3 the percentage of vehicles wishing to use the blocked link from the main road and side road are nearly the same. This causes the effect of the offset to be almost negligible.

The main conclusions from investigating the effects of turning movements are:

The turning movement can have an effect on the value of the offset which will maximise capacity. This is demonstrated in figure 4.

The optimum offset gives priority to those vehicles which do not wish to travel along the blocked link.

3.2 WASTED CAPACITY MEASURE

Having identified those conditions where the detailed control of the traffic signals can affect the capacity of the network a measure of the effect was required. In order to diagnose when there is a need to introduce special congestion strategies a 'new' congestion indicator, Wasted Capacity, has been derived. It is important that this indicator shows the likelihood of a loss of efficiency of SCOOT in its

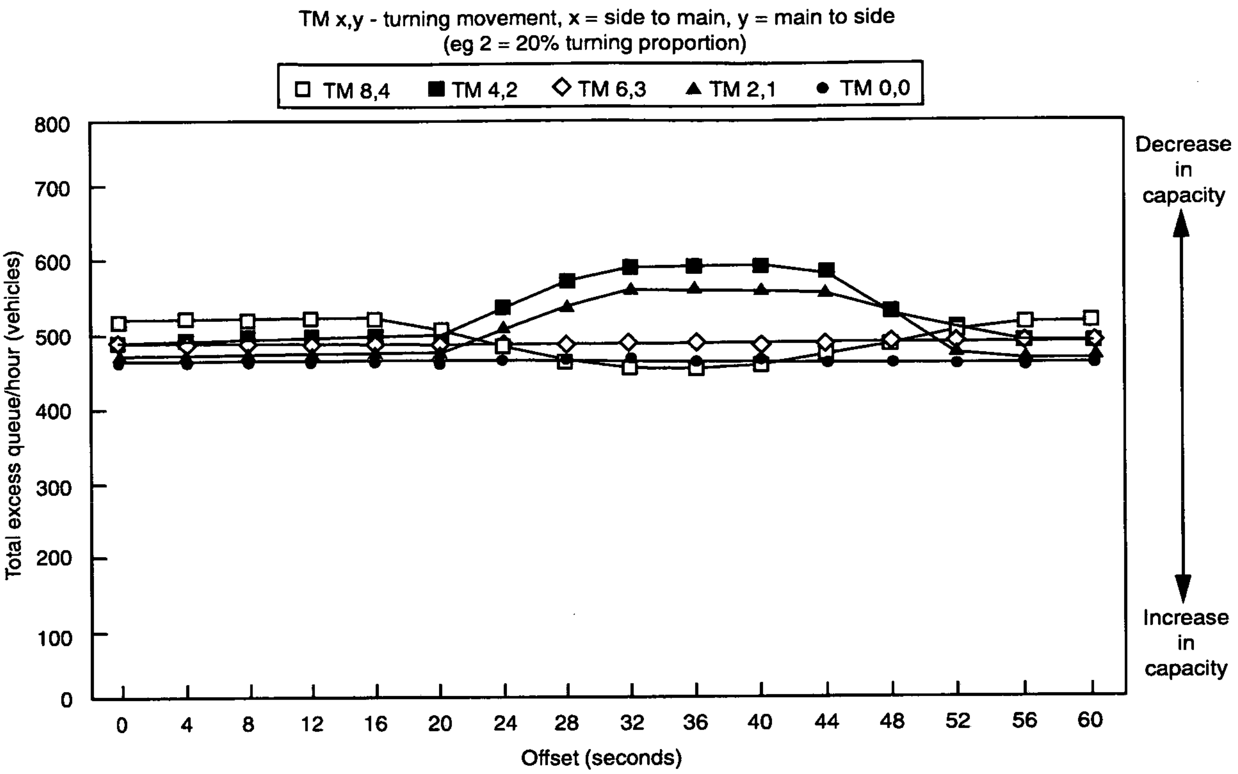

Fig 4. The effect of offset over different turning movements under fixed time control

normal mode of operation, rather than just indicating the actual level of congestion which could be measured, for instance, as a large increase in delay. It is important to determine not just that a network is congested but whether the congestion will spread to other parts of the network and cause further congestion.

A queue will cause a loss of capacity at an upstream junction by blocking the exit from that junction while the traffic signals are green. For each link at the upstream junction the amount of capacity potentially lost at the link is given by:

Lost Capacity = Time Blocked While Green x Saturation Flow

This Lost Capacity formula does not take into account that there may be no queue to disperse on the upstream link and therefore no vehicles would be prevented from moving.

To correct for this, Wasted Capacity is defined as:

Wasted Capacity = Min(Lost Capacity, Queue at End of Green)

and the Wasted Capacity at a node is the sum of the Wasted Capacities for the upstream links of that node.

The Wasted Capacity measure can be derived from standard SCOOT data. A method for calculating the measure has been implemented in the London SCOOT system where it has been used to monitor the Wasted Capacity information in real time. The Wasted Capacity in a SCOOT controlled network is measured in Link Profile Units (LPU) the SCOOT standard unit of demand. Normally a single vehicle is equivalent to 17 LPU.

4. AUTOMATIC DIAGNOSIS OF CAUSES OF CONGESTION

4.1 INTRODUCTION

A formal method of analysing congestion using Wasted Capacity has been developed and automated into a computer program called MONACO (MONitoring and Analysis of Congestion).

4.2 FUNCTIONAL SPECIFICATION

MONACO uses Wasted Capacity as a quantitative measure of the problems caused by congestion. Wasted Capacity is measured along routes that traverse a SCOOT network. A route is a series of SCOOT links that form a path through the area under SCOOT control. At present routes are entered into MONACO, but the program could be enhanced to read the SCOOT database and build up routes by using the "Main Downstream Link" of the current link to find the next link in the route. An example would be links:

N01/184e → N01/190a → N01/212e → N01/040a

Link N01/040a is the "main Downstream Link" of link N01/212e etc. As shown in figure 5, these particular links form a route travelling south along Regent Street across Oxford Circus. The interpretation of the levels of Wasted Capacity on different links in a route enables discovery of the critical link. The critical link is the link suspected of causing Wasted Capacity upstream on a route. Should several links have high levels of Wasted Capacity and then the next have little or none it is probable that the link has some form of obstruction on it. The critical link is defined as the first link with little or no Wasted Capacity in a route *after* (or downstream of) links with severe Wasted Capacity. For example, in the above route, if the levels of Wasted Capacity were 550, 650, 0 and 0 LPUs respectively (approximately 32, 38, 0 and 0 vehicles in the last analysis period) then link N01/212e would be the critical link and node N01/212 would be the critical node.

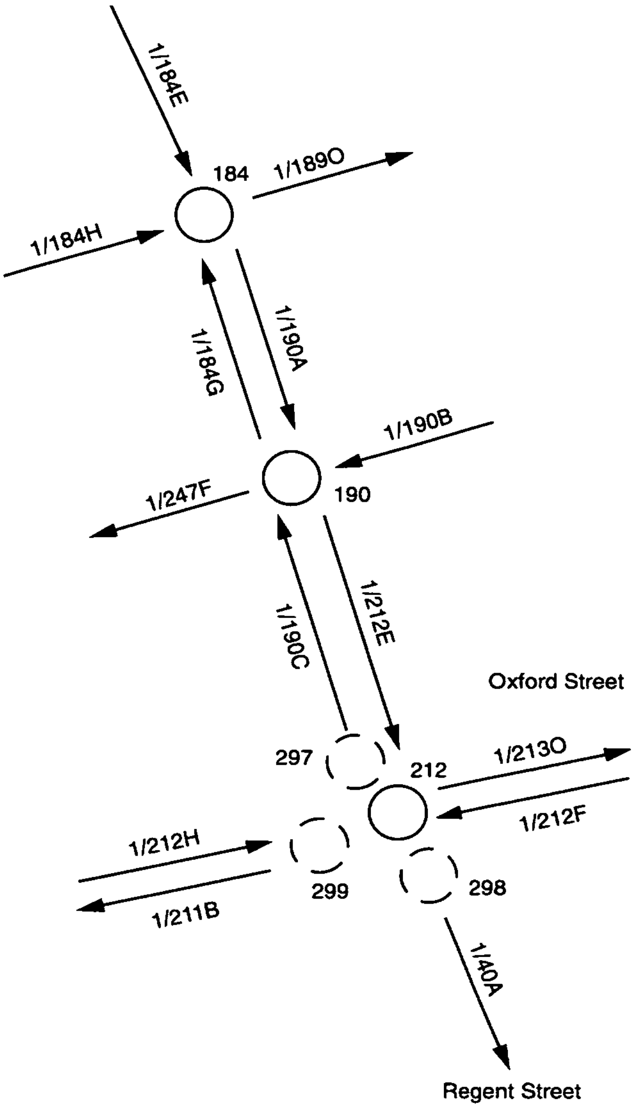

Fig 5. Route travelling south down Regent Street from N01/184e to N01/040a

MONACO is an IBM-PC (or compatible) based suite of programs that connects to, converses with and monitors an urban area under SCOOT control. Build ups of Wasted Capacity are calculated and measured and excessive amounts reported to the operator via a graphical display.

MONACO detects and diagnoses Wasted Capacity by the following method. Every five minutes each route is checked for excessive amounts of Wasted Capacity, should this be the case then MONACO investigates the route. During investigation MONACO will discover the critical link, the link where the congestion in spreading back from, and then attempts to diagnose the possible cause. A diagnosis given by MONACO will be governed by following the 'decision tree' in figure 6.

MONACO uses the decision tree to determine the probable cause of Wasted Capacity along a route in the monitored area. An example of how MONACO does this is given below, using the Regent Street route mentioned above:

Stage 1:	Wasted Capacity is measured along the above route, at the end of the five minute interval the levels of Wasted Capacity on this route are considered to be excessive.
Stage 2:	Each link in the route is examined. Wasted Capacity is found on links N01/184e and N01/190a, none is found on links N01/212e or N01/040a. This identifies link N01/212e as the critical link.
Stage 3:	MONACO starts at the root of the decision tree (the top in figure 6). The question here asks 'is the critical link heavily over saturated?'. This is the case, so MONACO follows the 'yes' path from its current position on the decision tree.
Stage 4:	MONACO arrives at another question in the decision tree. The question here asks 'is there a faulty detector at the critical node?'. This is also true, link N01/212f has a faulty detector. Again, MONACO follows the 'yes' path from its current position on the decision tree.
Stage 5:	MONACO arrives at a position in the decision tree with no questions. The position explains the probable cause on the Wasted Capacity as Default split wrong. So, for this example, MONACO would output a message stating the problem is likely to be associated with the default split for link N01/212f (the faulty link).

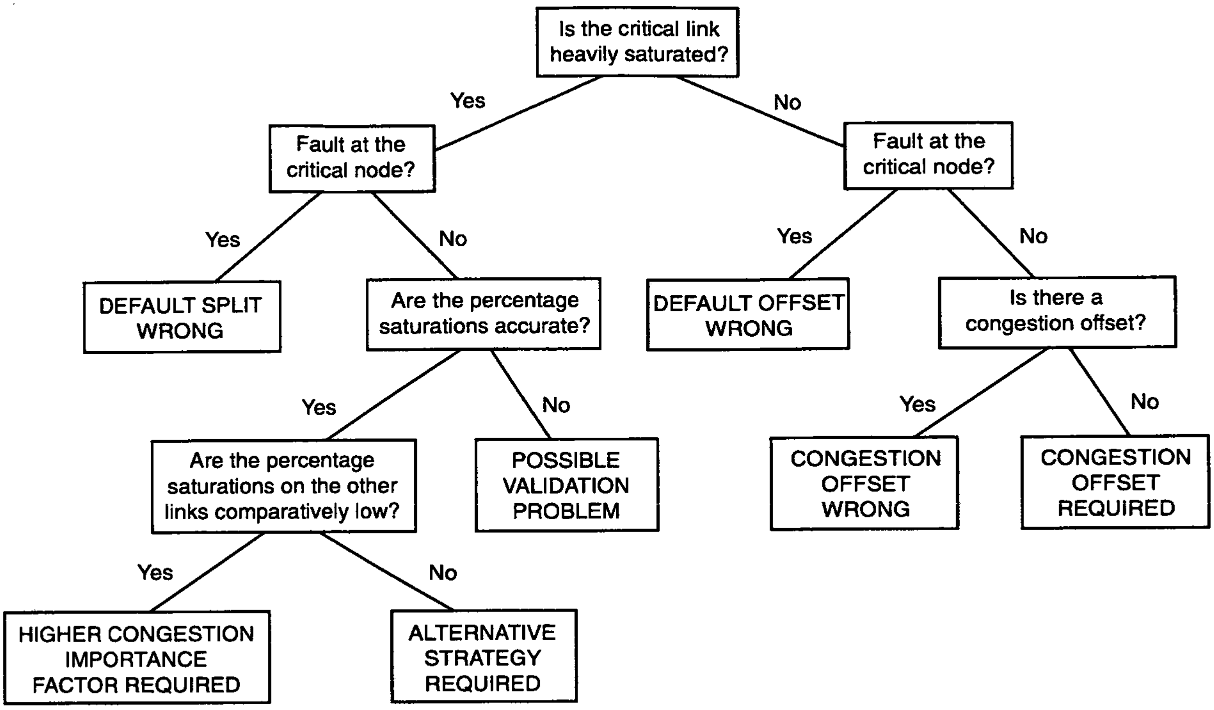

Fig 6. **The decision tree used by MONACO**

4.3 TESTING OF MONACO

Initially, MONACO was tested using off-line data already collected from the SCOOT computers in London; these tests enabled calibration of MONACO. After the off-line tests, the MONACO system was loaded onto a Viglen 386 Personal Computer and installed in the Operations Room at TCSU, Smiths Square, London for on-line testing.

4.3.1 Area to study

With MONACO installed in a single micro-computer, it was limited to studying SCOOT areas operated by one single SCOOT computer at a time. Therefore a decision had to be made as to which area, and therefore which computer, to choose. The decision was made to connect MONACO to the new VAX computer that now controls central London (referred to as SCOF in the TCSU computer network). This area, controlled by the SCOF SCOOT computer, stretches from Edgeware Road in the West to Shaftesbury Avenue in the North. SCOF is also responsible for the Trafalgar Square area.

4.3.2 Format of MONACO output

When MONACO detects and diagnoses the occurrence of congestion it displays messages to the operator and records information about the links affected. The links affected are considered to be all those links from the start of the congested route up to the critical node and in addition all those links attached to the critical node, for example:

	i	ii	iii
N01/091b	0	157	1020
N01/007e	0	106	234
N01/008a	0	78	0
N01/008b	1	-1	0

This data is concerned with a problem in Trafalgar Square, see Figure 7. The route affected was the Saint Martin's Lane entry to the north of the Square, travelling south. The congested route follows links N01/091b, N01/007e and N01/008a, which is the critical link. The only other link at the critical node is N01/008b, giving a total of four links. Each of the affected links in the route is stored with three values, these are:

i. detector status: this value is either 0, if the detector on that link is functional, or 1 if it is faulty.

ii. percentage saturation: this is the percentage saturation of the link (-1 indicates that the link is faulty).

iii. Wasted Capacity: this is the total Wasted Capacity calculated for the link, summed over the previous five minutes.

As well as the above information, which is logged to disk, MONACO outputs two messages to the operator informing him of the location of the congestion, its severity and probable cause. An example of MONACO's warning messages would be:

```
18:10:00 K01 N01/008b DEFAULT OFFSET WRONG
18:10:00 K02 CHAIN N01/091b, LINK N01/008a IS
                 CRITICAL, WASTED = 1254 LPUs
```

These two messages, which are linked to the example above, are output to the operator and logged to disk. The K01 message informs the operator of the probable cause of the congestion and the link from which the problem is emanating, in this example N01/008b. The K01 message can take seven formats depending on the probable cause of the Wasted Capacity, these are:

```
<time> K01 <link> DEFAULT OFFSET WRONG
<time> K01 <link> DEFAULT SPLIT IS WRONG
<time> K01 <link> POSSIBLE VALIDATION PROBLEM
<time> K01 <link> CONGESTION OFFSET REQUIRED
<time> K01 <link> CONGESTION OFFSET WRONG
<time> K01 <link> HIGHER CONGESTION
                  IMPORTANCE FACTOR REQUIRED
<time> K01 <link> ALTERNATIVE STRATEGY
                  REQUIRED
```

The K02 message takes just one format, this shows the time of the warning, the link at the start of the congested route, in this example N01/091b, the critical link on the route, N01/008a and the amount of Wasted Capacity created, in this case 1254 link profile units worth.

4.4 ANALYSIS OF DATA

MONACO collects and logs Wasted Capacity data over the monitored area plus any build ups of Wasted Capacity, recognition of Critical Links and a diagnosis of congestion detected. These log files were collected from the MONACO PC in London and examined at TRL.

Over a one week period, MONACO detected and diagnosed 156 congestion problems, as MONACO was being tested no immediate remedial action was taken and many of the 156 problems were repeat diagnoses of earlier problems. MONACO consistently found 10 critical links and 10 problem links on 8 different routes through the monitored area. The congestion problems MONACO can recognise fall into four broad categories.

4.4.1 Validation problems

If a problem is being caused by an unvalidated link in the monitored area, MONACO uses the message below to warn the operator:

```
POSSIBLE VALIDATION PROBLEM
```

This message is designed to tell an operator that a link in the area has possibly not been validated and is causing Wasted Capacity. This problem was detected in the earlier off-line testing of MONACO. There was only one occurrence of this problem in the area and, since then, the link has been validated. MONACO did not report this problem during the on-line testing.

4.4.2 Faulty detector messages

If a congestion problem is caused by a faulty detector, MONACO alerts the operator with either of the two messages below, depending on the circumstances:

```
DEFAULT OFFSET WRONG
DEFAULT SPLIT IS WRONG
```

Both of these messages warn the operator that a faulty link in the SCOOT network is causing congestion. The first states that a link is operating with an incorrectly validated default offset, suggesting that the operator should try alternative values. The second message informs the operator that a link at the critical node is operating at the wrong default split.

These two messages were reported in 43 cases over a one week period. Many of these were situated on link N01/008b in Trafalgar Square. The following three examples were the most frequently observed:

i. The following messages were output:

```
N01/091b   0   549  1204
N01/007e   0   156   369
N01/008a   0    78     0
N01/008b   1    -1     0

13:20:01 K01 N01/008b DEFAULT OFFSET WRONG
13:20:01 K02 CHAIN N01/091b, LINK N01/008a
             IS CRITICAL, WASTED = 1573 LPUs
```

The affected links in the route are:

N01/091b → N01/007e → N01/008a

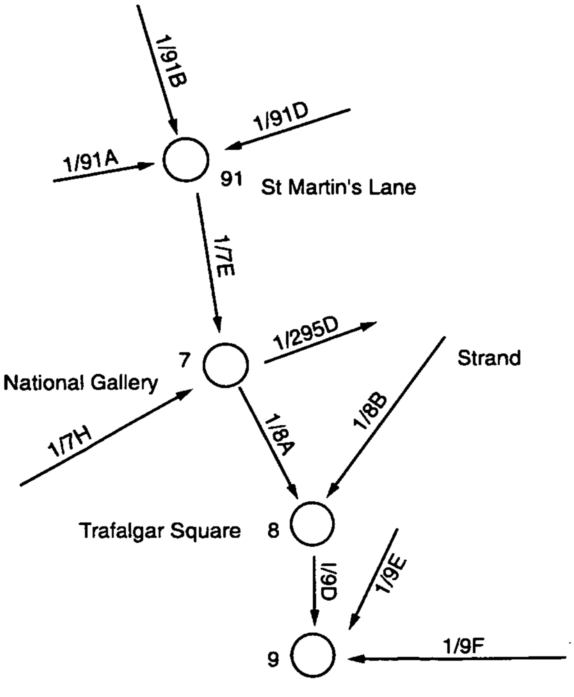

Fig 7. Route N01/091b through Trafalgar Square

This output states that link N01/008b is faulty and that there is a possible problem with the default offset. MONACO has reported a problem on this link as many as 60 times in a one week period. Investigation has shown that link N01/008b alternates between going suspect faulty and functional on a frequent basis. A detector goes suspect faulty when it has been occupied continuously for more than three minutes, study of link N01/008b has shown that this often occurs due to the geometry and traffic flow of the area. The action that should be taken for this link is to increase the suspect detector timeout so that the detector is not mistakenly set faulty.

ii. The MONACO information below was recorded for the route eastwards along Mortimer Street:

```
N01/184h    0   296  435
N01/185d    0   130  156
N01/186h    0   128  121
N01/187d    0    79  121
N01/188h    0   121  530
N01/189d    0   109    0
N01/189b    1    -1    0

15:40:00 K01 N01/189b DEFAULT SPLIT IS WRONG
15:40:00 K02 CHAIN N01/184h, LINK N01/189d IS
                   CRITICAL, WASTED = 1363 LPUs
```

Figure 8 shows the links comprising the route east along Mortimer Street and the links associated with the critical node. The links in the route are shown below.

N01/184h → N01/185d → N01/186h → N01/187d → N01/188h → N01/189d

Link N01/189d is shown as being critical in the route, i.e. there is no Wasted Capacity on this link but a significant amount upstream, suggesting that there is a problem with this link. Link N01/189b is flagged with the message that the default split is wrong, the probable cause of the congestion. This message informs the operator that the default split, currently in operation because the link is faulty, is incorrect. This is therefore causing a building up of congestion along Mortimer Street. MONACO has shown that the default split in operation for link N01/189b is incorrect, an investigation as to this being generally true or that the Wasted Capacity only occurs at certain times has not been carried out yet. It might be the case that this default split works effectively for the greater part of the day and that no change would be advantageous.

iii. A third example of an incorrect default split is:

```
N01/212g    0   750 1164
N01/190a    1    -1    0
N01/190c    0   108    0
N01/190b    0    98    0

13:10:00 K01 N01/190a DEFAULT SPLIT IS WRONG
13:10:00 K02 CHAIN N01/212g, LINK N01/190c IS
                   CRITICAL, WASTED = 1164 LPUs
```

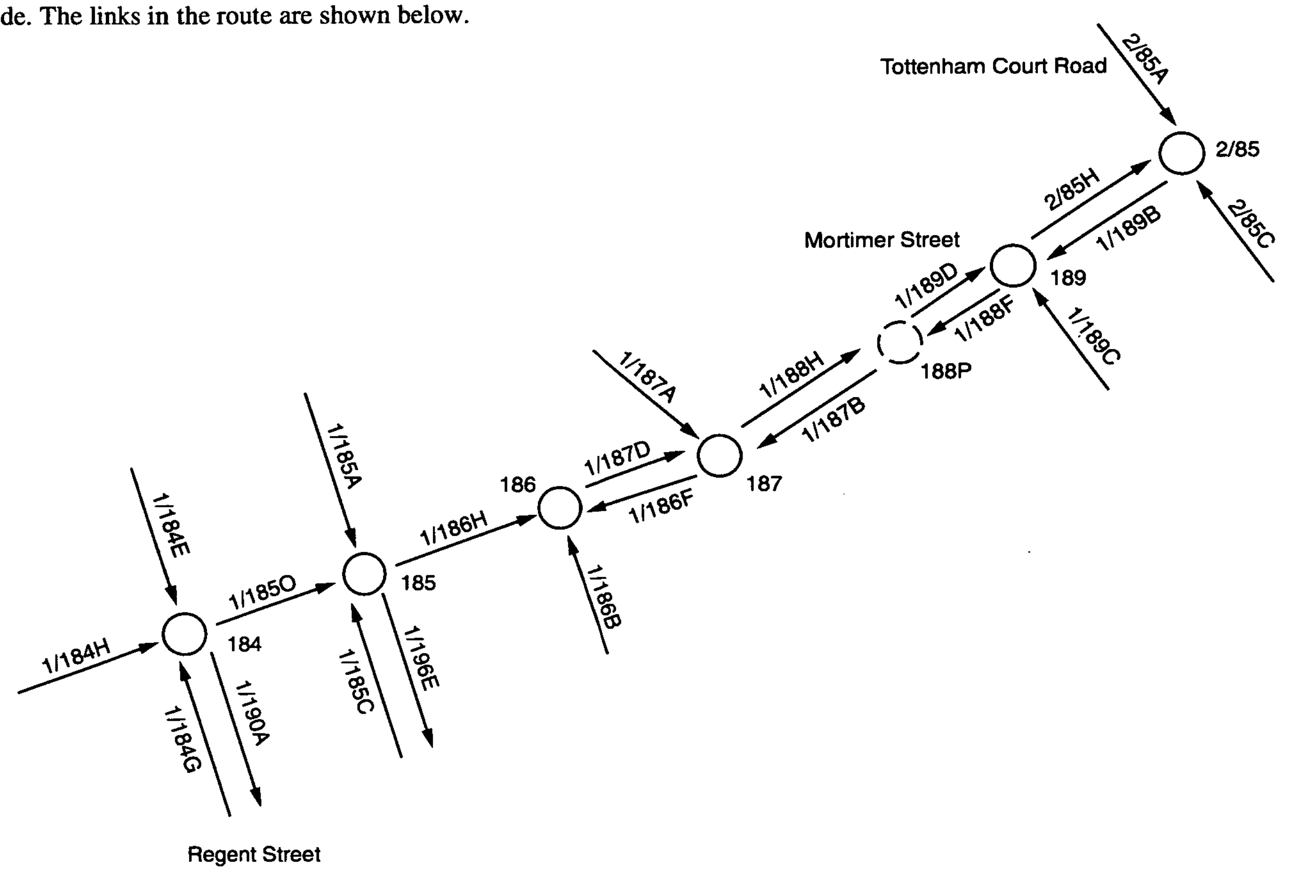

Fig 8. Route N01/184h east along Mortimer Street

This output was connected to the route shown in figure 9 that travels north across Oxford Circus. The links in the route are:

N01/212g → N01/190c

link N01/190a is flagged as the default split being too large. The default split on link N01/190a is too small, causing too little green time on link N01/190c which has caused the Wasted Capacity on link N01/212g. As with example ii. above, no investigation as to the ideal default split has been carried out, the current setting might be the best 'general case' value possible.

4.4.3 Congestion offset message

Should MONACO suggest that a congestion problem is caused by an incorrectly set, or the lack of, a congestion offset one of the messages below is displayed to the operator.

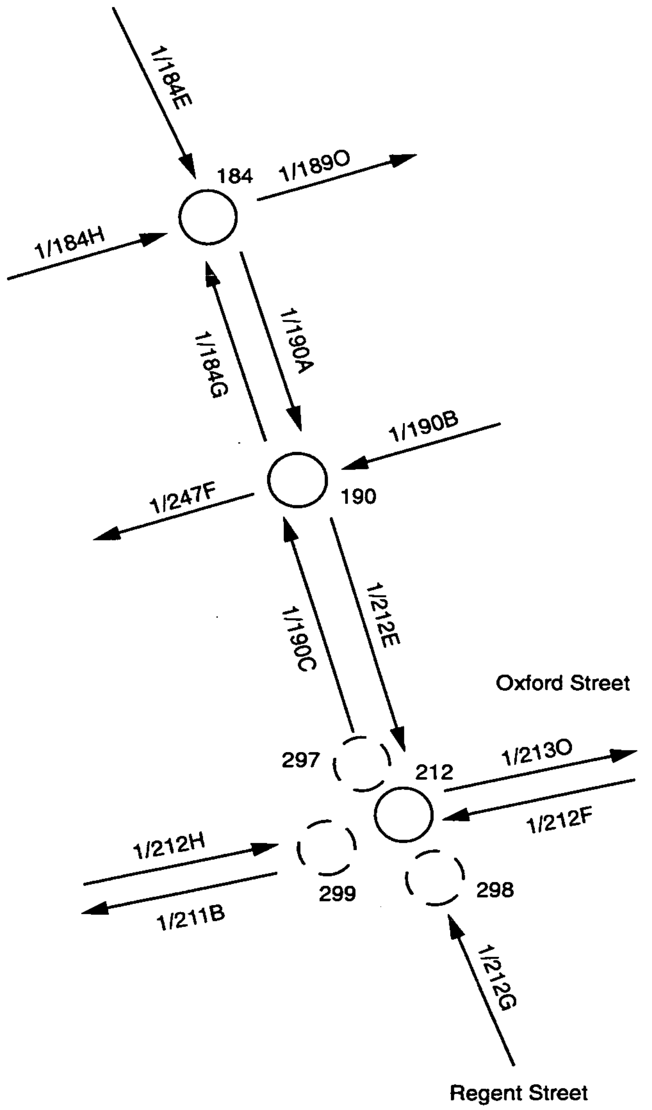

Fig 9. Route N01/212g north up Regent Street

```
CONGESTION OFFSET REQUIRED
CONGESTION OFFSET WRONG
```

The CONGESTION OFFSET WRONG message did not appear in the course of testing MONACO. This was due to there being no serious congestion occurring on the few links that have congestion offsets. The CONGESTION OFFSET REQUIRED message appeared for the following links:

N01/007e, N01/176x, N01/040a, N01/247f and N01/295d

Travelling west above Oxford Circus is a route crossing the northern section of Regent Street, see figure 10. The following MONACO data was recorded for this route:

```
N01/194f     0   134  1009
N01/190b     0   118   168
N01/247e     0    35     0
N01/247f     0    60     0
```

```
15:10:00 K01 N01/247f CONGESTION OFFSET
             REQUIRED
15:10:00 K02 CHAIN N01/194f, LINK N01/247f IS
             CRITICAL, WASTED = 1177 LPUs
```

The links in the route are:

N01/194f → N01/190b → N01/247f → N01/243b

Link N01/247f is identified as the critical link, link N01/190b has minor Wasted Capacity along it and, further upstream, link N01/194f has very heavy Wasted Capacity. The critical link is a 'short' link, i.e.. it is short enough that the offset along them is very important to the traffic flow. The importance of the offset and the Wasted Capacity that is being caused would suggest that a congestion offset on this link would be appropriate.

In fact, all the links in this route are 'short'. It would probably be wise to consider adding a congestion offset to link N01/190b as well. It can be seen that link N01/190b is also causing Wasted Capacity since the Wasted Capacity levels increase as you move upstream from the critical link, in addition, N01/190b has short green times, making the offset more important still. This observation has been made from the existing output shown above, had the operator merely placed a congestion offset on link N01/247f, as MONACO suggests, then MONACO would have detected the same problem at a later time, but with link N01/190b as the critical link.

4.4.4 Saturation messages

If a congestion problem is caused by an oversaturation of a link or node, MONACO alerts the operator with either of the two messages below, depending on the circumstances:

```
HIGHER CONGESTION IMPORTANCE FACTOR
REQUIRED
ALTERNATIVE STRATEGY REQUIRED
```

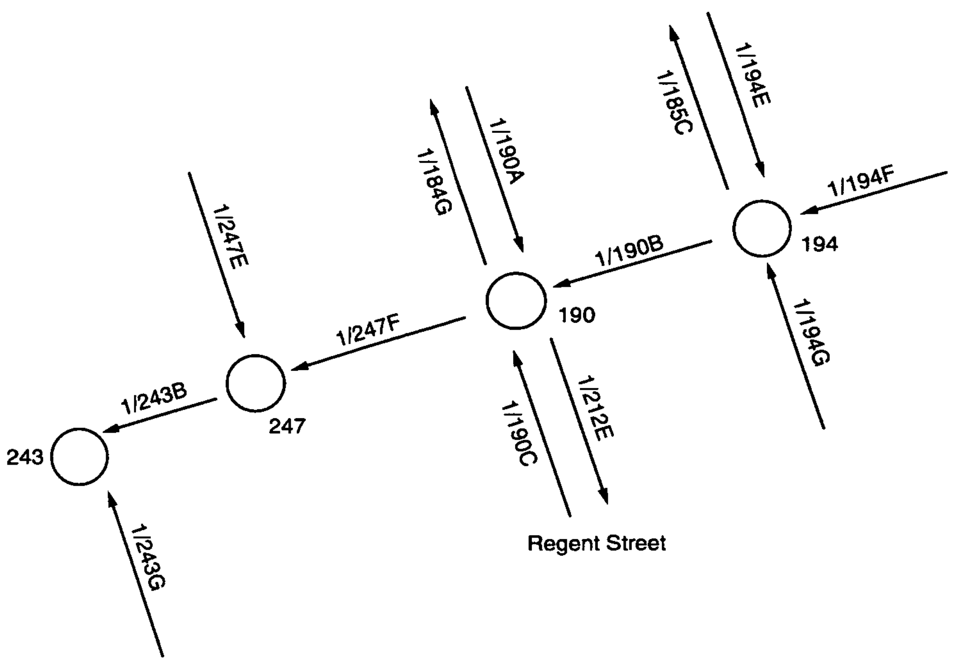

Fig 10. Route N01/194f west across Regent Street

The HIGHER CONGESTION IMPORTANCE FACTOR REQUIRED message informs the operator that critical link is oversaturated in relation to the other links at the critical node, meaning that the critical link requires a higher congestion importance factor to gain more of the available green time at that node.

The ALTERNATIVE STRATEGY REQUIRED message informs the operator that the critical link *and* other links at the critical node are oversaturated, indicating a more serious congestion problem requiring a more global solution.

The MONACO output:

```
N01/213b     0    0  1015
N01/212e     0   94   137
N01/212h     0  121     0
N01/212g     0   72   116
N01/212f     0  165     0

14:15:00 K01 N01/212f HIGHER CONGESTION
             IMPORTANCE FACTOR REQUIRED
14:15:00 K02 CHAIN N01/213b, LINK N01/212f IS
             CRITICAL, WASTED = 1015 LPUs
```

Indicated that a higher congestion importance factor was required.

The congested route is west bound along Oxford Street across Oxford Circus. The links involved are:

N01/213b → N01/212f

MONACO has informed the operator that the congestion importance factor on link N01/212f (western approach to Oxford Circus) is probably too low in relation to the other links. Congestion is building up along Oxford Street at link N01/213b. This message appears when the saturation on the problem link (N01/212f) is significantly higher than the other links at the critical node. Suggesting that the problem link should be allowed more weighting when SCOOT decides the green split.

An example of when the ALTERNATIVE STRATEGY REQUIRED message appeared occurred at Trafalgar Square:

```
N01/091b     0    0  1800
N01/007e     0  149     0
N01/007h     0  202     0
N01/007d     0   70     0

15:20:00 K01 N01/007e ALTERNATIVE STRATEGY
             REQUIRED
15:20:00 K02 CHAIN N01/091b, LINK N01/007e IS
             CRITICAL, WASTED = 1800 LPUs
```

Figure 7 shows the area affected, the links involved in the route are:

N01/091b → N01/007e

The ALTERNATIVE STRATEGY REQUIRED message appeared for link N01/007e three times during a one week period. The percentage saturation on the critical link (N01/007e) is very high while the competing links also show high saturation. In this case, no simple node based solution will have a positive overall effect because the node is completely oversaturated. Possible actions are to reroute vehicles to avoid the congested area and therefore decrease

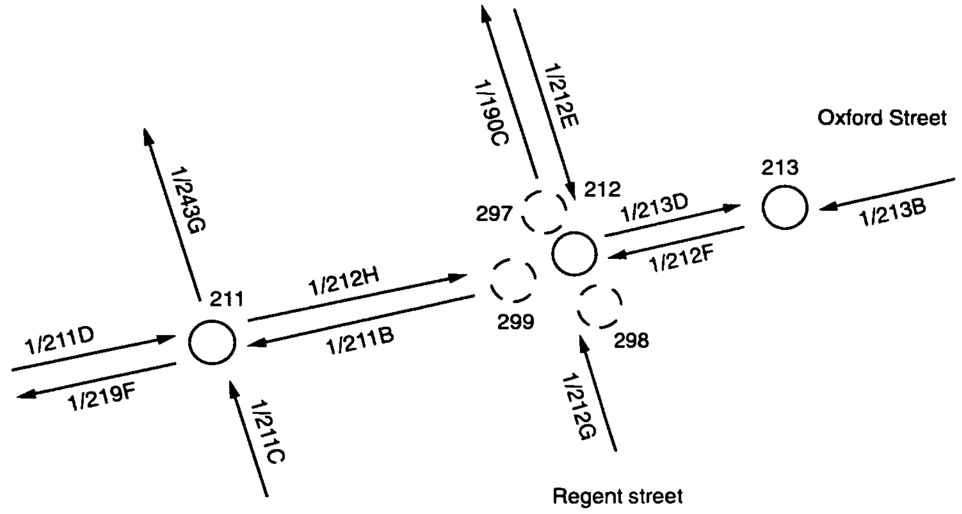

Fig 11. Route N01/213b west across Oxford Circus

demand or set up a gating system to relocate the traffic queues to a less sensitive location.

4.5 CONCLUSIONS FROM ANALYSIS

4.5.1 Benefit of MONACO

It is clear that MONACO is finding consistent and relevant information about the SCOOT area it has been tested on. MONACO has produced information about congestion problems, their location and severity along with the probable cause. The probable cause of the congestion can currently fall into one of four categories, each of which is reported by one or more MONACO messages:

i. Validation problems, where a link has not been validated, this occurred once in the early testing stages of MONACO and the problem has been rectified.

ii. Faulty detector problems, where the loss of traffic information from one or more detectors causes SCOOT to lose some of its responsiveness. MONACO will only detect problems due to a faulty detector when the default logic is unable to prevent congestion. Normally this will only happen if the default parameters are inappropriate. The operator can then either prioritise the detector repair schedule or adjust the default parameters.

iii. Congestion offset problems, where a congestion offset is either required or an existing one is inaccurate.

iv. Saturation problems, these can inform the operator that a particular approach to a junction is being penalised, or more seriously, when a junction is oversaturated. Where saturation problems are indicated, it may be possible to solve them by fine tuning the junctions that MONACO indicates are causing problems. This may be achieved using congestion offsets, more accurate default values for when a link goes faulty or increasing the priority a link has at a junction. If, however, there is persistent oversaturation in the network, these measures will not be sufficient and other solutions will need to be applied. MONACO indicates this by outputting an ALTERNATIVE STRATEGY REQUIRED warning message. This is discussed in the following section.

4.5.2 Alternative strategy required warning

Whenever the ALTERNATIVE STRATEGY REQUIRED warning message is presented to the operator simple adjustments to the SCOOT database, the 'fine tuning' mentioned above, are not sufficient to solve the congestion problem. This warning is given when a particular junction is over saturated on several links, caused by the level of traffic approaching, or exceeding, the maximum the road network can accommodate. In these situations the traffic engineer has the option of employing a traffic management technique such as a SCOOT procedure or gating to restrict the volume of traffic approaching the junction. These methods are discussed below. Should the engineer fail to solve the problem using traffic management techniques then he might have to resort to civil engineering solutions, such as redesigning the junction, with the associated cost and disruption that this would cause. In some situations, junction redesign could be impossible because of the positions of local buildings etc.

Where traffic management solutions are inadequate and civil engineering solutions are not possible, or too expensive, the congestion and resultant delay may just have to be accepted.

4.5.3 Alternative congestion strategies

As mentioned in the previous section, an operator can use traffic management techniques to remedy a congested situation. Two methods of doing this are:

SCOOT Procedures

Procedures are a set of predetermined SCOOT commands for split weighting, offset weighting etc. These commands are grouped together much like a computer program and will change the behaviour of a certain part of the network when started. The traffic engineer, in response to a regular occurrence of congestion, determines the SCOOT commands required to relieve the situation and places these commands into a procedure. From then on, when that particular situation occurs, the operator calls the procedure and the remedial action is carried out. When the problem eases, the procedure is removed.

Procedures are very flexible, they can contain whatever commands the engineer sees fit, some might reduce the green splits on links approaching a problem area thus keeping the problem area relatively free of vehicles, others might force the SCOOT computer to provide ideal offsets for a particular route through the network, providing a 'green wave' effect that can be switched on and off at will. However, procedures have two drawbacks, they are only useful for regular occurrences of congestion and they must be manually started and then stopped again by the operator.

The only location where the ALTERNATIVE STRATEGY REQUIRED warning occurred during the testing of MONACO was for link N01/007e, in Trafalgar Square. This area already has a set of procedures to temporarily change certain database values in order to clear the area of traffic.

Gating

Gating, as described in section 2, is a traffic management tool that automatically responds to congestion. The traffic engineer on noticing a recurrent congestion problem, sets up a gating strategy to automatically adjust green splits leading to or away from the problem area. The engineer identifies a 'bottleneck link', the saturation of which is constantly monitored. If the saturation exceeds a preset level, the green splits on one or more 'gated links' are adjusted, either to restrict the flow of vehicles into the problem area, or to enable a greater flow of vehicles out of it.

Gating is less flexible than procedures because only the green splits on the gated links are changed. However, once set up, it requires no manual observation to operate. As yet no gating systems have been set up in the London SCOOT area monitored by MONACO.

5. AREA WIDE STRATEGIES

As stated in the introduction, capacity maximising signal timings have the potential to increase the throughput of a road network and could form an area wide strategy for the alleviation of congestion. TRL has been investigating the potential of capacity maximising signal timings. The best known method of producing network wide capacity maximising timings is P0 (Smith 1980). However, in this work a less extreme approach has been taken. The SCOOT split optimiser equalises the maximum degrees of saturation at each junction. At the ideal node cycle time, such equisaturation settings approximately minimise delay at the junction. In general, capacity in a network can be increased under congested conditions by moving away from equisaturation timings to signal timings biased in favour of roads with the greatest capacity at a junction. For example, a dual carriageway main road would be run at a lower degree of saturation than a single carriageway side road crossing it. Although P0 would, theoretically, give greater network capacity than the simple approach of biasing high capacity links, the simple approach could be easily implemented using the existing SCOOT weighting facilities. The objective function of the SCOOT Split optimiser could be changed to P0, rather than equi-saturation, however, such a modification could create a large change to the signal timings. The effect of this change would be unpredictable, particularly in the short term, as it would depend on drivers' reactions and rerouting in response to the changes.

To obtain some information on the likely effects of modified timings, CONTRAM has been used to model the changes. From a set of origin-destination flows, CONTRAM models the junction delays and assigns traffic to the modelled road network to minimise the total delay in the network. In assigning traffic to the network, CONTRAM assumes perfect knowledge by the drivers and, therefore, perfect rerouting in response to any change in signal control. Within this project, CONTRAMI, the incident version of CONTRAM, has been used to investigate how the benefit predicted by CONTRAM depends on drivers' rerouting. To model an incident, CONTRAMI starts with an optimum traffic assignment from a CONTRAM run for normal operation of the road network. An incident is then imposed on the network, by reducing the saturation flow of one or more links. Drivers will not have full knowledge of the "new" network with the incident. Therefore, a complete reassignment using CONTRAM is not appropriate. CONTRAMI models driver response by allowing a user defined proportion of drivers to divert in response to abnormally large queues on links. Within this project, CONTRAMI has been used to treat the new signal timings as an "incident" and indicate how the network delay varies as different proportions of drivers respond to the "incident."

5.1 NETWORK

A real road network has been simulated using CONTRAM and CONTRAMI. The network chosen was the CONTRAM network of Kingston-upon-Thames that has been used in previous studies by TRL, including work on how best to deal with incidents in congested conditions. Figure 12 shows the network. It should be noted that, since the roads were coded for CONTRAM, the one-way system in Kingston town centre has been extended and modified. However, the old road network, as used here, is more suitable for our work. The new one-way system so dominates the road network that there is little practical alternative to using it and so little scope for rerouting. The network has 41 nodes and a total of 150 links, of which 43 are controlled by traffic signals. It is large enough to provide realistic scope for rerouting, but small enough to have reasonable computer run times and for the results to be interpreted sensibly.

5.2 DERIVING THE BASE SIGNAL TIMINGS

The timings to be tested have been generated by TRANSYT. However, the signal timings depend on the traffic flows in the network and in turn these flows depend on the routes chosen by drivers in the network. Therefore, it is necessary to iterate CONTRAM to assign the traffic and produce the network flows and TRANSYT to produce appropriate timings for those flows. The computer program INTRESS was used to iterate TRANSYT, operating without any link biases, and CONTRAM. TRANSYT optimisation of splits differs in detail from that of SCOOT as it aims to minimise a weighted sum of delay and stops rather than equalise degrees of saturation. However, the resulting timings are a good approximation to those from normal SCOOT operation for this project. The resulting delays modelled by CONTRAM represent the network delay for minimum delay control with drivers optimising their routes for these signal timings. The iteration of CONTRAM and TRANSYT did not converge very well. The total CONTRAM network journey time fluctuated between runs. It was not part of this project to investigate the stability of the convergence process; such studies would constitute a significant research programme on their own.

TRANSYT and CONTRAM are iterated because it should be possible to reduce delays by optimising the signal timings for the routes drivers take through the network.

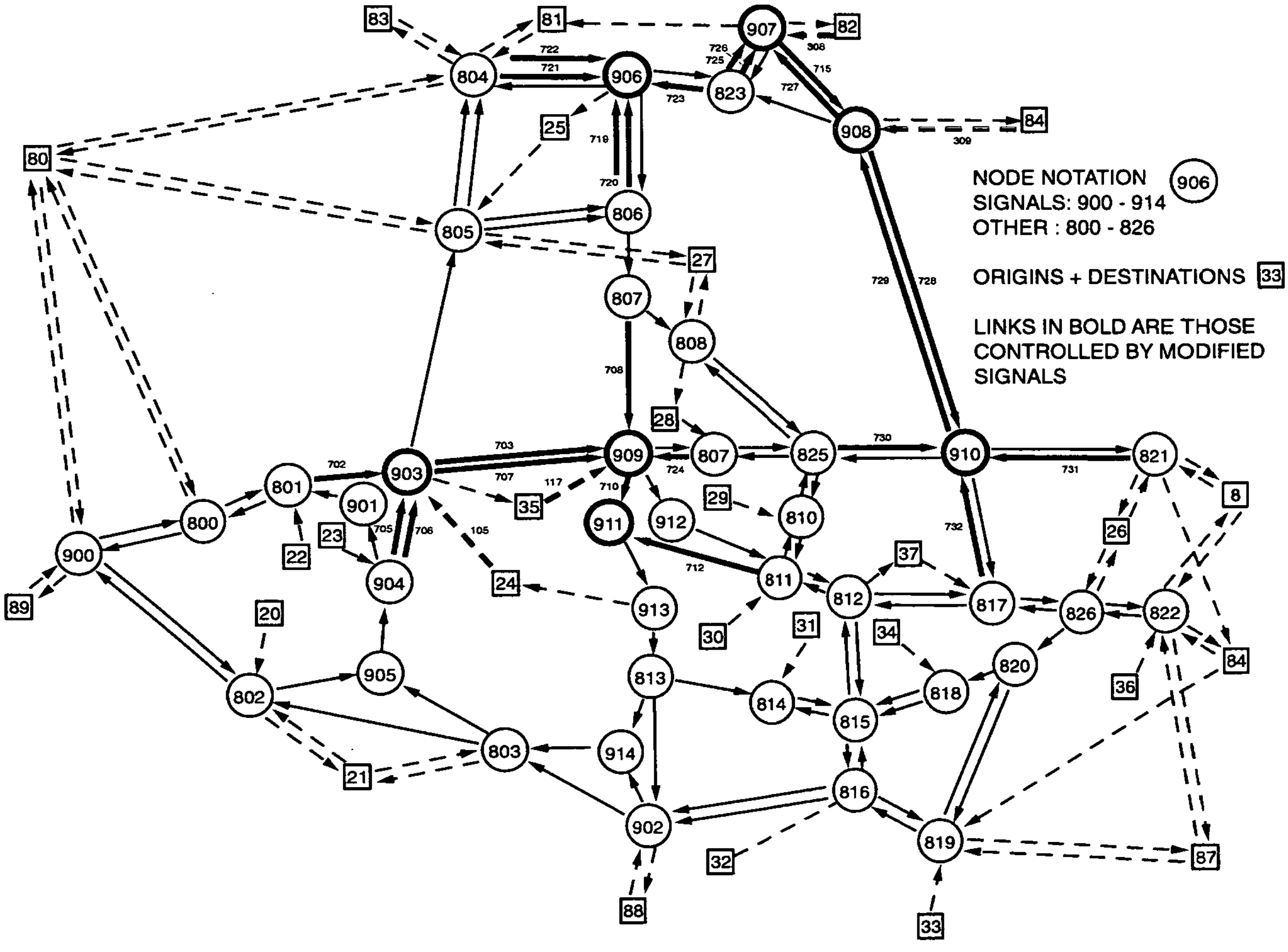

Fig 12. Kingston-upon-Thames CONTRAM network

It was clear from the intermediate results of INTRESS that little extra benefit was being obtained by the later iterations. Although INTRESS had not converged to a stable result, the trend towards reduced delay stopped after a few iterations. Therefore, a set of signal timings and traffic assignment was chosen for the base case. The chosen set corresponded to a low point of the oscillation when the trend to reduced delay had flattened out. This base case was for an optimum traffic assignment to a set of traffic signal settings that produced low network delay. It is very unlikely that the network delay could be significantly reduced by further iterations with the same control method. Therefore, the chosen base case provided a good basis for testing possible improvements to the signal control method.

5.3 BIASING THE SIGNAL TIMINGS

Having calculated the base signal timings and associated delays, the next stage was to look at the effects of biasing timings in favour of the higher capacity signal controlled links. The TRANSYT link bias facility was used to modify the signal timings. TRANSYT works by minimising its "Performance Index," a weighted sum of delay and stops. For a biased link, the delay contribution to the performance index is the product of the calculated delay and the bias. Therefore, a bias of greater than 100% increases the importance of delay on that link and the TRANSYT optimiser will derive greater benefit by reducing delay on that link. Conversely, biases of less than 100% will make delay less important.

To simplify the task of understanding the results, biases were first applied at single junctions to gain an understanding of the effects. The tests at individual junctions were a preliminary stage, before biasing all the signal controlled junctions. Therefore, it was desirable to minimise the computer time needed for these tests. A full run of INTRESS to iterate CONTRAM and TRANSYT was not justified. TRANSYT was run with link bias and the flows from the base case. CONTRAM then took the resulting timings and reassigned the traffic with full driver knowledge of the modified signal timings. A range of biases was tested at each junction, typically up to 200% on favoured links and down to 50% on disbenefitted links.

Considerable care was needed when choosing which links to favour with the bias. A stage based control system is being considered and, therefore, it is necessary to bias in favour of the stage that has the largest saturation flow when the saturation flows for all the links that are green during that stage are summed. The individual link with the highest saturation flow at a junction may not be the one that should receive the favourable bias. This consideration will be the same in a real network as well as a simulated one.

A second point, with an assignment program such as CONTRAM, is that traffic enters the network at fixed entry points. It is not possible to divert such traffic away from its fixed entry link. Therefore, a link that is directly connected to an origin point is unlikely to be a good candidate for a positive bias. A similar consideration will apply to selecting biases in a real network. In a SCOOT network, it is unlikely that large biases will be applied to entry links. The possibilities for diversion may well be less for entry links but, more importantly, giving extra green time to an entry link is likely to be giving extra green to traffic that is entering the congested network; it is more likely that restrictions will be needed for entering traffic.

Having gained an understanding of the effects of biasing single junctions, biases were applied at all the appropriate signal controlled nodes and INTRESS run. The results of this run show the maximum benefit from the change in signal control as CONTRAM was used and assumed full driver knowledge and response to the new method of control. CONTRAMI was then used to test the effects of limited knowledge and response to the new signal timings.

5.4 RESULTS OF CONTRAM NETWORK TESTS

From the results on individual junctions it appeared that weights in the middle of the range tested, 150% on favoured links and 60% on disbenefitted links were most likely to give an overall benefit. However, the effects of network-wide weighting of different strengths needed to be investigated. This was done by varying the weightings in different ways, but using the same weightings at all signalled junctions within any one run. The selection of which links to benefit, bearing in mind the factors discussed in the previous section, had already been done for the individual node tests and no further selections were needed. INTRESS was used to iterate TRANSYT and CONTRAM with the prescribed link weightings in TRANSYT to provide a stable set of routes and signal timings.

As with the base case, the iteration of TRANSYT and CONTRAM did not converge quickly to a single, stable solution. It is possible that the use of weightings does make the process intrinsically less stable because TRANSYT will be more sensitive to delay on the positively weighted links. Again, investigation of the stability was outside the scope of this project.

Figure 13 shows the results of the CONTRAM (i.e. full driver knowledge) tests for different strength weightings over the whole network.

It is clear from Figure 13 that weighting the links to favour the higher capacity roads can reduce the delay in the test network, but as the weightings are increased beyond the optimum the delay increases; the optimum weightings agree with the results on individual junctions. With these flows, the maximum reduction in journey time was just over two per cent, equivalent to a reduction in delay of about seven per cent.

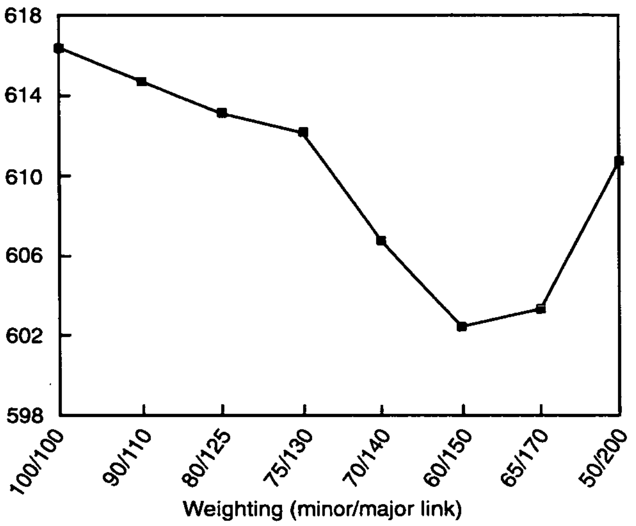

Fig 13. Network journey time (veh-h) against TRANSYT link weighting

5.5 RESULTS OF CONTRAMI TESTS

The incident version of CONTRAM, CONTRAMI, was used to investigate drivers responses to the modified signal timings obtained using the "best" set of weightings, 60/150 from the CONTRAM tests. The changes in signal timings were treated as an "incident," that is some drivers were allowed to change routes in response to queues caused by the new signal timings and the others were constrained to their original routes. This use of CONTRAMI is an extension from its original purpose.

CONTRAM and CONTRAMI respond to the change in signal timings in two different ways. In CONTRAM drivers are assumed to have full knowledge of the network and to choose their best routes. Therefore, when the signal timings are changed a completely new optimisation takes place. With CONTRAMI, the basic routes are taken from a previous CONTRAM run and some drivers are allowed to change from their base route in response to queues encountered on their journey. This is a very different process from re-optimising the whole network in response to the changed signal timings. Therefore, even when all drivers are allowed to divert in CONTRAMI they would not be expected to find as good routes as they would with full knowledge in CONTRAM.

The results of the CONTRAMI tests are given in table 4. The signal timings were for the optimum set of weights from the CONTRAM tests, 60% to the minor links and 150% to the major links.

Several effects can be seen in the results in Table 4. Firstly, even when no drivers are allowed to divert the total network journey time is less with the biased signal timings than with the unbiased ones. Some of this effect may be due to the fluctuations between CONTRAM runs and some to the benefit of giving extra green time to the higher capacity links with the normal traffic pattern. The second effect is a small further reduction in journey time as drivers are allowed to divert within CONTRAMI. The effects of diverting in response to traffic queues are not as great as reassigning the traffic with full knowledge of the new signal timings as shown in the last line of the table.

Methods to increase the capacity of a network become more important as the level of congestion in a network increases. Therefore, the tests on the Kingston-upon-Thames network were repeated with increased traffic flows. It was not possible to increase the flows much without grossly oversaturating the network, hence the flows to the network were uniformly increased to 105 per cent. As for the base case, the effects of different proportions of drivers diverting was tested using CONTRAMI for the increased flow level.

With the increased flow levels, all the benefit of biasing the timings was seen even when no drivers could divert. For this road network, it appears that with the high flows and greater congestion, the optimum traffic assignment for normal signal operation is effectively the same as that for

TABLE 4

Results of CONTRAMI tests

	Percentage eligible for diversion	Total network journey time (veh-h/h)
Base case no weightings	n/a	616.5
CONTRAMI	0%	606.3
CONTRAMI	25%	605.5
CONTRAMI	50%	605.5
CONTRAMI	75%	605.0
CONTRAMI	100%	605.0
CONTRAM	n/a	602.5

the biased timings. All drivers who can benefit by using the high capacity roads are using them with the normal timings. Weighting the signal timings in favour of those roads reduces network journey time, but does not cause significant traffic re-routing.

The use of CONTRAMI to model driver response to modified signal timings is an extension to its normal operation. The details of the internal operation mean that the modelling of driver response will not be as good as for a traffic incident. Therefore the results should be taken as indicative rather than definitive. The conclusions from the modelling are:

> The full CONTRAM runs show that there is potential benefit from using capacity maximising timings.
>
> The level of the benefit in the network testes did not appear very sensitive to the proportion of drivers responding to traffic queues. However, simply responding to queues did not produce the full benefit. Therefore, some time would probably be needed for drivers to become accustomed to the new timings and adjust their journeys appropriately.

5.6 POSSIBLE ON-STREET TESTS

Having explored the possibility of reducing congestion by modifying the signal control, the next stage would be for a future research project to extend this work with practical on-street trials. Such trials would require a co-operative highway authority, controlling a SCOOT network with frequent congestion problems, and a method of assessing the trial. As described in the introduction, the existing SCOOT split weightings could be used to implement a policy of favouring higher capacity roads, with the highway authorities agreement. The SCOOT system would not need to be modified. As a second stage, an experimental SCOOT system with an alternative objective function, P0, for the split optimiser could be developed, if the on-street trials were promising.

Because of the expense of manual data collection, it would be necessary to use automatic data collection for any extended trial. ASTRID could be used to collect SCOOT data for an evaluation, but would need to be used with care. The accuracy of the SCOOT model data has been assessed for evaluation purposes (Carden 1988). That study concluded that the SCOOT data are good for evaluation, but the delay estimates are less reliable in congested condition when links become exit blocked. Therefore, to evaluate the effect of a change of control in congested conditions it would be necessary to use both delay and congestion from ASTRID. As the aim of the modified control strategy is to increase the capacity of the network, the total flow in the controlled region would also be used in the evaluation. The SCOOT flows, which ASTRID stores, are normally good for following changes in the level of traffic flow in the network. However, in this test, the level of congestion may change and affect the average speed of vehicles over the SCOOT loops. As SCOOT uses both flow and occupancy, a change in speed over the loop can change the SCOOT "vehicle flow" when the actual level of flow does not change. Therefore, it would be necessary to calibrate the SCOOT flow against the true flow at the start and end of the survey period. The calibration is straightforward, an observer on-street measures the flow for 1 hour and the calibration factor is the ratio of the measured flow to the SCOOT flow for the same period.

The main difficulty in arranging a test on-street would be the length of time required and the variability of traffic. A learning period of several weeks would be required for drivers to respond to the modified control and change their routes if appropriate. The learning period would be considerably longer if the network became congested only infrequently. Congestion is very variable. Therefore, several weeks' measurements would be needed for each control method. The standard practice to minimise the difference in background traffic conditions between control methods is to alternate the control methods daily, or weekly. However, because of the learning period required for drivers to choose routes appropriate to the control strategy, such alternation would be unacceptable for this survey. Hence, the survey periods for each method of control would be at least one or two months apart and seasonal variations in traffic would confound the analysis of the results. If two similar regions with comparable congestion problems can be found in a city one region could be used as a control to reduce the effects of the variations in traffic on the analysis of the results. However, good controls are difficult to obtain, and the comparability would have to be checked for a period prior to the trial.

5.7 CONCLUSIONS FROM THE AREA WIDE STRATEGY WORK

The modelling work reported here has shown that there can be a potential benefit in modifying the SCOOT split optimiser in congested conditions.

If the aim of the highway operator is to move the maximum amount of traffic under congested conditions then SCOOT weighting should help achieve this aim. It may take some time for drivers to reroute to obtain maximum benefit from the change in control if re-routing is necessary. However, in some congested networks, drivers may already be choosing the high capacity roads. This was the case found for the CONTRAM model of Kingston-upon-Thames with artificially increased flows. Experiments with an operating SCOOT system and extensive automatic data collection, through ASTRID, would be needed to confirm the benefits.

The aims of highway authorities are moving away from reducing vehicle delay to a much more positive attitude to helping public transport vehicles. There has been no investigation in this project of the interaction between changing the normal operation of the SCOOT split optimiser and the operation of bus priority. Such an investigation would be best undertaken as part of a general project on public transport priority in congested conditions. It is often the case, in areas where congestion is likely, that bus routes coincide with main road links to which capacity would be biased by the proposed method. An option might be to use capacity increases for public transport priority, for example by shortening the set-back of bus lanes.

6. CONCLUSIONS

An extensive study has been made of the control of urban traffic in congested conditions.

Simulation studies have been used to identify the conditions of link length and turning movements when offsets are critical to the capacity of a network.

A new quantitative measure of the problems caused by congestion, Wasted Capacity, has been developed.

This measure has been used as the basis of an automatic method to analyse congestion and diagnose the cause. The method has been automated in the MONACO computer program. MONACO found consistent and relevant information about the SCOOT area that it was tested on. It produced information about congestion problems, their location and severity along with the probable cause.

The SCOOT gating facility has been tested and shown to be a valuable traffic management tool. It can control where queues form in a congested road network and can relocate queues to less critical areas of the network. As part of a bus priority scheme, Gating can move queues of delayed vehicles to roads where it is possible to give buses priority, by bus lanes, or other means, to by-pass the queues of congested traffic. Therefore, Gating enables the traffic engineer to give priority to buses in situations where, otherwise, there would not be sufficient road space to reserve a separate lane for buses.

A user guide to SCOOT Gating has been written explaining how to employ Gating.

A global capacity maximising strategy may be of benefit under very congested conditions. However, it may be desirable to allocate the extra capacity to public transport vehicles rather than general traffic where possible.

7. REFERENCES

Carden, P, N Hounsell and R D Bretherton (1988). SCOOT model accuracy. *PTRC Summer Annual Meeting. Bath, UK.* London: PTRC

Department of the Environment (1975). Bus demonstration project - Summary report NO.8 Southampton. London.

Hounsell N and F McLeod (1990). ASTRID: Automatic SCOOT Traffic Information Database. *TRRL Contractor Report 235.* Crowthorne: Transport and Road Research Laboratory.

Smith, M J (1980). A local traffic control policy which automatically maximises the overall travel capacity of an urban network. *Traffic Engineering and Control,* Vol. 21, No. 6, pp 298-302.

Wood, K and R T Baker (1994). User guide to the Gating method of reducing congestion in traffic networks controlled by SCOOT. *TRL project Report PR/TT/101/94.* Crowthorne: Transport Research Laboratory. *(Unpublished report available on direct personal application only. SCOOT users should apply to their system supplier.)*

8. ACKNOWLEDGEMENTS

The work described in this paper was undertaken for the Driver Information and Traffic Management Division of the Department of Transport.

TRL is grateful to the following for their cooperation in providing the Gating data used in this report:

R Morris and J Hammond :-
Southampton City Council,
ROMANSE Project Office

J Landles and A Earl :-
Traffic Control Systems Unit,
Corporation of London

K Huggett and P Drummond :-
Royal Borough of Kingston upon Thames

Insp. Dudley and Sgt. Hills :-
Metropolitan Police,
New Scotland Yard

Sgt. Upex :-
Kingston Police,
Kingston upon Thames

N Adams :-
Surrey County Council, Epsom

MORE INFORMATION

The Transport Research Laboratory has published the following other reports on this area of research:

LR 1014	SCOOT - a traffic responsive method of coordinating signals by P B Hunt, D I Robertson, R D Bretherton and R I Winton. Price £10.
PR41	Urban traffic control, system review by K Wood. Price code J.
TRL135	Accidents at 3-arm traffic signals on urban single-carriageway roads by Marie c Taylor (TRL Ltd), R D Hall (University of Southampton) and K Chatterjee (University of Southampton). Price code N.
TRL217	Traffic congestion incident congestion, D J Bowers, R D Bretherton, G T Bowen, and G T Wall. Price Code E.
TRL220	Review of the potential benefits of Road Transport Telematics. K Perrett, A Stevens, I Wilkinson and P Masurel. Price Code, Vol 1 - Q, Vol. 2 - Y.

If you would like copies, photocopy and fill in the slip below. There is a 20% discount if you take all the reports listed above. Prices include postage and are correct at the time of publication. Please see the enclosed letter for current price code values and handling charge. Enquiries to TRL Library Services, Tel: 01344 770784, Fax: 01344 770193.

To: Publication Sales, TRL Library, PO Box 304, CROWTHORNE, Berkshire, RG45 6YU.
Please send me the following Transport Research Laboratory reports (state report Nos and quantity)

Report no .... *Quantity* ....
Report no .... *Quantity* ....
Report no .... *Quantity* ....

Report no .... *Quantity* ....
Report no .... *Quantity* ....
Report no .... *Quantity* ....

Name ....
Address ....
....
....
Postcode ....
Telephone ....
....
Credit card address (if different from above)......
....

PAYMENT:
- I enclose a cheque for £ ....

payable to TRL Ltd.
- Please debit my Deposit Account

no ....
- Please debit my Credit Card by £...............
- Credit card no ....

Expiry date ....
Signature ....