Akcelik & Associates Pty Ltd
PO Box 1075G, Greythorn, Vic 3104 AUSTRALIA
info@sidrasolutions.com

Management Systems Registered to ISO 9001
ABN 79 088 889 687

REPRINT

Route Control of Traffic in Urban Road Networks Review and Principles

R. AKÇELIK and M. J. MAHER

REFERENCE:

R. AKÇELIK and M. J. MAHER (1977). Route Control of Traffic in Urban Road Networks - Review and Principles. *Transportation Research, Vol. 11, pp 15-24*

NOTE:

This paper is related to the intersection analysis methodology used in the SIDRA INTERSECTION software. Since the publication of this paper, many related aspects of the traffic model have been further developed in later versions of SIDRA INTERSECTION. Though some aspects of this paper may be outdated, this reprint is provided as a record of important aspects of the SIDRA INTERSECTION software, and in order to promote software assessment and further research.

Transpn Res. Vol. 11, pp. 15–24. Pergamon Press 1977. Printed in Great Britain

ROUTE CONTROL OF TRAFFIC IN URBAN ROAD NETWORKS

REVIEW AND PRINCIPLES

R. AKÇELIK
Traffic and Transport Engineering Section, National Capital Development Commission, Canberra, Australia

and

M. J. MAHER
Department of Probability and Statistics, Sheffield University, Sheffield S3 7RH, England

(*Received* 23 *February* 1976)

Abstract—An important type of control which can be used for the control of traffic in urban areas is termed route control. In this paper, route control is defined and a comprehensive review of the related literature is presented. There is no single study in the literature which deals with route control in a comprehensive manner, but it has been possible to derive the relevant information from previous work in a wide range of subjects. The paper describes the basic problems and principles of route control by drawing together the conclusions from the previous work, and shows how they lead to the setting up of a series of route control experiments, using simulation. The simulation model, validation tests and the route control experimental results are given in a subsequent paper.

1. INTRODUCTION

Area traffic control (ATC) systems have been developed for the integrated control of traffic in urban road networks essentially by the use of a central computer. At present ATC systems are generally limited to the optimum control of networks of signalised intersections. It has been well established that the benefits which can be derived from this type of control system are considerable. Also, the installation of ATC systems is easier and less expensive than any alternative measures for the alleviation of urban traffic congestion. In fact, the scope of ATC is wider than mere signal control; the control of critical traffic facilities such as urban motorways and tunnels integrated with the control of adjacent street networks, emergency and special traffic arrangements, control and guidance systems for car parking, and public transport priority, have already been considered or implemented in this context.

An important type of control which can be implemented in ATC systems is the route control of traffic. Route control is the way of distributing traffic over the existing (and/or new) alternative routes in an optimum manner. In other words, route control is a quantity control which aims at obtaining an optimum distribution of traffic in the network by means of a control over the route selection decisions of drivers. Such control has been suggested by a large number of authors, but strong recommendations have come from Gazis (1971) and Brand (1972). After reviewing the present state of traffic signal control techniques, Gazis suggested that probably the limit of improvements which can be obtained by signal control alone is being approached, and the biggest payoff in traffic control might be expected from route control. Along similar lines, Brand suggests that probably this is as far as urban traffic control can be taken within the constraints of preserving manual driver control of the vehicle. Recent research results on traffic signal control support this idea, and it is the aim in this and succeeding papers to describe research which indicates that substantial savings can be obtained from route control implemented in conjunction with signal control, over signal control alone.

The interaction between signal control which is a control in time and route control which is a quantity control and a spatial control is important, since route control means changing flow patterns in the network and this requires the re-adjustment of signal plans. The research reported here was designed to investigate the potential savings of traffic delay and travel time which could be achieved by the use of route control techniques, and to study the implications of such interaction between route control and signal control in the context of area traffic control.

This paper reviews the previous work in the area under several headings, and then draws together the conclusions from this work and shows how they lead to the setting up of a series of route control experiments, using simulation. A second paper then describes the simulation model, the validation tests and the route control experimental results.

Route control of traffic and related studies in the literature can be considered in the following contexts: (1) Area traffic control systems as implemented at present, that is, the centralized computer control of traffic in urban road networks where traffic signal control is dominant. (2) Control of critical traffic facilities such as urban motorways and tunnels integrated with the control of adjacent surface street networks. (3) Static traffic management schemes such as one-way systems, turn prohibitions, tidal flow operations, bus priority systems, and environmental schemes. (4) An experimental electronic route guidance system (ERGS). (5) Driver behaviour.

The review presented here concentrates attention on the first three of these headings since they are the subjects most relevant to the route control experiments to be described later. Some important work in the other areas are mentioned only briefly. Readers interested in a more detailed review of these topics are referred to Akçelik (1974).

2. ROUTE CONTROL AND TRAFFIC SIGNAL CONTROL

From the early days of the development of area traffic control, a great number of authors have suggested the route control of traffic, to be implemented as part of area traffic control systems. Among the authors who have suggested route control and discussed the subject only briefly are Hillier (1962), Cobbe (1964), Irwin (1965), Miller (1965), Baker (1967), Hewton (1967), and the OECD Road Res. Group (1972a). The earliest route control suggestions are due to Gerlough (1954), Rexworthy and King (1960), and Parker (1962), who proposed dynamic route control systems consisting of sensing equipment to measure traffic conditions at critical points in the network and variable direction signs (also placed at key points) controlled by a central computer to re-route traffic from congested routes to less crowded and quicker routes, according to the prevailing traffic conditions in the network. A review of these early works has been given by the Road Res. Lab. (1965), as well as descriptions of a number of real-life "automatic diversion" experiments (including control of holiday traffic and tunnel traffic). Although positive results were obtained in some of these experiments, it was concluded that it was difficult to find sites where suitable alternative routes were available. It should be noted that an analytical approach to the route control problem has not been employed in the studies mentioned so far.

An early study of several traffic signal and route control techniques by means of computer simulation is due to Grimsdale *et al.* (1963). Fixed-time and vehicle-actuated signal control, entry control (imposing a limit on the number of vehicles entering the network by means of traffic signals) and route control (route advice signs placed at suitable points in the network to indicate the directions giving savings in travel time) were considered. Simulation experiments were devised to see to what extent a particular control or combination of controls gave the desired performance in a network of nine intersections. The control objectives were preventing congestion forming inside the network and increasing the system output rate (the rate of vehicles reaching their destinations). It was found that the development of congestion was prevented with the signal/entry control and the signal/entry/route control combinations. Considering the system output rate, the most efficient was the route/entry/signal control combination, and the least efficient was the simple fixed-time signal control. Grimsdale *et al.* also discussed the practical difficulties in implementing such a dynamic control scheme involving a central control computer.

It should be noted that, although the route control criteria used by Grimsdale *et al.* (such as a minimum time saving of 10% and a maximum degree of saturation of about 0.55) seem to be chosen arbitrarily, they represent some route control requirements reasonably, e.g. the limit on the degree of saturation assures the availability of unused capacity on the alternative route. The following points should also be noted about this paper. The signal control methods used in the experiments are not coordinated control methods, and hence the finding that the fixed-time signal control was the least efficient is not surprising. Although entry controls are shown to increase the system output rate and to prevent congestion inside the network, the delays to the vehicles waiting for entry are not mentioned. In addition, it is not clear how the distribution of traffic over the network in the absence of route control was determined. However, Grimsdale *et al.*'s paper should be noted as the first investigation of route control in an analytical manner.

Strong recommendations for route control have come from Gazis (1971) in addition to his suggestions and studies in a number of papers (1965, 1966, 1967, 1968), in a paper with Bermant (1966) and in a paper with Chu (1974). In fact, the term "route control" is attributable to Gazis. After reviewing the present state of traffic signal control techniques, Gazis has suggested that probably the limit of improvements which can be obtained by signal timing alone is being approached, and the time has come to tackle the real problem in traffic control which is the problem of allocating traffic facilities to traffic movements. The biggest payoff in traffic control might be expected from a route plan for an entire metropolitan area, or at least a large reasonably self-contained part of it. Ideally, every driver should be instructed when to start his journey and which route to follow in order to reach his destination. However, a practical system which could be implemented sooner would be a system which allocates traffic facilities to observed traffic, at key points, in such a way as to minimize the overall delay to the drivers. Gazis has also suggested that it is necessary to interrogate individual vehicles as to their destinations, at least on a sampling basis, for the dynamic allocation of traffic facilities.

In an earlier paper, Gazis and Potts (1966) described a method of relieving traffic congestion at critical intersections by coordinated signal and route control of traffic platoons. In this method, by using adjacent streets, the simple intersection is replaced by an intersection complex through which traffic platoons can be guided by properly coordinated traffic signals and diversion signs. The basic principle is that the capacity of an intersection can be increased by diverting some traffic through peripheral roads so that more than one gap in the cross traffic can be exploited.

Gazis and Potts described several ideally constructed examples to illustrate this principle, and discussed the difficulties involved. They suggested that route control would require a diversion sign which would make it obligatory for traffic to choose one of the two alternative routes, for example using green arrows throughout and warning the drivers, as they near the system, with preliminary signs such as "traffic must follow green arrows". Essentially the same route control system has been proposed by Bellis (1966), who called it "switch-point design", as an alternative to the use of a grade-separated intersection.

In another recent paper, Brand (1972) has recommended route control, emphasizing that it allows operational control to substitute for physical capacity by spreading the demand throughout the network. He suggests that probably this is as far as urban traffic control could be taken within the constraints of preserving manual driver control of the vehicle. Brand considered two types of route control, namely (a) advisory route control without O–D information, and (b) mandatory route control with real-time O–D information. He suggested that substantial benefits would come only from mandatory route control with O–D information in real time. In the absence of O–D information, more extensive use of existing methods to provide advisory information on congested routes would be a better strategy than mandatory control. He believed that limited savings could be obtained from this type of route control and from other evolutionary improvements such as computer control of traffic signals. Generally, the success of route control would depend on the network configuration, the pattern of already utilized capacity in the network, and most importantly, the degree to which the diversion of vehicles to more free flowing routes changes network occupancy and O–D travel time individually and in aggregate.

Clark (1973) reports that the West London computer system is already used to control diversion signs in such a way that traffic is discouraged from using a residential road except for short periods when the use of the road is essential (without computer control the alternative would be to route traffic along the residential road at all times).

Recently, a theoretical framework has been described by Allsop (1974) for affecting the number of journeys made, i.e. trip distribution, as well as the routes taken, i.e. traffic assignment, in a network of intersections controlled by coordinated traffic signals.

3. CONTROL OF URBAN MOTORWAYS AND TUNNELS

The route control of traffic using critical facilities such as urban motorways and tunnels plays an important role when the control of such facilities is considered as integrated with the control of adjacent surface street networks. In contrast with area traffic control systems where route control has been widely considered but never applied, the route control of traffic in urban motorway-arterial street networks has been developed progressively, although it is known as motorway (or freeway) control rather than route control. In the early freeway control systems, the objective was the control of peak period traffic to obtain optimal operations on the freeway itself (local control). This has been accomplished by means of entrance ramp control methods, i.e. ramp closure and ramp metering. However, this has resulted in the diversion of some traffic from the freeway to the alternative surface street routes. Then, the control objective has been the optimal use of the freeway and adjacent street system, and more recently "variable route signing" has been considered as a separate method of control, in addition to the ramp control, speed control, and lane control methods. The development of freeway control systems from simple local freeway control towards overall network control has resulted in the production of very useful information for route control purposes, as described below. Although specifically concerned with the urban motorway-surface street networks, the results are equally related to the route control problems in urban networks without motorways.

The problem of the optimum distribution of traffic on an urban motorway-surface street network has been studied by Pinnell (1964a, b), Leygue (1968), Taylor (1968), Yagar (1971a, b) and Hurdle (1974). Pinnell emphasized the need to develop a systems approach to the peak period congestion problem, and suggested that traffic could be controlled or routed at or near its origin, rather than at the critical links such as freeway sections or major arterial streets, in order to achieve optimum system operation by spreading the traffic load over the entire system. Pinnell stated the basic question as "given a set of O–D requirements for a network, how should traffic be routed over this network to obtain optimum system operations", and used a multi-copy LP model to produce the routeings that minimize total system travel time (system-optimising assignment). Pinnell suggested that these optimum routeings should be compared with the shortest path routeings (user-optimising assignment) to find the traffic which should be diverted from the critical links of the network by some type of operational control, such as ramp closures, one-way street operations, advisory signing, and so on.

Indeed, the user and system-optimising assignment principles indicate the way to develop specific route control techniques. A system-optimising route plan (flow pattern) can be computed using an appropriate assignment technique, the existing flow pattern in the network can be compared with this optimum pattern, and the flow differences between the two patterns would be subject to route control. Because the existing traffic pattern may differ from the user-optimising pattern, the existence of non-Wardrop phenomena (Simoes Pereira, 1968) should be taken into account in computing the existing flow pattern, if real-life measurements are not used for this purpose.

Taylor (1968) studied the optimum distribution of traffic over a simple network consisting of a freeway and two parallel arterial streets between a single O–D pair, using assumed speed-flow relationships for each route. The system-optimising distribution was found by maximizing the total system output in terms of vehicle-miles per hour, and the user-optimising distribution was found by equating the speeds on the two routes because equal route lengths were assumed. The results showed that the speed on the freeway should be kept higher than the arterial streets to obtain the system-optimising pattern, which could be accomplished by diverting some freeway vehicles to the arterial streets; the increase in system output would be between 4 and 15%. The percentage increase in efficiency was found to be lower as the system approached capacity. Taylor also investigated the effect of increasing the capacity of arterial streets and found that the maximum output of the system increased substantially with an increase in the capacity of arterials.

In a more recent study, Yagar (1971a, b) investigated alternative routeing patterns, namely the user-optimising, system-optimising, and existing traffic patterns, for a given set of peak period O–D demands on a freeway-arterial street network, in order to obtain an upper bound on the total amount of travel time which might be saved through perfect controls. In this study, a deterministic model was used, and provision was made for the time-varying demand on the system using time slices with constant demand. To find the system-optimising traffic distribution, a single commodity LP formulation was used as an approximation to the multi-commodity formulation. An iterative assignment technique using minimum path method without capacity restraint was used to find the user-optimising flow pattern. These methods were applied to a real network consisting of a signalized arterial with four intersections, one direction of a parallel freeway and its frontage road, and four two-way roads connecting the freeway and the arterial street. Yagar found that the possible amount of decrease in total peak-hour travel time in the system with perfect controls would be in the range from 0 to 26% (349 veh-h/h), but added that this range could be narrowed further with additional studies, because the flows obtained by the LP model were not necessarily completely realistic. Yagar also found that the traffic in the actual network distributed itself in such a way that the total travel time was somewhere between the respective totals resulting from the user- and system-optimising flow patterns, and it corresponded more to the former than the latter. In another theoretical study, Gordon (1972) computed more than 30% saving in total delay in a hypothetical freeway-arterial street network, resulting from the diversion of traffic from the freeway by means of ramp control.

Several authors have reported total system improvements due to entrance ramp controls observed in real-life networks. Generally, the diversion of vehicles from freeway to alternative surface street routes results in decreased travel times on the freeway and increased travel times on the surface street routes, giving a net gain in the total system travel time. The following savings in total network travel time have been reported: May (1965) 21 veh-h/h (2%), Pinnell *et al.* (1967) 360 veh-h/h (25%), Wattleworth *et al.* (1968) 377 veh-h/h (12%), Newman *et al.* (1970) 420 veh-h/h (27%). Wattleworth *et al.* have reported travel time savings on the arterial street system as well, which was a result of the re-timing of traffic signals in the arterial street network to allow for the diverted traffic. When all the theoretical and experimental results from the papers mentioned above are considered, it can be concluded that the possible total travel time savings due to route control of traffic in urban motorway-surface street networks is in the range 0 to 30%. It should be noted that the amount of diversion required was quite small in most of these studies (approximately in the range 100–500 vph).

Most of the controls implemented (or assumed) in the reports mentioned above were static in nature, i.e. based on the assumption of steady-state conditions, with O–D patterns assumed to be constant over such time periods as 15 minutes or half an hour. However, these reports have emphasized that the time variation of traffic demand patterns is very important and the success of controls depend on the accurate estimation of these values. They have also suggested that more dynamic controls, and some form of real-time computer control would be useful for overall system optimisation. The central computer would monitor the system operation and provide appropriate information to drivers so that they could follow the appropriate routes to their destinations either by choice or by enforcement.

Among other authors discussing traffic diversion and the use of variable-message direction signing in urban motorway-surface street networks are Brenner *et al.* (1961), Gervais (1964), Surti and Gervais (1967), McDermott (1967), May (1970), Duff (1971), Wall and Burr (1972), Moskowitz (1973) and Linde (1973). The OECD Road Research Group (1971) has presented a review of the freeway control systems in operation or under development in many countries around the world, and has given detailed suggestions for future research work on route control as well as lane control, speed and headway control, and ramp control in these systems. The OECD report emphasizes that traffic control on a freeway should take account of the road system of which the freeway forms part, and electronic control systems on freeways should be integrated within area traffic control systems, as applicable. Network control would then consist of influencing drivers in order to distribute them over the road system in the best way. Proposals of the report for future research related to route control include complete or selective diversion of traffic from the freeway when the traffic demand exceeds the normal freeway capacity. The OECD report also suggests that providing drivers with information concerning traffic conditions on the freeway and on the alternative routes and suggesting the best possible detour under the given conditions is an essential part of the network control. Methods such as utilization of electronically controlled variable information signs and radio broadcasting are already in use in some countries to provide advance information. Detailed information on the use of variable-message signing for the control of traffic in urban areas is given in a report by the Highway Research Board (1972). The OECD report suggests further research on the type of information required for this purpose: the prediction of journey times on alternative routes, the effectiveness of communication to drivers, and the response of drivers to the given information. A discussion of these subjects are given later.

The problems involved in the control of traffic in urban tunnels (usually river crossings) and adjacent networks with respect to route control are essentialy the same. An early experiment on route control of traffic between two parallel tunnels crossing the river Thames in London has been described by the Road Res. Lab. (1965). This experiment was successful but discontinued, because it was thought that, since both tunnels were congested, diversion to the alternative route was not beneficial (in terms of individual driver travel time). Recently, tunnel traffic control as part of an ATC scheme has been implemented in Liverpool (Stockley *et al.* 1969; Honey 1973). In this scheme, "part-time signs" are used to give drivers an advance notice of relative delays and closures in the two Mersey tunnels, displayed at points on approaches to the tunnels where a change of route decision is still possible. These signs are changed by the central control computer according to the conditions detected on the approaches and inside the tunnels. Entry control is also implemented by means of signal timings.

4. STATIC TRAFFIC MANAGEMENT TECHNIQUES

Such traffic management techniques, which affect the routeing of vehicles in the road network, as the introduction of one-way systems, banning of right-turn, tidal flow operations, control of entry to city centres, bus priority systems and environmental schemes should be considered as some form of route control, although the relevance of these techniques to the route control of traffic has not been recognised in the literature. In fact, a number of points useful for the understanding of route control in general can be derived from the existing knowledge on the use of such systems known as traffic management techniques; see, for example, Kell and Johnson (1970), Cleary (1972), Rae (1970), Duff (1963), Huddart *et al.* (1973), Page (1971), Trench and Slack (1973).

Obviously, the introduction of one-way schemes and prohibition of turning movements are important static route control systems. As emphasized by Hershdorfer (1965), the effects on network performance (e.g. total travel time) should be taken into account when designing such systems. Usually, this kind of systems are implemented in city street networks where traffic signal control is dominant. Hence, it is important to predict flows and turning movements at all intersections affected by the re-routeing of traffic for the re-timing of traffic signals controlling such intersections. An investigation of the effects of both introducing one-way street operations and optimising signal settings on network performance has been reported by Flynn and Siu (1972).

Recently, bus priority systems have been given much emphasis for the improvement of traffic in urban areas (the OECD Road Res. Group 1972b, Transp. and Road Res. Lab. 1973, Huddart *et al.* 1973 and Pretty 1975). It should be noted that the use of bus only streets, i.e. prohibiting all traffic except for buses, as well as other methods of giving priority to buses, namely contra-flow and with flow bus lanes, and priority to buses at traffic signals are also related to route control because they affect the distribution of traffic (a) directly as a result of the re-routeing of vehicles other than buses, or (b) indirectly as a result of the changes in travel conditions to other vehicles. The use of variable turn prohibition, entry prohibition and direction signs as part of a traffic management scheme which involved the introduction of one-way street and bus priority systems has been reported by Earp (1973). The re-routeing of traffic resulting from the introduction of bus priority was taken into account in this scheme.

Another recent example which demonstrates the need to use route control principles in traffic management is the Oxford Street, London, environmental scheme reported by Turner (1973) and Carstens (1973). In this scheme, only buses and taxis are allowed to enter Oxford Street in order to improve pedestrian safety and the shopping environment, and to assist buses. A comprehensive traffic re-assignment was carried out using an all-or-nothing type traffic assignment-simulation model developed by Giannopoulos (1971) in order to predict the new flow pattern in the network and to revise signal timings accordingly, to estimate delays at various intersections, and to compute travel times for cars, taxis and buses and compare them with those prior to the introduction of the scheme. More recently, Coombe *et al.* (1974) has reported a study of the wider use of bus and taxi-only streets, in which re-routeing of traffic was considered and a more systematical analysis made using another traffic model developed for this purpose. It should be noted that, although analytical approaches were employed in these studies prior to the introduction of new traffic management schemes in order to predict their results at a detailed level, the determination of the schemes and the resulting flow patterns was based on engineering judgement rather than an optimization technique, which should ideally be employed in this kind of traffic management systems.

5. AN EXPERIMENTAL ROUTE GUIDANCE SYSTEM

Perhaps the only comprehensive work on route control has been the development of an experimental electronic route guidance system, ERGS. The development of the system from its conception to its future applications have been described by Rosen *et al.* (1970). The system works as follows: When the driver starts on a trip, he looks up the intersection nearest his destination in a directory, and enters the corresponding code word for his destination into the vehicle equipment. As he approaches an instrumented intersection, an antenna in the pavement activates the ERGS equipment in the vehicle. The vehicle's ERGS unit transmits the destination code to the roadside equipment via a communication loop. The roadside equipment processes the code, and communicates an appropriate routeing instruction back to the vehicle. The received instruction (turn right, turn left, etc.) is displayed visually. A similar two-way exchange of information occurs at each instrumented intersection during the trip. The system is destination oriented, if a driver failed to follow or elected to ignore a routeing instruction, the next instrumented interesection would route him to his destination from that point.

Rosen *et al.* have discussed static and dynamic routeing problems, and the use of ERGS for these purposes. The static routeing problem is exemplified by the difficulties unfamiliar drivers experience while traversing the highway network using only conventional maps and signs, and results in driver uncertainty at decision points and excessive travel in the system. The dynamic routeing problem is exemplified by the difficulties both familiar and unfamiliar drivers have while seeking best routes through the network wherein traffic conditions are varying with time. Dynamic routeing problems arise primarily because individual drivers lack knowledge of alternative routes and traffic control systems do not include comprehensive surveillance capability and responsiveness to varying demand. A more detailed fundamental analysis of the routeing problem as considered in ERGS has been given by Stephens *et al.* (1968). An investigation of the user benefits associated with alternative static electronic guidance systems for a sample network has been reported by Bellomo and Young (1972).

6. DRIVER BEHAVIOUR

The degree of acceptance of route control measures by drivers and the information requirements of drivers for route control purposes are obviously matters of great importance. In order to predict the likely success of any traffic control scheme, some knowledge of driver response to road signs, advisory messages and diversions is necessary. In the literature there is a limited number of studies on this subject. The results of a questionnaire survey of the attitudes and behaviour of drivers regarding route diversion in a freeway-arterial street network have been reported by Heathington *et al.* (1971). In this study, the driver was placed in a hypothetical situation and given traffic information about the conditions on the freeway (but not on any alternative routes). The responses of drivers regarding the voluntary diversion from their normal routes were related to perceived attitudes and were in part based on past behaviour. The following are some of the results of this study:

(a) When reporting on their current behaviour, the drivers indicated that they diverted more because of an accident than because of heavy congestion.

(b) The drivers were more receptive to diversion to avoid a delay or to save travel time in the journey to work than on the journey home, and generally they did not attach much significance to small time savings.

(c) Statistically there was no difference between the proportion of freeway and non-freeway drivers diverting to avoid a delay on the journey to work.

A fundamental discussion of the subject of driver information needs has been given by Allen *et al.* (1971). At present, the ways of providing route information to drivers are (a) outside or inside the vehicle, (b) static or variable-message, and (c) visual or aural. Reviews of such systems have been given by Stephens *et al.* (1968) and the OECD Road Res. Group (1971). In general, the existing systems are based on static visual signing (direction signs, lane markings, turn prohibition signs, etc.). Variable-message direction signing is finding wider use in recent years as mentioned above. Studies to evaluate various information systems have been reported by Dudek *et al.* (1971), Dudek and Jones (1971), and Hoff (1971). Generally, it was found that drivers would react to real-time traffic information by changing their routes.

One of the most important problems in route control is the understanding of drivers' decision-making process for route selection. This is a fundamental problem for the traffic assignment techniques, and also very important in the design and evaluation of traffic control systems in general. Generally, travel time has been employed as the route selection criterion in most traffic assignment methods, and it is the most widely used measure of effectiveness for traffic control purposes. In spite of the extensive use of route selection criteria in traffic studies, there has been limited research into the actual behaviour of drivers regarding route selection. Several important research results are summarized below.

Wachs (1967), Benshoof (1970), Ueberschaer (1971), Ratcliffe (1972), Tagliacozzo and Pirzio (1973) have reported the results of studies which involved both questionnaire surveys and measurements of travel times, distances, and flows on the actual network in order to understand the major elements of the route selection process in urban areas. A number of route selection criteria, namely minimum travel time (quickest route), shortest distance, minimum cost, less traffic, fewer stops, greater safety, and so on, were investigated by these authors. From the results of these studies it can be concluded that travel time remains the best single criterion for route selection. It is also the most convenient criterion for reasons of accuracy and practicality of measurements. In addition, other parameters are always related to the travel time, e.g. travel costs can be expressed in terms of travel time. It is also useful to make a distinction between the drivers' route selection decisions and the route selection criterion used for simulating traffic distribution. As suggested by Wattleworth in the discussion of a paper by Heathington *et al.* (1970), it is possible that drivers do not evaluate alternative routes in terms of travel time or delay, but rather that they tend to think in terms of speed or some other parameter such as comfort or convenience. However, this is not necessarily to say that the drivers do not end up taking the minimum travel time route. On the other hand, it is important to note that route selection is a stochastic process, and that there will be some difference between an existing flow pattern and the corresponding user-optimising pattern. This could be allowed for by assuming that drivers finally choose routes which differ from the minimum time route, in totality, by a certain amount. This fact was taken into account in the design of route control experiments reported in a subsequent paper.

7. PRINCIPLES OF ROUTE CONTROL

Various problems related to the route control of traffic in urban areas have been indicated throughout the previous sections. An overall view is presented in this section, which leads to the principles and assumptions used in the simulation experiments for the investigation of route control.

In general terms, a traffic control system consists of the road network and the traffic demand on the network. The traffic demand is expressed in terms of origin and destination requirements and the amount of flows. This demand is served by the road network which consists of links to facilitate directional movement of traffic and nodes where turning movements are executed. The nodes are major decision points for route selection during the journey from origin to destination. A route consists of a sequence of links between the origin and destination. Individual movements between origins and destinations in the network will result in a traffic flow (distribution) pattern which can be expressed by flow rates and turning percentages for each link.

In the context of area traffic control, traffic signal control can be accepted as the fundamental control for the purposes of optimising traffic operations because the signal timings determine the capacity and delay characteristics of the links, and consequently, link travel times depend on signal control parameters. The computation of a network signal timing plan is based on a given flow pattern, i.e. it assumes that link flows are uncontrollable variables. On the other hand, a flow pattern is a result of individual drivers' route selection decisions which depend on route travel times (sum of link travel times and turn penalties), and hence it is in turn a function of the network signal settings. The re-distributional effects of signal control in the absence of route control has been discussed in detail in the paper by Maher and Akçelik (1975).

In contrast with traffic signal control, route control treats link flows as controllable variables, and aims at finding an optimum flow pattern in the network; in this case, the O–D demands are uncontrollable variables. In the absence of route control, each driver will ideally try to minimize his own route travel time, and this will result in a user-optimising traffic pattern, under a given set of traffic signal timings. Because of inaccuracies in the drivers' route selection decisions, the actual traffic pattern may be different from the user-optimising pattern. In most cases, this pattern will be different from the system-optimising traffic pattern, because some individual drivers will be introducing larger travel times to the overall system than their individual savings by choosing the quickest route with respect to their own journeys. This will result in sub-optimisation, which is not necessarily equivalent to the overall system optimisation. As a result of the sub-optimisation of individual movements over alternative routes between origin and destination, the major routes will be over-utilized whereas the secondary routes will be under-utilized. In the case of overall system optimisation, the capacities of alternative routes will be used in a more balanced manner.

Two types of route control strategies can be derived from the existing literature related to route control which has been reviewed in the previous sections:

(a) User-optimising route control strategy: Route control would aim at diverting individual vehicles to the quickest routes assuming that, in the absence of route control, drivers are not able to choose their minimum routes because they can not evaluate traffic conditions on alternative routes accurately. Most of the route control suggestions made with relation to ATC systems, some of those with relation to urban motorway control systems, and the principle of ERGS fall into this category which represents the existing understanding of route control.

(b) System-optimising route control strategy: Route control would aim at diverting vehicles to alternative routes so as to minimise total travel time to all drivers in the network. Route control suggestions in relation to traffic assignment studies, and most of the diversion systems for urban motorway control fall into this category. On the other hand, the existing network signal control techniques aim at minimising total system travel time without considering redistributional effects of signal settings, which approaches a user-optimising solution.

The objective of a route control system should be to obtain a system-optimising traffic flow pattern, rather than a user-optimising flow pattern. The problem of selecting performance criteria is the same as for traffic signal control. The total system travel time is the most convenient measure of performance for route control as well. Therefore, more precisely, the objective of a route control system will be the minimization of total network travel time, by controlling the distribution of traffic over alternative routes. It is often suggested in the literature that, as a result of this policy, there will be increases in individual travel times of diverted vehicles. This may be true in certain cases, but it is also possible that the travel times of both diverted and undiverted vehicles are decreased by the re-timing of traffic signals. This is because the capacities of alternative routes are better utilized as a result of employing such a system-optimising control policy. These aspects of route control are demonstrated in the route control experiments reported in a subsequent paper.

Thus, route control will in turn effect the traffic signal timings, because the traffic distribution pattern will be changed. Therefore the interaction between the signal control and route control should be understood, if they are to be used together in the context of area traffic control. Essentially, traffic signal control is the control of demand in time, while route control is the control of demand on space. However, signal control results in the distribution of demand on space as well, by affecting link travel time characteristics; and route control affects the distribution of demand in time as well, because diverted flows will be subject to different signal controls. Hence, the interaction between signal control and route control can be explained by the fact that both controls jointly affect the distribution of demand on the network in both time and space.

In order to estimate the benefits of route control, it is necessary to predict the existing the objective flow patterns so that the amount of diversion required and possible travel time savings can be determined. Existing traffic assignment techniques are useful in developing the methods of predicting the user- and system-optimising flow patterns, but the model must have the fine details of a network as used in network signal timing methods.

Computer simulation can be used as a reliable approach for both determining alternative flow patterns and evaluating their effects on the system and individual vehicle performance. A traffic assignment-simulation model was developed for this purpose during this research. As an alternative (or complement) to traffic assignment, search methods can be employed for computing user- and system-optimising flow patterns. A mathematical analysis of route control in simple network situations has been given and the need to a simulation approach in complex network situations has been demonstrated in the work of Akçelik (1974).

Before implementing a route control plan, the amount of travel time savings should be predicted. This would depend critically on the degree of utilization of each of the available alternative routes. These would in turn depend on the geometry of the network, the demand-capacity situation on the network scale, and the other controls used in the network. The amount of savings could vary from zero to a certain maximum value depending on these conditions. The mode and the degree of route control to be implemented should be decided on this basis.

Two types of route control are possible, namely advisory and mandatory. The selection of one of these should be based on various considerations. Generally, each control type has advantages and disadvantages. The advisory route control would have the advantages of user acceptance and compatibility with existing controls. But it would probably result in a traffic flow pattern between the user-optimising and system-optimising patterns, therefore the savings from advisory route control would be limited compared with mandatory route control. In addition, advisory route control systems would require the provision of more information to drivers including road and traffic conditions, routeing and directional guidance. Theoretically, maximum savings could be obtained by the use of mandatory route control which aims at the system-optimising flow pattern. The implementation would be easier to accomplish by means of simple diversions signs. However, there may be practical difficulties, for example, because of the inaccuracy in O–D data, or because of a sudden change in traffic conditions, a congestion might be created on a route where traffic has been diverted from other routes. Or, even under the best operating conditions with mandatory route control, some vehicles may be forced to take routes with longer travel times for the sake of overall travel time savings in the system, which will not have the advantage of user acceptance. However, a limit could be imposed on the amount of acceptable increase in the travel times due to diversion. It might be that a significant improvement could be obtained in the system by increasing the travel times of a small number of drivers only slightly.

In addition, it may be necessary not to interfere with public transport and some essential services, and advisory route control would be advantageous in this case. Otherwise, mandatory route control which is applied to certain classes of vehicles (e.g. only private cars) might be considered. It is important that the choice of mandatory or advisory route control should be based on the route control strategy accepted, and the type of information to be displayed should be selected accordingly. For example, if in-vehicle display equipment is to be used as in ERGS, it could be difficult to use mandatory routeing and force drivers towards a system-optimising flow pattern, contrary to the suggestion by Brand (1972). This is simply because, it would not be practical to monitor the movements of individual drivers who might be instructed to take routes with travel times longer than their minimum time routes.

Another important problem in developing and implementing route control systems is the selection of a static or dynamic route control strategy. A static route control system would be based on historical data assuming steady-state conditions during the control period, whereas a dynamic route control system would require real-time data on O–D demands and travel times in order to update continuously optimum route plans and the corresponding routeing information. In the literature, almost every author discussing this subject considers the use of dynamic route control methods and expects large savings over static route control. The situation is very similar to the early discussions on network traffic signal control prior to the implementation of such controls. Almost everybody expected that dynamic traffic signal control techniques would result in large savings, but the area traffic control experiments carried out so far have shown that dynamic signal control methods do not produce better results than static control methods. In addition, both the hardware and the on-line computing time requirements are heavier for dynamic control schemes; therefore static systems have the advantage of cost-effectiveness over dynamic systems. Problems of control stability in dynamic systems, which is partly due to the time lag in reacting to changing traffic conditions may be more important in route control systems, because of the quantity control of traffic flows. Also the frequent change of routeing information and/or instructions may not be desirable, at least from the user acceptance viewpoint.

The interaction between signal control and route control should be considered when the responsiveness of controls is to be decided. Because one set of control parameters must be fixed, at least one control should be static, and hence three control strategies would be possible: (1) static signal control/dynamic route control, (2) dynamic signal control/static route control, and (3) both static signal and route controls. At the present, there is little (or no) information about the implementation of route control systems in real-life conditions, and following the example of signal control, it would be a better strategy to consider the implementation of static route control at the beginning. Fixed route and signal plans established according to the time of day (each peak and non-peak period) could be selected from a library of such plans. Considering the changes in origin-destination patterns over shorter time intervals during peak periods, control plans for shorter intervals such as 30-min or 1-hr could be prepared. If the implementation of route control during off-peak periods is not desirable for practical reasons, blank or fixed route signs could be used during these periods.

Also for practical reasons (cost, driver acceptance, etc.) the application of route control may be limited only to the critical routes in the network. These routes can be selected on the basis of the amount of congestion and the contribution to the total travel time saving. The latter can be obtained by comparing the existing and the optimum traffic flow patterns. The limited application of route control may be justified when other reasons, in addition to cost-effectiveness and driver acceptance, are taken into account, namely the availability of alternative routes with sufficient capacity in the network, and the adverse environmental effects on the alternative routes. The latter may be an important factor in the planning of route control schemes, because it may be undesirable to divert more vehicles to the routes which have residential and similar characteristics. In fact, route control techniques might be used to improve the environmental conditions of such routes by limiting the number of vehicles using them.

8. CONCLUSION

Although there is no single study in the literature which deals with route control in a comprehensive manner, sufficient information and guidance can be derived from the papers referred to in the previous sections to decide on, and to define, the nature of the route control problem to be investigated in the series of simulation experiments. It was decided, therefore, that the most relevant and useful context would be:

(i) An urban road network, in which traffic signals are the dominant control mechanism.

(ii) Static route control is applied over a sufficiently long time period to allow steady-state conditions to be established (anologous with fixed-time ATC applied by time-of-day).

(iii) Route signals situated outside the vehicle (i.e. variable direction signs) are the means by which route control is accomplished; manual control of the vehicle is preserved.

(iv) Average travel time is used as the route selection criterion of individual drivers, and total network travel time is used as the measure of performance of any control system.

The objective of the experiments are to find the user- and system-optimising control solutions when the network structure and the O–D data are given, to investigate the interaction between route control and signal control, and to determine possible travel time and delay savings from various types and degrees of route control under various flow and network conditions. In order to achieve these objectives, it was decided to use a semi-macroscopic stochastic simulation-assignment model. The model is described and the results from these experiments are reported in a subsequent paper.

REFERENCES

Akçelik R. (1974) Traffic signal and route control in urban areas. Unpublished Ph.D. Thesis. University of Leeds.

Allen T. M., Lunenfeld H. and Alexander G. J. (1971) Driver information needs. *Highw. Res. Record* **366**, 102–115.

Allsop R. E. (1974) Some possibilities for using traffic control to influence trip distribution and route choice. In *Transportation and Traffic Theory* (Edited by D. J. Buckley). Elsevier North-Holland, Amsterdam.

Baker R. F. (1967) A practical view of the systems approach. *Traffic Q.* **21**, 471–487.

Bellis W. R. (1966) Intersection design: switch point. *Highw. Res. Record* **105**, 8–20.

Bellomo S. J. and Young G. K. (1972) Benefit measurements of highway electronic route guidance. *Transpn Engng J., A.S.C.E.* **98**, (TE1), 51–69.

Benshoof J. A. (1970) Characteristics of drivers' route selection behaviour. *Traff. Engng Control* **11**, 604–606, 609.

Brand D. (1972) Urban traffic control: how far can we take it. *Traffic Engng* **42**, 50–55, 65.

Brenner R., Telford E. and Frischer D. (1961) A quantitative evaluation of traffic in a complex freeway network. *Highw. Res. Bd Bulletin* **291**, 163–206.

Carstens R. S. (1973) Oxford Street experimental traffic scheme. *Traff. Engng Control* **15**, 79–81, 85.

Clark D. G. (1973) Area traffic control. In *Traffic Management.* P.T.R.C. Seminar Proceedings. London.

Cleary E. F. (1972) The challenges of systems management. *Traffic Engng* **43**, 26–30.

Cobbe B. M. (1964) Development of traffic signals for area traffic control. *Traff. Engng Control* **5**, 543–546.

Coombe R. D. and Buchanan C. M. (1974) Bus priority in Inner London. A series of two papers. *Traff. Engng Control* **15**, 480–486, 575–580.

Dudek C. L., Messer C. J. and Jones H. B. (1971) Study of design considerations for real-time freeway information systems. *Highw. Res. Record* **363**, 1–17.

Dudek C. L. and Jones H. B. (1971) Evaluation of real-time visual information displays for urban freeways. *Highw. Res. Record* **366**, 64–78.

Duff J. T. (1963) One-way streets. *Traff. Engng Control* **4**, 518–520.

Duff J. T. (1971) Accomplishments in freeway operations outside the United States. *Highw. Res. Record* **368**, 9–25.

Earp J. H. (1973) Bitterne traffic scheme, Southampton. In *T.R.R.L. Rept LR* 570, pp. 99–106. Transp. and Road Res. Lab. Crowthorne, Berkshire.

Flynn T. E. and Siu C. M. (1972) Comparison of traffic circulation alternatives through SIGOP. *Highw. Res. Record* **321**, 46–59.

Gazis D. C. (1965) Traffic control, time-space diagrams, and networks. In *Traffic Control Theory and Instrumentation* (Edited by T. R. Horton), pp. 47–63. Plenum Press, New York.

Gazis D. C. and Bermant O. (1966) Dynamic traffic control systems and the San José experiment. *8th Int. Study Week in Traff. Engng, Theme V: Area Control of Traffic.* O.T.A., London.

Gazis D. C. and Potts R. B. (1966) Route control at critical intersections. *Proc. Aust. Road Res. Bd* **3**, Part 1, 354–363.

Gazis D. C. (1967) Mathematical theory of automobile traffic. *Science* **157**, 273–281.

Gazis D. C. (1968) Traffic flow theory. *Proc. Inst. Electrical and Electronic Engrs* **56**, 459–471.

Gazis D. C. (1971) Traffic control: from hand signals to computers. *Proc. Inst. Electrical and Electronic Engrs* **59**, 1090–1099.

Gazis D. C. and Chu K. (1974) Dynamic allocation of parallel congested traffic channels. In *Transportation and Traffic Theory* (Edited by D. J. Buckley). Elsevier North-Holland, Amsterdam.

Gerlough D. L. (1954) Control of automative vehicular traffic on an area basis. Unpublished Tech. Memo. LA-5A. Inst. Transpn and Traff. Engng, Univ. of California, Los Angeles.

Gervais E. F. (1964) Optimization of freeway traffic by ramp control. *Highw. Res. Record* **59**, 104–118.

Giannopoulos G. A. (1971) Towards a comprehensive assessment of traffic management. *Traff. Engng Control* **13**, 105–107.

Gordon R. L. (1972) A concept of deterrence controls for freeway to freeway interchanges. *Traff. Engng* **42**, 23–27..

Grimsdale R. L., Mathers R. W. and Sumner F. H. (1963) An investigation of computer-controlled traffic signals by simulation. *Proc. Instn Civ. Engrs* **25**, 183–192.

Heathington K. W., Worrall R. D. and Hoff G. C. (1970) An analysis of driver preferences for alternative visual information displays. *Highw. Res. Record* **303**, 1–16.

Heathington K. W., Worrall R. D. and Hoff G. C. (1971) Attitudes and behaviour of drivers regarding route diversion. *Highw. Res. Record* **363**, 18–26.

Hershdorfer A. M. (1965) Optimal routing of urban traffic. Unpublished Ph.D. Thesis. Massachusetts Institute of Technology.

Hewton J. T. (1967) The Metropolitan Toronto signal system. In *Area Control of Road Traffic.* Symposium Proceedings. Instn Civ. Engrs. London.

Hillier J. A. (1962) A review of developments in area traffic control. *Proc. Aust. Road Res. Bd* **1**, Part 1, 416–429.

Highway Research Board (1972) The changeable-message concept of traffic control. *Special Rept* **129**.

Hoff G. C. (1971) A comparison of selected traffic information devices. *Highw. Res. Record* **366**, 116–129.

Honey D. W. (1973) Liverpool's wide-area traffic control scheme. *Traff. Engng Control* **15**, 302–304.

Huddart K. W. *et al.* (1973) Bus priority in Greater London. A series of six articles. *Traff. Engng Control* **14**.

Hurdle V. F. (1974) The effect of queueing on traffic assignment in a simple road network. In *Transportation and Traffic Theory* (Edited by D. J. Buckley). Elsevier North-Holland, Amsterdam.

Irwin N. A. (1965) The Toronto computer-controlled traffic signal system. In *Traffic Control Theory and Instrumentation* (Edited by T. R. Horton), pp. 209–218. Plenum Press, New York.

Kell J. H. and Johnson B. C. (1970) Optimising street operations through traffic regulations and control. *N.C.H.R.P. Rept* **110**. Highw. Res. Bd, Washington, D.C.

Leygue F. (1968) Operating motorways in urban areas *9th International Study Week in Traff. Engng.* Theme VIII. O.T.A., London.
Linde R. (1973) Planning for recreational traffic. *Traff. Engng* **43**, 24–25, 52–56, 62.
Maher M. J. and Akçelik R. (1975) The re-distributional effects of an area traffic control policy. *Traff. Engng Control* **16**, 383–385.
May A. D. (1965) Experimentation with manual and automatic ramp control. *Highw. Res. Record* **59**, 9–38.
May A. D. (1970) Traffic control on urban motorways in the United States: a European assessment. *Traff. Engng Control* **12**, 370–372, 379.
McDermott J. M. (1967) Operational effects of automatic ramp control on network traffic. *Highw. Res. Record* **202**, 1–31.
Miller A. J. (1965) Computers in the analysis and control of traffic. *Traffic Q.* **19**, 556–572.
Moskowitz K. (1973) Motorway surveillance and control. *Traff. Engng Control* **14**, 526–528.
Newman L., Dunnet A. M. and Meis G. J. (1970) An evaluation of ramp control on the Harbor Freeway in Los Angeles. *Highw. Res. Record* **363**, 27–42.
O.E.C.D. Road Research Group (1971) Electronic aids for freeway operation. Road Res. Programme. O.E.C.D., Paris.
O.E.C.D. Road Research Group (1972a) Area traffic control systems. Road Res. Programme. O.E.C.D., Paris.
O.E.C.D. Road Research Group (1972b) Optimization of bus operation in urban areas. Road Res. Programme. O.E.C.D., Paris.
Page M. H. W. (1971) Tidal flow in Portsmouth. *Traff. Engng Control* **12**, 473–476.
Parker E. (1962) Trafficometry. *Traff. Engng Control* **4**, 318–323.
Pinnell C. (1964a) Development of a linear programming technique for a systems analysis study of arterial street operation. Unpublished Ph.D. Dissertation. Texas A & M University.
Pinnell C. (1964b) Optimum distribution of traffic over a capacitated street network. Res. Rept 24-2. Texas Transpn Inst. Texas A & M University. College Station, Texas.
Pinnell C. *et al.* (1967) Evaluation of entrance ramp control on a six-mile freeway section. *Highw. Res. Record* **157**, 22–76.
Pretty R. L. (1975) The reduction of traffic delays for buses. *Aust. Road Res. Bd Rept No.* **39**.
Rae T. E. (1970) The preparation of a traffic and transportation plan. *Traff. Engng Control* **12**, 66–71.
Ratcliffe E. P. (1972) A comparison of drivers' route choice criteria and those used in current assignment processes. *Traff. Engng Control* **13**, 526–529.
Rexworthy D. R. and King D. W. W. (1960) Freeing city traffic by computer. *New Scientist* **7**, 197–199.
Road Research Laboratory (1965) Automatic diversion of traffic. In *Research on Road Traffic*, 351–357. H.M.S.O., London.
Rosen D. A., Mammano F. J. and Favout R. (1970) An electronic route guidance system for highway vehicles. *Inst. Electrical and Electronic Engrs Transactions* **VT-19**, 143–152.
Simoes Pereira J. M. S. (1968) Some comments on the traffic assignment problem, *Proc. Fourth International Symposium on the Theory of Traffic Flow* (Edited by W. Leutzbach and P. Baron). *Strassenbau und Strassenverkehrstechnik* **86**, 220–223.
Stephens B. W. *et al.* (1968) Third generation destination signing: an electronic route guidance system. *Highw. Res. Record* **265**, 1–18.
Stockley N. H., Evans A. J. R. and Davison A. (1969) City of Liverpool: central area congestion control scheme. *Proc. Instn Civ. Engrs* **44**, 119–135.
Surti V. H. and Gervais E. F. (1967) Peak period confort and service evaluation of an urban freeway and an alternate surface street. *Highw. Res. Record* **157**, 144–178.
Tagliacozzo F. and Pirzio F. (1973) Assignment models and urban path selection criteria. *Transpn Res.* **7**, 313–329.
Taylor W. C. (1968) Optimization of traffic flow splits. *Highw. Res. Record* **230**, 60–77.
Trench S. and Slack J. A. (1973) Nottingham's new transport policy. *Traff. Engng Control* **15**, 200–204.
Transport and Road Research Laboratory (1973) Bus priority. Symposium Proceedings. *T.R.R.L. Rept* **LR 570**. Crowthorne, Berkshire.
Turner E. D. (1973) Oxford Street experiment: environmental improvement. In *Traffic Management* P.T.R.C. Seminar Proceedings, London.
Ueberschaer M. H. (1971) Choice of routes on urban networks for the journey to work. *Highw. Res. Record* **369**, 228–238.
Wachs M. (1967) Relationships between drivers' attitudes toward alternate routes and driver and route characteristics. *Highw. Res. Record* **197**, 70–87.
Wall J. S. and Burr M. A. (1972) Aston Expressway signing and signalling system. *Traff. Engng Control* **14**, 370–375, 381.
Wattleworth J. A., Courage K. G. and Carwell J. D. (1968) An evaluation of two types of freeway control systems. Res. Rept 488-6. Texas Transpn Inst. Texas A & M University. College Station, Texas.
Yagar S. (1971a) A study of the travel patterns in a corridor with reference to the assignment principles of Wardrop. In *Traffic Flow and Transportation* (Edited by G. F. Newell) American Elsevier, New York.
Yagar S. (1971b) Dynamic traffic assignment by individual path minimization and queuing. *Transpn Res.* **5**, 179–196.