A MODEL TO EVALUATE
PUBLIC TRANSPORT SYSTEMS
IN URBAN AREAS

by

Eduardo Duarte C.

Thesis submitted in fulfilment of the requirements
for the Degree of Doctor of Philosophy in the
Department of Town and Regional Planning,
University of Glasgow, Glasgow, Scotland

July,1980

TO LIA

CONTENTS

LIST OF TABLES

LIST OF FIGURES. Page

LIST OF FLOWCHARTS

LIST OF MAPS.

ACKNOWLEDGEMENTS

The help received from several individuals and organizations is most gratefully acknowledged. First of all I wish to express my gratitude to my supervisor Mr. D.M. Allan for his kind assistance throughout this research in both academic and non-academic matters. To Professor G.C. Cameron, former Head of the Department. I am also grateful to Dr. A.D. Ochojna of the GGPTE who kindly provided me with the material to test the model presented and devoted valuable time to solve my problems. To J. Cullen of Scott Wilson Kirkpatrick & Co.

The computing work of this thesis has been carried out in the Glasgow University Computing Service and I wish to thank the assistance of its consulting staff, in particular those in the Adam Smith Building.

I would like to thank the British Council and the University of Glasgow for providing the necessary financial support for the undertaking of this research. To the Departamento de Planeación Distrital in Bogotá where I was encouraged to carry out this research.

Above all to my wife for her encouragement and valuable comments throughout this study and for the typing of the report.

SUMMARY

The purpose of this study has been to develop a model capable of evaluating existing transport systems in urban areas and of supplying relevant information helpful for planning in the short run. The proposed model is intended as a decision-making tool and is developed on a modular approach in order to deal with the different aspects of the public transport assignment process. The demand for public transport is assumed to be completely determined and given. The analysis is intended for studies of transport by bus although some other modes of public transport may be handled provided that the main assumption concerning the independence of the arrivals of vehicles and passengers at boarding points is taken into account.

Given the public transport demand and information about the corresponding supply, the model assesses the performance of the system in terms of parameters of efficiency and generalised costs to the users. Then, by making changes to the basic variables, a set of alternatives to the existing system may be proposed and evaluated in order to establish the best design or to make minor modifications to the system.

The assignment problem is subdivided into three different stages: i) the determination of individual options to make a trip; ii) the process of loading passengers onto selected options; and iii) the process of loading passengers onto routes and vehicles. Six main linked submodels are used: PATH, ROUTE-2, ASSIGN, TRAVAR, ARRIVE and SIMULA.

PATH is a matrix-notation algorithm which obtains the

shortest paths in networks. This submodel is the basis of ROUTE-2 which deals with minimum paths for public transport networks and to do so it uses total travel time as a measure of travel resistance. Apart from the optimum options to make a trip ROUTE-2 is capable of registering alternative ways to travel where these are feasible. This submodel does not require the coding of dummy links in order to consider the respective level of service of the walking mode.

Having obtained the different alternatives to make a trip and given the number of people wishing to travel, submodel ASSIGN apportions passengers among the options. To do so it is assumed that the number of passengers taking each option is inversely proportional to the option's total travel time. This submodel transforms the door-to-door O-D matrix into a stop-to-stop matrix.

The last three submodels TRAVAR, ARRIVE and SIMULA deal with the third stage of the assignment process. This stage relies heavily on simulation techniques which are very useful in the analysis of complex queueing problems. For the simulation process TRAVAR and ARRIVE play an important role. Variations in travel time between bus stops are assumed to follow a gamma distribution and given the mean and mimimum possible travel time between two linked stops submodel TRAVAR estimates the respective times for random numbers between 0 and 1, creating in this way a table and eventually a file with the set of tables for the whole bus network. Thus, the process of obtaining inter-stop travel times is reduced to a simple table search given a random number.

The arrival of passengers at bus stops is assumed to follow a random pattern which means that there is no direct association between arrival times and expected bus departures or other passengers arrivals. Submodel ARRIVE generates exact arrival times by transforming a random number by means of the inverse cumulative function of the exponential distribution. Following this procedure and with information supplied by previous submodels ARRIVE creates a file of passenger registers containing for each passenger his origin, destination, time of arrival and bus services to make the journey.

Submodel SIMULA controls the simulation of the public transport operations and the gathering of meaningful statistics. Buses are considered as the entities of the system and events are represented by their arrival at bus stos. When a bus arrives at a stop, passengers waiting for that service are placed on board on the basis of first come first served, but always taking into account bus seat availability. Passengers are set down according to their destination.The fact that passengers are considered individually until they board the bus allows the respective waiting and total travel times to be differentiated. The passage of time in the simulation is recorded by means of a number referred to as clock time and this submodel guarantees that events are taken into account in their chronological sequence.

The main data input to the model are composed of: bus stop spacing, inter-node distances, bus frequencies, bus route description, O-D matrix and set of parameters. The model gives detailed information about: minimum distances in the network and path description, options to travel by public transport between any two pair of nodes in the network and relevant information on the walking mode, average bus occupancy, average waiting times, average total travel times, passengers boarding and alighting, passengers not boarded due to seat unavailability and other factors.

The model has been tested using real data gathered in Springburn, a suburb of Glasgow. Taking into consideration some limitations in the information used, the results are encouraging and suggest that further attempts to reach a more extensive validation of the model may be successful.

INTRODUCTION

1. The importance of public transport planning in urban areas.

Transport planners have recognized the potential of transport provision to shape the urban environment by influencing the accessibility within urban areas, and its influence will be proportional to the quality and quantity of service provided. In other words, land use and traffic flow can be regarded as inter-dependent.

On the other hand, the automobile has also influenced the design of the city to a considerable extent and has changed the pattern of urban life. In general, the resources which are considered essential for a satisfactory standard of living are dispersed and people without access to a private car find difficulty in coping with this pattern of urban design, partly because other modes of transport cannot serve as well the diffuse pattern created by the automobile.

The increasing demand for transportation due to a growing population and to a rising standard of living has imposed considerable strain on the existing infrastructure. As the efficient organization of a city requires the adequate provision of access and mobility, many cities have tried to solve their transportation problem by providing additional capacity, especially in highways. But this solution has proved so expensive that society must now consider whether it might be more desirable to halt the increase in private car usage in congested areas and to encourage public transport.

Therefore, the impracticability of providing enough road capacity to cope with unrestrained demand, the need to provide an alternative to the private car, the provision of adequate mobility to those dependent on public transport, the best use of scarce financial resources, and other considerations have encouraged planners to look at public transport with renewed interest.

The primary role of a public urban transport system remains that of giving people mobility, and in order to fulfil this role adjustments to the system are continually necessary to keep pace with the natural growth of the city. Where public transport services have remained static over the years new patterns of land use and changing travel standards necesitate rationalisation of these services. There has been a shift in population away from the centre of the cities to the suburbs and therefore it is essential to adapt the transit system to these changes and to update transport services in order to keep pace with changes in distribution and activity levels.

Transportation demand is characterized by having cyclic fluctuations with similar patterns over different days of the week. As a result there are big differences between off-peak and peak periods which make it necessary to provide additional transport capacity in order to meet peak demands. This imbalance between supply and demand usually causes under-utilization of equipment during non-peak periods.

To meet these challenges public transport operators have responded from a managerial point of view, trying to get a favourable balance between their inputs and outputs; in other words thinking in terms of efficiency. But in most cases this approach has not prevented permanent loss of patronage. This unexpected result has been analysed by transport planners and no unique explanation has been given. It is not completely known what value people place on privacy, door-to-door service, comfort, and the like, but it seems that a rising standard of living plays an important role and that because of this people are switching to private transport.

What may be drawn from this is that bus operators should think not only in terms of efficiency but should also consider the quality of service provided. If the previous analysis is correct it would be desirable to improve the level of service in order to keep pace with rising standards of living, adjusting bus services to the requirements of the whole community, while taking account of operating costs.

2. Framework of the study.

The transport planning process may be regarded as consisting of three different stages: inventory, forecast and evaluation. The inventory comprises the development of the data base for the whole process: land use data, travel pattern data and existing transportation facilities information. The next stage is that of forecasting the future urban land-use patterns that the transportation system is to be designed to serve and of trying to describe the travel patterns and the way people will use the transport facilities. The evaluation stage attempts to assess whether the transportation proposals put forward satisfy the estimated demand from the social, economic and operational point of view. The forecasting of travel patterns and use of transport facilities constitute the so called Transport Model which can be seen as composed of four submodels: trip generation, trip distribution, modal split and assignment.

Trip generation is concerned with the number of trips which will be generated by and attracted to each zone of the study area. In other words, it deals with the prediction of future levels of travel. Trip distribution seeks to establish the links for the attracted-generated trips calculated in the trip generation process. Modal split is the process of allocation of trips among the available modes of transport. Assignment is concerned with the allocation of trips to a given transport system and to its different elements. The first three sub-models are the major elements in forecasting travel demand while the assignment process allocates this demand to the existing transport facilities.

Land use, distribution of population, socio-economic characteristics of the population and patterns of activity determine the transportation demand for each mode of transport. Furthermore, transportation demand may itself be influenced by the installation of transport facilities and the provision of service that may change the locational patterns of the economic activity. In this sense it is possible to talk about effective demand and latent demand, the latter being the potential number of people that would use transit if different services, new or improved, were provided.

In other words the interaction between transport supply and demand is well defined by the level of service variables. The volume, composition and time dependency of the demand depend upon the level of service at which transportation is supplied. At the same time the level of service depends upon the volume, composition and time variation of the demand. Therefore, the public transport demand is affected by the service provided which in turns depends upon the community support. Consequently, this is a dynamic process in which variables constantly interact with each other.

The elements of the supply of public transport could be regarded as classified in two main groups: the infrastructure and the transport service elements. The infrastructure consists mainly of the street network, termini, depots, etc. The service elements are the equipment itself and the transport network characterised by routes and frequencies of service. The human factor also plays an important role in public transport systems as transport operation is usually regarded as a labour intensive industry.

Improvement or deterioration of the transport system may influence people's behaviour: as an immediate reaction to a change people can either do nothing, or switch to another route of the system, or switch to another mode of transport or be attracted from other mode to the public transport system.

Thus, the evaluation of a public transport system could be accomplished by analysing carefully the different impacts it may produce. However, a conflict arises in the sense that what is the optimum for some people may not be for another: the transport operator usually seeks to maximise his profits, the user thinks in terms of quality of service and minimum costs, and society as a whole wants a system which causes minimum congestion and damage to the environment. As variables involved in the process interact with each other when attempts are made to improve one particular aspect, another is often made worse.

Therefore, as in any planning or evaluation process it is essential first of all to define the goals of the transportation planning process, taking into account that certain improvement in one particular aspect could possibly deteriorate another one. Some goals could be:

- Improving urban life by providing better quality of transportation;

- Increasing the economic and technical efficiency of the system in order to reduce operating costs and transportation time;

- Improving service availability specially for those who do not have other alternatives.

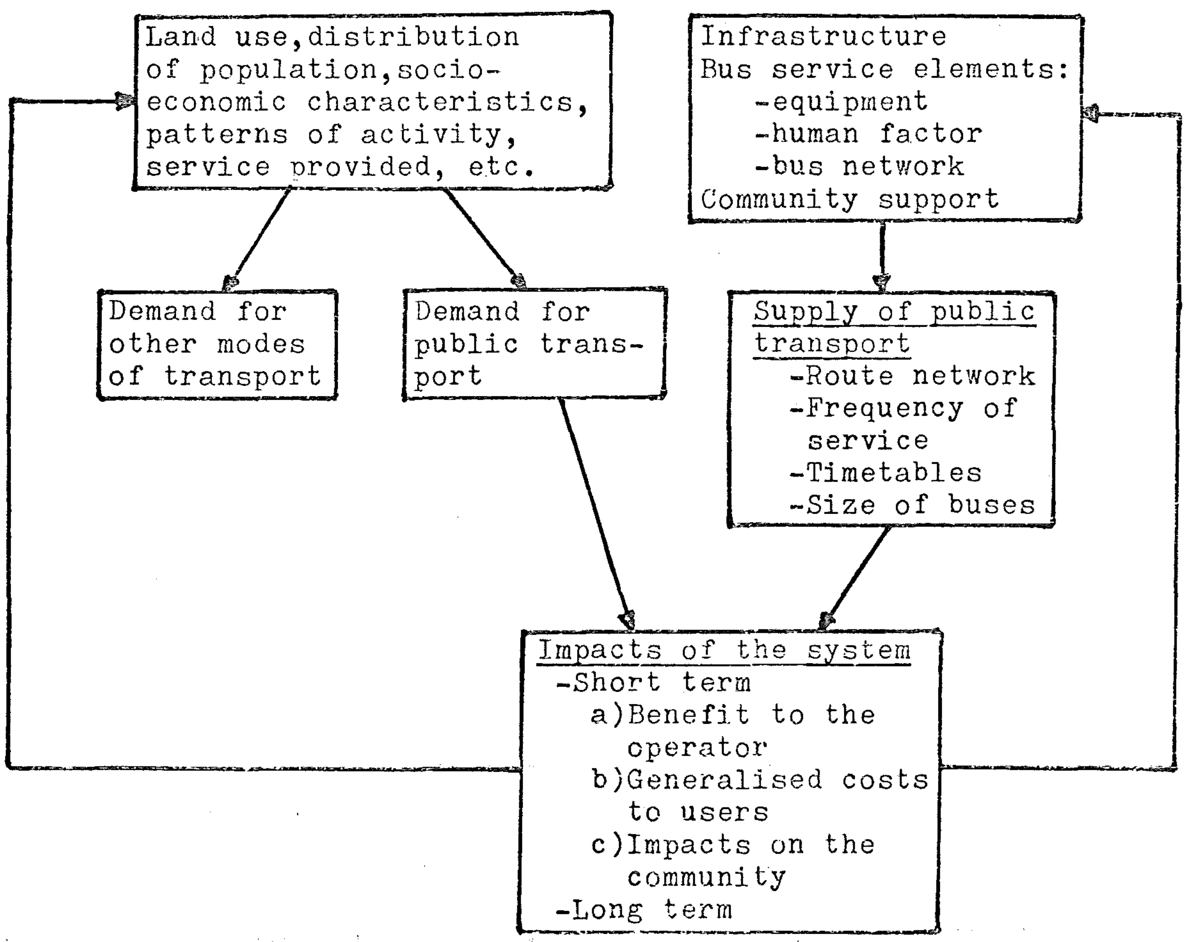
These goals may be translated into costs, which carefully assessed can serve in the evaluation process of a public transport system. The main costs are: operating costs, passengers generalised costs, time and operating costs of all other vehicles on the street network and relevant environmental costs. The public transport system which minimises the total of these costs may be regarded as the optimum.

The design of alternative transportation plans which accord with community aims and objectives meet the estimated demand for transportation should be evaluated in both economic and social terms to derive the optimum solution. But a public transport system is so complex and there are so many variables interacting that it is considered to be desirable to have a model in order to evaluate changes and to help planners in the task of updating services to the real requirements of the community, within the context of providing some benefits to the operator and good degree of mobility to the users.

The purpose of this study has been to develop a model capable of evaluating existing transport systems and of supplying relevant information helpful to planners in seeking a better match between supply of and demand for public transport in such a way that with little or no capital investment the system could be improved and the existing resources used in the most efficient way.

The public transport planning process may be represented by a diagram (see Figure 1). It is clear that the effects or impacts produced by the system will certainly have repercussions on the subsequent levels of supply and demand as this is a dynamic process. However,as the proposed model is to be used for planning in the short run, it was assumed here that these impacts will not affect the demand. A further assumption was that the demand for public transport is completely determined and given, either by a forecasting process similar to the one explained above or by information derived from a survey. Secondly, the analysis was restricted to transport by bus in urban areas, though some of the concepts could easily be applicable to other modes of public transport.

FIGURE 1. Public transport planning process.



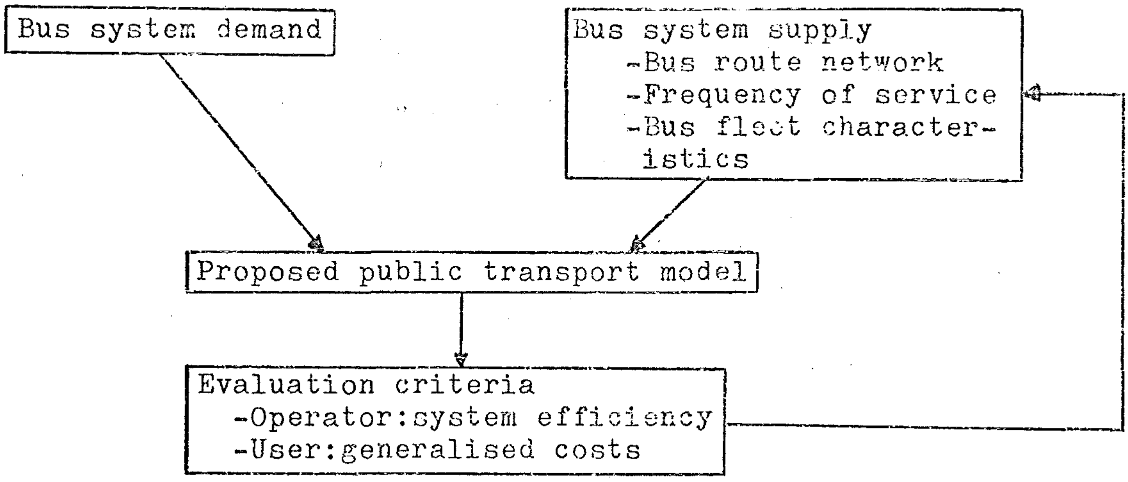
In the choice of criteria to measure the impacts of a bus system it is advisable to choose standards that consider the process as a whole and are as far as possible quantifiable. Therefore, environmental impacts such as noise and air pollution were beyond the scope of this model. Traffic aspects, on the other hand, were only considered indirectly as it was assumed that by using the system in a more rational and efficient way the burden of the street network could be diminished.

From the operator's point of view the system is evaluated by using appropriate measures of efficiency; from the user's stand point perhaps the most important factor is how long it will take him to get his final destination. Then, the model concentrates in the different components of a trip such as : walking, waiting and riding times, and fares expressed in generalised costs. The approach proposed is presented in Figure 2.

In order to evaluate the quality of a bus system the proposed model should be capable of representing the interaction between the different variables and of measuring a set of standards able to characterised the effectiveness of the whole system. At this point of the analysis two different approaches might have been considered:

1) To seek the optimisation of the bus system by means of a linear programming model and then to measure how far the existing system is from the optimum; i.e. given the public transportation demand, the question is to design an optimum route network with a frequency of service in such a way that generalised costs to passengers are minimised. Due to the complexity of the variables involved in the process this approach is considered to be rather difficult to carry out.

FIGURE 2. Approach proposed.



2) To develop a model capable of evaluating the performance of a bus system and by means of some changes in the basic variables evaluate a set of alternatives to the existing system in order to implement the best design or to make minor modifications to the system in operation. In other words, given the corresponding public transportation demand and information about the supply the model assesses the performance of the system in terms of parameters of efficiency and in terms of generalised costs to the users. Thereafter, by analysing reasonable alternatives, the system could be either improved or kept unchanged. By exploring a range of operating strategies the system could be improved by either reducing the amount of resources keeping constant the quality of service or improving the level of service given a fixed amount of resources or by a combination of both. This second approach was chosen in view of the complex analytical relationships involved in an optimisation process.

It is clear that a theoretical model capable of doing the required task could be large and complex; but the problem could be broken down into two stages: i) analysis of the patterns of travel and the way people may use the transport system in order to make a trip; ii) process to allocate trips to the existing transport facilities. It is worth noting that the previous stages constitute the basic elements of a public transport assignment model and for the purpose of this research several linked submodels have been used to accomplish the required functions.

The first stage relies heavily on network analysis techniques as it was assumed that people choose their routes to make a particular trip so as to minimise their total travel time. The second stage uses simulation techniques that have been demonstrated to be adequate for the investigation of systems involving queues and complex interactions between them. Furthermore, simulation is a practical approach when a large volume of computations is required and when analytical tools are inappropriate. This technique is not to be considered as an optimisation process by itself but as a tool to improve systems and to enable decisions to be taken.

Thus, the proposed model is intended as a decision-making tool in the sense that it may provide basic information relevant for transport operators and policy-makers involved in the public transport planning process. The data base comprises on the one hand the transit demand, and on the other hand the route network, frequencies of service, and bus fleet characteristics. Since the demand is assumed to be given, the process of seeking better alternatives to the existing transport system depends upon changes in the supply side.

The layout of the thesis is as follows. The first chapter provides a revision of the main literature related to the topic;

it also presents the main concepts of simulation techniques that play an important role in the modelling process of this study. The second chapter presents in detail the different steps of the proposed model. The third chapter is devoted to the application and testing of the model using basic data collected in Springburn, a suburb close to the city centre of Glasgow. Finally, a summary and conclusions are presented, emphasising the possible implications of such a study in the planning process.

CHAPTER I

REVIEW OF CONCEPTS IN MODELLING AND ANALYSIS OF PUBLIC TRANSPORT OPERATIONS

The main purpose of this chapter is to provide theoretical insight into the different aspects of model building and the factors required for a better understanding of public transport operations in urban areas. The content of the chapter is divided into three parts for convenience and simplicity. The first part is a basic examination of the most important principles of simulation modelling, making particular emphasis on the main concepts involved in stochastic simulation, its relevance and utility for the present study. The second part covers the impacts of a public transport system on the different sectors of society and some important factors affecting service reliability. The last part presents some recently developed models which deal with a variety of facets of public transport planning.

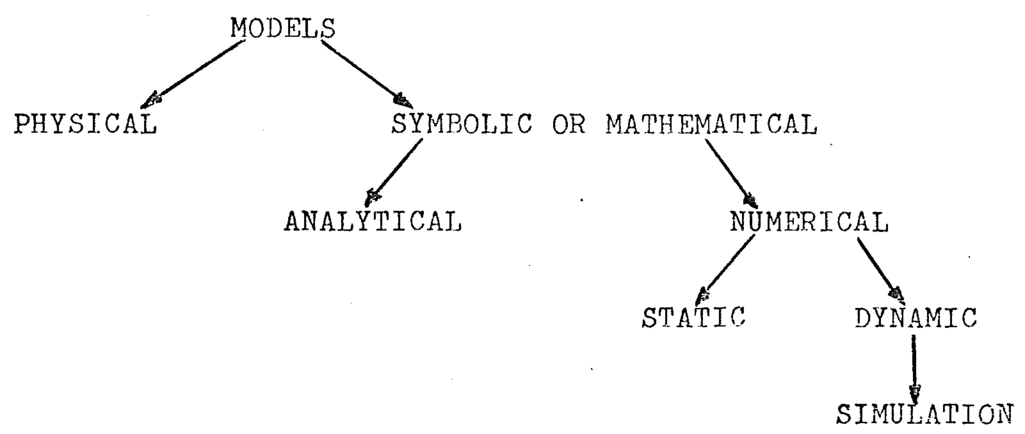
A review of other work in this field was considered essential for the design of the proposed model; one relevant factor that remains to be presented in the next chapter is that of minimum paths in networks and its extension to public transport analysis.

1. Some important concepts in simulation models.

A system may be defined as a set of interdependent or interacting elements, acting as a unit to fulfil a specific objective. In this context, Systems Analysis is a term used to describe an approach to the study of large and usually complex systems; this approach attempts to study the total unit performance rather than to concentrate on the parts.

It must be stressed, however, that any representation or model of a system will differ inevitably from the system itself. There are two different kinds of models: physical and symbolic. A symbolic model represents a system by using mathematical or logical terms, and according to the techniques used to solve the mathematical models it may be classified either as analytical or as numerical. This classification is illustrated in Figure 3.

FIGURE 3. Classification of models



Perhaps the most satisfying form of model solution is an analytical solution. In order to apply this method it is necessary to fulfil two general conditions:

i) that the behaviour of the elements involved and the relationships between them can be fully described in mathematical terms; and

ii) that a mathematical theory is available to solve the model.

Such a solution can only be obtained when the system to be modelled has a very tight logical structure, and the internal relationships are uncomplicated. When feasible, this approach can be a cheap and quick way of reaching a solution. However, in many cases the analyst is not able to formulate the model with the mathematical precision required or to solve the model. In some complex queueing models, for instance, a complete mathematical formulation of the problem does not exist.

When the previous conditions are not fulfilled the analyst should then consider the use of numerical techniques. The range of queueing models whose behaviour can be examined by analytical techniques is very limited. Where analytical techniques cannot be applied however, approximate solutions with sufficient accuracy can often be obtained by numerical approaches. A numerical solution manipulates numbers obtained by substituting them for the variables and parameters of the model. Monte Carlo and simulation techniques constitute two numerical approaches which are in common use.

Monte Carlo techniques use random numbers to obtain solutions. Simulation refers to certain mathematical and logical models that describe the behaviour of a system over a period of time. In this way, simulation is a dynamic approach in contrast to those static models in which time does not play any role at all.

A simulation where any of the inputs or outputs are represented by a distribution of values rather than by single numbers is said to be stochastic. Since these models use random numbers it could be said that a stochastic simulation is synonymous of Monte Carlo techniques; but it is important to clarify that the latter uses random numbers to solve deterministic problems,eg. to find the value of an integral.

In the formulation of a stochastic simulation model it is necessary to specify very clearly the components or entities of the system under study, the variables and parameters and the functional relationships (Naylor et al, 1968). These are illustrated in Figure 4.

FIGURE 4. Elements of a simulation model

Elements of a simulation model
- Entities
- Variables
  - Endogenous
    - Output
    - Intermediate
  - Exogenous : Input
- Parameters
- Relationships
  - Identities
  - Operating characteristics

A component or entity is an object of interest in a system. Variables are used to represent corresponding features of the system and can be classified as exogenous and endogenous variables. Exogenous or independent variables denote characteristics external to the system being modelled; in other words they are input variables that may act upon the system while the system does not affect them. According to the control that the analyst has on these variables, they are classified as controllable and non-controllable variables.

Endogenous variables denote characteristics internal to the system and can be classified as intermediate variables that describe the state of the system and output variables. Therefore, the variables of a model can be classified as input, intermediate and output variables.

Parameters are quantities that influence the endogenous variables and are constants that can be derived on the basis of statistical inference. (statistical inference deduces the characteristics of the totality, based on the observations of less than the totality).

Relationships describe the connection between variables and components of a system.They can be classified as identities or definitions, and operating characteristics that relate input and intermediate variables to output variables normally by means of a mathematical function.

In a stochastic simulation model at least one of the operating characteristics is given by a probability function; in other words, a system may be regarded either as deterministic or stochastic depending upon the causal relationship between input and output.

In looking at a simulation model it would be useful to have a different view of the elements of the model. The system could be viewed as a set of entities (components) which have certain characteristics expressed by means of numerical or logical values called attributes (variables and parameters). On the other hand, the entities interact with specific activities or changes in the system according to some prescribed conditions (relationships) that determine the sequence of interactions or events.

The description of all the entities, attributes and activities of the system as they exist at one point in time is called the state of the system. The dynamic behaviour of a system implies the concept of state. Thus, simulation is a technique that can register and trace over a period of time a life history which consists of a sequence of changes in the system state.

If a bus system, for instance, is assumed to be composed of passengers and buses travelling along the different routes, at any instant t the state of the system can be described by the instantaneous position and conditions of each element. Similarly, at any time t + dt the system can also be described, and so on, which means that it is possible to obtain a set of instantaneous pictures of the different states of a system over a period of time, thus establishing the history of the system. A simulation model can reproduce the historical state of a system and if appropriately defined this historical state can be taken as that of the real system.

One of the key stages in this process is to ensure that the model accurately represents the situation to which it is going to be applied so that the state descriptions it makes will give a reliable indication of what will actually happen in any given set of circumstances.

A change in one aspect of a system may produce changes or even create the need for changes in other parts of the system, and even if its individual elements are optimised a system may be non-optimal due to interactions between the parts, so it must be emphasised strongly that it is important to study a system as a whole. In order to do this it is necessary to observe it in the real world, to formulate hypotheses about the behaviour of the system and to reduce these to a level of abstraction that allows the analyst to formulate the model in a convenient and suitable way.

Simulation has been widely used to solve a variety of problems concerning existing and proposed systems and is a valuable tool in scientific research. However, as with any other technique it has advantages and disadvantages; therefore, the decision to use it must depend on the judgement of the analyst. The advantages include the following:

1) It permits experimentation with:

- consideration of many factors;
- manipulation of several individual parameters;
- ability to consider alternative policies; and
- no disturbance of the actual system.

2) It enables an analysis of a system to be made in a chronological sequence;

3) It provides operational insight;

4) It allows the use of observed distributions rather than idealised approximations; and

5) Simulation models can be more easily understood by non-technicians than analytical models which may involve sophisticated mathematical reasonings.

Among the disadvantages:

1) They may be very expensive especially because of the time required to develop them;

2) They may require scarce and expensive resources: data and computer processing for example; and

3) They may require extensive field studies.

Stochastic simulation involves basically four different concepts: probability functions, random number generation, random variables and time keeping techniques.

A. Probability functions. In statistical terms, a variable is a characteristic that can be measured; an individual measurement of a variable is called a variate. If a variable can assume specific values it is called a discrete variable, while a continuous variable can assume any values whatsoever between certain limits.

Consider a discrete variable which can take I different values, X_i (i = 1,2,.....I). The set of possible outcomes is called the sample space and is denoted by S. If the event is repeated n times and the numbers of occurrences of each value X_i are defined as the absolute frequencies n_i then $\sum n_i$ over S is equal to n. A relative frequency is obtained by dividing each absolute frequency by n. For large numbers of n, the relative frequency will approximately be equal to $p(X_i)$ which is the probability of occurrence of the value of X_i. The set of values $p(X_i)$ is said to be a probability mass function.

A theoretical probability mass function associates a probability with each possible outcome. In the case of an empirical distribution (based upon actual observations), the probability mass function associates relative frequencies with outcomes.

The probability that a variable takes on a value less than or equal to X_i is given by the cumulative distribution function which in mathematical terms is simply called the distribution function. If the variable being observed is continuous, an infinite number of possible values can be taken by the variable; therefore, the probability that the variable takes on the value of X is to be considered zero. Thus, to describe the variable, a probability density function f(X) is defined. The probability that the variable takes on a value in the range a to b is given by:

$$p(a \leqslant X \leqslant b) = \int_a^b f(X)dX \qquad f(X) \geqslant 0$$

For a continuous variable, the cumulative distribution function F(X) which gives the probability that the random variable is less than or equal to X is given by:

$$F(X) = \int_{-\infty}^{X} f(t)dt$$

It follows by definition of the probability mass function and the probability density function that:

$$\sum_{i=1}^{I} p(X_i) = 1 \qquad \text{and} \qquad \int_{-\infty}^{\infty} f(X)dX = 1$$

Probability functions are summarized by certain key parameters that can provide insight into some of the most important characteristics of the distributions. The <u>nth moment</u> of a distribution is given by:

$$\sum_{i=1}^{I} X_i^n \, p(X_i) \qquad \text{for dicrete variables, and}$$

$$\int_{-\infty}^{\infty} X^n \, f(X)dX \qquad \text{for continuous variables}$$

The first moment, n=1, is called the mean $\bar{X}$ which indicates the central tendency of the distribution. It is said that if a single number is to be chosen to represent the variable without any randomness, the mean is the best one to use.

It is possible to take moments about some constant value such as the mean $\bar{X}$:

$$\sum_{i=1}^{I} (X_i - \bar{X})^n \, p(X_i) \qquad \text{for discrete variables}$$

$$\int_{-\infty}^{\infty} (X - \bar{X})^n \, f(X)dX \qquad \text{for continuous variables}$$

Moments about the mean are called central moments; the second central moment is called the variance, and its positive square root the standard deviation. The standard deviation which has the same dimensionality as the observations, measures the degree to which data are dispersed from the mean. Another useful measure of dispersion is the coefficient of variation which is defined as the ratio of the standard deviation to the mean. The third moment about the mean characterizes the skewness of the distribution, while the fourth moment is the measure of Kurtosis (flatness or peakness) of the curve.

Some of these terms are illustrated by the figures in Table 1. This records the number of 1-minute intervals in which X_i passengers arrive at a bus stop, and gives the corresponding probability and cumulative distributions.

TABLE 1. Probability and cumulative distribution of passenger arrivals per one-minute interval.

Arrivals	Number of 1-min interval	Probability distribution	Cumulative distribution
X_i	n_i	$p(X_i)$	$P(X_i)$
0	30	0.32	0.32
1	31	0.33	0.65
2	18	0.19	0.84
3	9	0.10	0.94
4	5	0.05	0.99
5	1	0.01	1.00
	94	1.00	

Source: Calculations made by the author.

Mean $\bar{X}$ = 1.26 arrivals

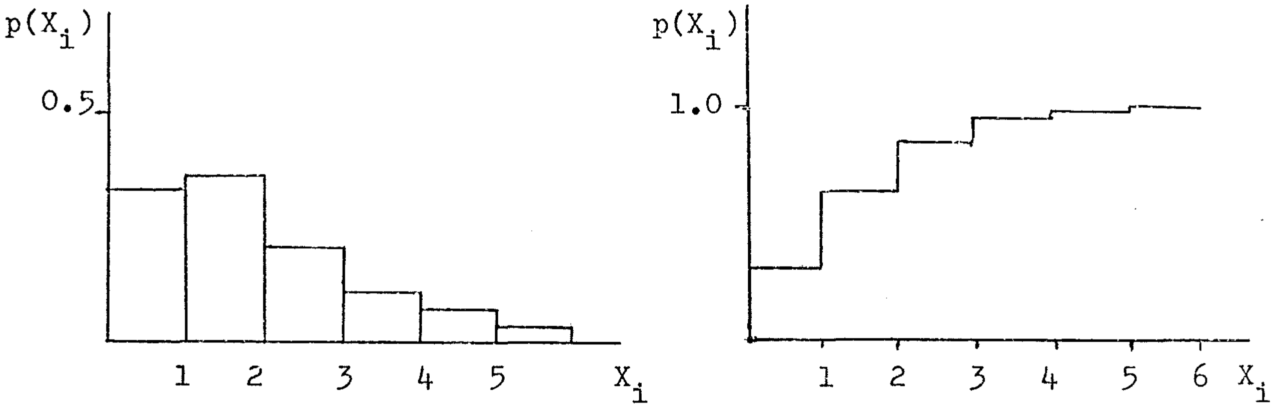
Variance σ^2 = 1.45

Standard deviation σ = 1.21

Coefficient of variation $\sigma/\bar{X} = 0.96$

The probability mass function displayed graphically (Figure 5) appears as a bar histogram, while the respective cumulative distribution appears as a step function type of graph.

FIGURE 5. Relative and cumulative distributions of data in Table 1.



The following theoretical probability functions are of particular relevance in this study: Poisson, exponential, normal and gamma distributions, and a brief description of their characteristics follows.

Poisson distribution. The probability mass function of the Poisson distribution is given by:

$$p(n) = \frac{\bar{n}^{\,n}\; e^{-\bar{n}}}{n\,!}$$

where p(n) is the probability that the discrete variable takes on integer value n.

$\bar{n}$ is the mean of the variable.

n is an integer value of the variable.

A probability function may incorporate more than one variable. An example of a multivariate function is the following version of the Poisson distribution:

$$p(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

where t is the time interval over which n occurs

λ is the mean rate of the occurrence of n

$\bar{n} = t$

For the Poisson distribution mean $= \lambda t$ = variance.

Exponential distribution. This distribution can be derived from the Poisson distribution when n=0, which means that no events occur during a time interval of length t.

$$p(0) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t} = p(h > t) \qquad t \geqslant 0$$

where $p(h > t)$ is the probability that an event occurs with a gap greater than t. Then,

$$p(h \leqslant t) = 1 - e^{-\lambda t} = F(t) \qquad t \geqslant 0$$

which is the cumulative distribution of the exponential function. Thus,

$$f(t) = \frac{dF(t)}{dt} = \lambda e^{-\lambda t} \qquad t \geqslant 0$$

If the probability that an event occurs in a small time interval is very small, and if the occurrence of this event is independent of the other, then the time interval between occurrences can be represented by an exponential distribution.

Normal distribution. The probability density function of the normal distribution is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

where μ = mean and σ = standard deviation

This distribution is symmetric about its mean. If the transformation $Z = (x-\mu)/\sigma$ is made, the normal distribution is transformed to a form in which $\mu=0$ and $\sigma=1$:

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-Z^2/2}$$

which is the density function of the standard normal distribution. This has the advantage of removing the complication of expressing the variate in different units in different contexts.

<u>Gamma distribution</u>. The gamma distribution is described by the following density function:

$$f(t) = \frac{\alpha^K t^{(K-1)} e^{-\alpha t}}{(K-1)!}$$

whre $\alpha > 0$ and $K > 0$ for positive values of t. The exponential distribution can be obtained from the gamma when K=1. For a positive integer value of K, the gamma distribution is identical to the so called Erlang distribution found representative of certain types of telephone traffic. The mean and standard deviation of the gamma distribution are given by:

$$\bar{t} = K/\alpha$$

$$\sigma = \bar{t}/K^{\frac{1}{2}}\text{ , then coefficient of variation } = 1/K^{\frac{1}{2}}$$

When the analyst has obtained an empirical distribution, selection of the possible theoretical probability distributions from which it may have derived is a matter of judgement and experience. If the sample data do not differ significantly from the theoretical or expected frequency, it is usually convenient to use the theoretical distribution to represent the underlying population.

The analyst normally determines the main distribution parameters such as the mean and the variance. Thereafter, he selects one or more theoretical distributions that he believes might fit the sample data. Then, he can proceed with statistical testing.

Such a test may be defined as a mathematical procedure applied to empirical sample data for deciding on the grounds of probability whether or not a hypothesis is tenable. The choice of test, on the other hand, depends very much on what is known about the population from which the sample was drawn. Two goodness-of-fit test are widely applied: the Chi-square test and the Kolmogorov-Smirnov test.

The Chi-square test can be applied to test a hypothesis regarding the particular distribution from which a given set of empirical data were sampled. The discrepancy existing between an observed and an expected frequency is given by the statistic χ^2 (Chi-square):

$$\chi^2 = \sum_{j=1}^{c} \frac{(O_j - E_j)^2}{E_j}$$

where O_j is the observed frequency of values falling in the Jth class;

E_j is the expected frequency of values falling in the Jth class; and

C is the number of classes.

There are some practical restrictions on the values for E_j and C.

It is clear that the larger the value of χ^2, the greater is the discrepancy between the two distributions; on the other hand, if $\chi^2 = 0$ the observed and theoretical distributions agree exactly.

There are practical and theoretical reasons for believing that the arrival of people at a bus stop follows a Poisson distribution. Nevertheless, in the previous example (Table 1) if the empirical data had come from a Poisson distribution the mean and the variance would be equal, which in fact is not so. Despite this discrepancy, on the grounds of experience it would be worth testing statistically the hypothesis (null hypothesis) that there is no significant difference between the observed frequency distribution and that of a Poisson distribution with the same parameters.

For the example in question, $\bar{n} = \lambda t = 1.26$, and the variance $\sigma^2 = 1.45$, so the values of the expected Poisson distribution could be obtained using as mean an average value of $(1.26 + 1.45)/2 = 1.355$ (Shannon, 1975). In this way:

$$p(n) = \frac{1.355^n \, e^{-1.355}}{n!} = \frac{0.2579 \times 1.355^n}{n!}$$

and the probabilities are multiplied by 94 to get the frequencies.

TABLE 2. χ^2 calculations for Table 1.

n	O_j	E_j	$(O_j - E_j)^2/E_j$
0	30	24	1.50
1	31	33	0.12
2	18	23	1.09
3	9 } 15*	10 } 14*	0.07
4	5	3	
5	1	1	
	94	94	$\chi^2 = 2.78$

Source: Calculations made by the author.

* Expected frequencies for each class should equal 5 or more. If not, classes can be grouped.

Degrees of freedom $v = c-1-m$; where c is the number of classes and m is the number of sample data population parameters necessary to calculate the expected frequencies. Here,

$$v = 4-1-1 = 2$$

From the tabulated values of χ^2, using a confidence level of 0.90 and two degrees of freedom, $\chi^2_T = 4.61 > 2.78$ which means that the null hypothesis is not rejected.

The Kolmogorov-Smirnov test can also be used to test the degree of agreement between the sample data and a theoretical distribution. The test compares the respective cumulative distributions in order to obtain the largest absolute deviation, which is to be compared to the critical values to determine if this deviation could occur owing to random variation. For the example under study Table 3 shows the respective calculations.

The largest absolute deviation is 0.06 for zero arrivals. From the tabulated critical values, for 94 intervals, $\alpha = 0.10$, the critical value is given by $1.22/\sqrt{94} = 0.126 > 0.060$. Therefore, there is no reason to reject the null hypothesis.

TABLE 3. Kolmogorov-Smirnov calculations.

n	$p_O(n)$	$p_T(n)$	$P_O(n)$	$P_T(n)$	(D)
0	0.32	0.26	0.32	0.26	0.06
1	0.33	0.35	0.65	0.61	0.04
2	0.19	0.24	0.84	0.85	0.01
3	0.10	0.11	0.94	0.96	0.02
4	0.05	0.03	0.99	0.99	0.00
5	0.01	0.01	1.00	1.00	0.00

Source: Calculations made by the author.

B. Random number generation. Random numbers form a set of numbers produced entirely by chance or assumed to be free from statistical bias for a specific purpose. In computer applications it is possible either to read a lengthy table of random numbers into a computer memory or to use a random number generator. A generator may be either a hardware unit designed to produce random numbers in specified quantities or an algorithm capable of producing a sequence of numbers sufficiently random for any desired degree of statistical accuracy. A uniform random number, on the other hand, ranges between zero and unity and it is assumed to be randomly selected from the uniform probability distribution function. A very commonly used method for generating uniform random numbers is called the congruential method, based on the following relationship:

$$U_{i+1} = aU_i + c \qquad \text{(modulus m)}$$

The initial value of U_i is called the 'seed', U_o; and where a, c, m and U_o are all non-negative integers. The (i+1)th number is derived from the ith number by multiplying by a, adding c and then taking the residue upon dividing by m. The method as described above is called the mixed congruential method. If c = 0 it is called the multiplicative congruential method while if a = 1 it is called the additive method.

According to the definition given above where the input U_i becomes equal as at some previous stages, the whole sequence will start repeating itself. Therefore, in the multiplicative method the choice of a, U_o and m is influenced by the need to obtain a maximum period with a minimum degree of correlation. In order to satisfy these conditions a should be prime to m, in other words it must be an odd number, then a is given by:

$$a = 8t \pm 3 \quad \text{where } t \text{ is any positive integer}$$

For a binary computer with b digits in a word, Naylor et al (1968) suggests that if a is close to $2^{b/2}$ it will minimise the first order correlation between random numbers. Shannon (1975) accepts that this is true but points out that such a value will produce strong correlation between triplets. So, he recommends that a be chosen so as to be five or more digits and not contain long strings of zeros and ones; at the same time, U_o should be positive and odd.

A commonly used modulus is $m = 2^b$ and it is possible to prove that the maximum attainable period is given by $h = 2^{b-2}$ (Naylor et al, 1968), which means that only a finite number of distinct integers can be generated, after which the sequence repeats itself.

If N is defined as an integer of 32 bits, and $A = 5^{13}$ = 1220703125 and $M = 2^{31}$ = 2147483648, the following FORTRAN subroutine generates random numbers according to the multiplicative congruential method:

```
SUBROUTINE RANDOM (N,RN)
N = N * A
RN = ABS(N) * M ** (-1)
RETURN
END
```

The multiplicative method has been found to behave statistically well in the sense that it generates uncorrelated and uniformly distributed numbers. However, it is clear that a reproducible sequence of random numbers which has been generated in a deterministic way is not to be taken as truly random. Nonetheless, on the grounds of statistical tests of randomness, a sequence of generated numbers can be proved acceptable and are referred to as pseudo-random numbers.

The multiplicative method generates pseudo-random numbers requiring a minimum amount of computer memory, which has the advantage of being a reproducible and fast method. One check for the randomness of N uniform random numbers U_i is to compute the mean and partial variance in such a way that:

$$\frac{1}{N}\sum_{i=1}^{N} U_i \simeq 1/2$$

and

$$\frac{1}{N}\sum_{i=1}^{N} U_i^2 \simeq 1/3$$

As mentioned previously, the Chi-square and the Kolmogorov-Smirnov tests are also widely used to compare the distribution of a set of numbers generated against a uniform distribution. For instance, using the uniform random number generator mentioned above, 10,000 numbers were generated in the range 0 to 99. It is expected that if the generator is truly random each digit shows up about 100 times. The observed frequencies are given in Table 4.

TABLE 4. Observed frequencies for generator under test.

108	90	118	88	112	86	105	99	119	116	92	90
86	116	92	97	100	111	96	88	108	100	117	99
105	104	122	90	105	112	92	88	89	94	91	107
92	85	91	118	121	121	94	105	98	110	99	92
107	97	105	96	90	109	105	100	95	96	97	96
112	103	113	103	102	87	107	102	102	107	101	83
86	115	101	105	94	97	112	83	96	110	89	96
88	119	89	102	101	92	81	90	72	101	102	121
87	98	100	100								

Source: Calculations made by the author.

$\chi^2 = 110.42$. Using a significance level of 0.90, for v = 100-1 = 99 degrees of freedom, from the table $\chi^2_T = 117.41 > 110.42$. Alternatively:

$$\frac{1}{10,000} \sum_{i=1}^{10.000} U_i = 0.4952 \simeq 0.5$$

and

$$\frac{1}{10,000} \sum_{i=1}^{10.000} U_i^2 = 0.3283 \simeq 1/3$$

Although the previous results are quite satisfactory they do not guarantee the complete randomness of the generator; however, the discussion of more complicated and diverse tests is beyond the scope of this research. In any case, it is worth noting that no set of tests can guarantee the complete suitability of a sequence generated (Shannon, 1975).

C. Random variates: sample techniques. When a source of random numbers is available, it is then necessary to have a procedure to translate these numbers into random variates. This process of sampling from a probability distribution is summarized as follows:

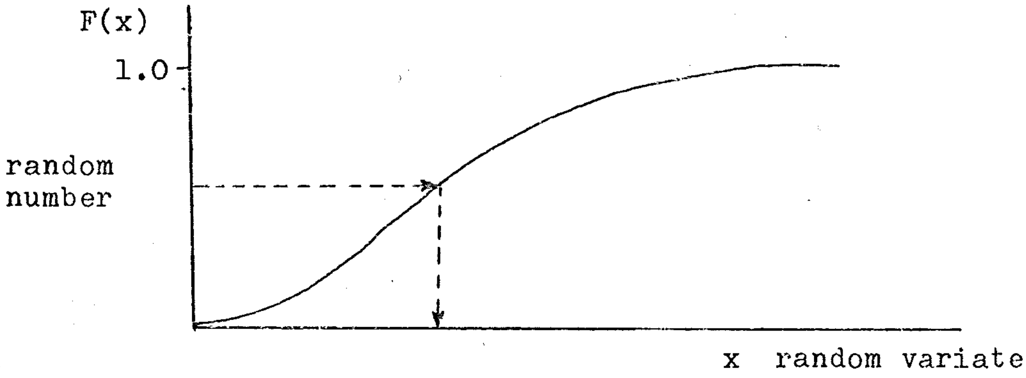
a) Obtain the cumulative distribution function and plot it with the values of the variate on the X axis and the probabilities on the Y axis;

b) Generate a random number between 0 and 1;

c) Obtain the sample value by projecting horizontally on the Y axis the random number generated and where this projection intersects the cumulative curve, projects down in order to get the value of the random variate (see Figure 6). This method can also be used for dicrete variables.

In the case of a discrete probability distribution or when the function of the continuous distribution is not known

FIGURE 6. Generation of random variates from a continuous distribution



(only pair of values are available) it is necessary to design a routine to search in a table in order to accomplish the process explained above. A table search may be made in one of the following ways:

a) If the search is made sequentially, the best is to organize the table selecting the intervals in decreasing order of probability. For the example developed previously using the Poisson distribution, Table 5 shows how to organize the information.

TABLE 5. Sequential search for case of Table 1 using a Poisson distribution.

X	Y
1	0.35
0	0.61
2	0.85
3	0.96
4	0.99
5	1.00

Source: Calculations made by the author.

Given a random number U, the search is made from top to bottom until the following condition is satisfied $Y_i > U \geq Y_{i-1}$. Having organized the table in decreasing order of probability, 35% of the searches will reach only the first entry while with the original organization the corresponding figure would have been 26%.

b) The binary search method relies upon the items in the table being in ascending order of sequence by key value, in this case the cumulative distribution. The key is compared with a midway value in the table in such a way that one half of the table is rejected according to whether the key is greater or lesser than the value selected from the table. Then, the remainder of the table is taken and the process is repeated until the desired value or interval is found. The method is illustrated by means of the example shown in Table 6.

TABLE 6. Binary search for case of Table 1.

I	X(I)	Y(I)
1	0	0.26
2	1	0.61
3	2	0.85
4	3	0.96
5	4	0.99
6	5	1.00

Source: Calculations made by the author.

For instance the generated random number RAND = 0.95 and the number of classes N = 6. Initial values N1 =0, N2 = N = 6.

1) I = (N1+n2)/2 = (0+6)/2 = 3

Y(I) = Y(3) = 0.85

RAND $>$ 0.85 $\rightarrow$ N1 = I = 3

2) $I = (N1+N2)/2 = (3+6)/2 = 4$

$Y(I) = Y(4) = 0.96$

$RAND < 0.96 \longrightarrow N2 = I = 4$

But N2 - N1 = 4 - 3 = 1, therefore the answer is N2 = 4 which means that the value of the random variate is X(4) = 3.

In the case of a continuous distribution, besides the table search it is necessary to make an interpolation. If the random number generated RAND falls between Y(N1) and Y(N2) then the random variate can be approximate by the following linear interpolation:

$$X = X(N1) + \left[X(N2)-X(N1)\right] \frac{RAND-Y(N1)}{Y(N2)-Y(N1)}$$

The following subroutine called SEARCH was written in FORTRAN in order to perform a binary search.

```
      SUBROUTINE SEARCH (RAND, N, Y, N2)
      DIMENSION Y(N)
      N1 = 0
      N2 = N
   3  I = (N1+N2)/2
      IF (RAND.GE.Y(I)) GO TO 1
      N2 = I
      IF ((N1+1).EQ.N2) RETURN
      GO TO 3
   1  N1 = I
      IF ((N1+1).NE.N2) GO TO 3
      RETURN
      END
```

In this subroutine RAND is the random number generated previously, N is the total number of classes or intervals, Y is the table with dimension N, and the answer is given in N2 in such a way that for discrete variables the random variate is X(N2), and for continuous variables the random variate is given by the interpolation mentioned above. This method is recommended for lengthy tables.

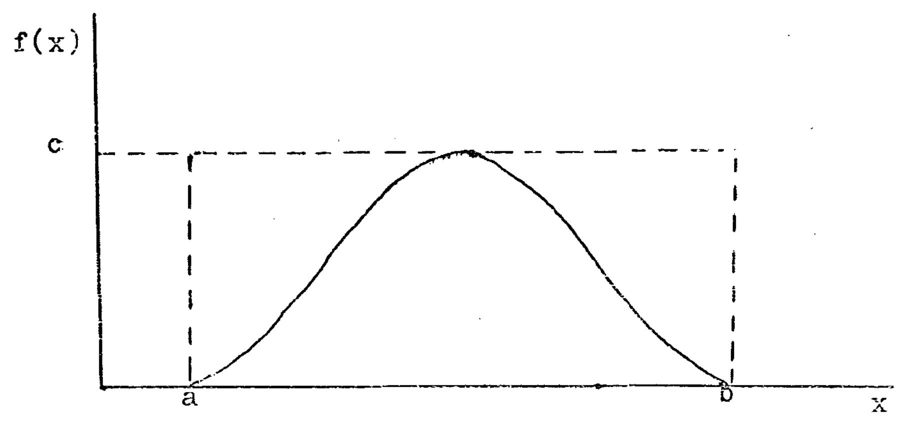
In addition to the table search method explained above, there are two main methods for converting a random number into a continuous random variate: the inverse transformation and the rejection methods.

- Inverse transformation: if $F(x)$ is the cumulative distribution function, $F^{-1}(x)$ is the inverse c.d.f. for the random variable x, and U is a random number, then a random variate is defined by $x = F^{-1}(U)$. The application of this method depends upon the existence of $F(x)$ in an explicit form and on the possibility of obtaining its inverse.

- Rejection method: if $f(x)$ is the probability density function of a distribution under study (see Figure 7) by means of generating two random numbers U_1 and U_2, the method is applied in the following three steps:

1) $x_1 = a+(b-a)U_1$
2) $x_2 = cU_2$
3) If $x_2 \leqslant f(x_1)$ then x_1 is the next random variate, otherwise repeat the process with two new uniform random numbers.

FIGURE 7. Rejection method.



When the area under the curve f(x) is too small in comparison with c(b-a), this method becomes very inefficient since a large number of random numbers has to be generated for every random variate. It can be seen that this method has the disadvantage that two or even more random numbers are to be obtained for each trial point.

Next, different methods of generating random variates are going to be considered for the following theoretical distributions: Poisson, exponential, normal and gamma distributions.

Poisson distribution. A method suggested by Kahn (Martin, 1968) multiplies N uniform random numbers U_i successively until the following inequality is satisfied:

$$\prod_{i=1}^{N} U_i < e^{-\lambda t}$$

Then, the respective random variate R is equal to N-1. One test of randomness for a series of M Poisson random variates is given by:

$$\frac{1}{M} \sum_{i=1}^{M} R_i \simeq \lambda t$$

and

$$\frac{1}{M} \sum_{i=1}^{M} R_i^2 \simeq \lambda t(\lambda t+1)$$

Exponential distribution. The cumulative distribution function of the exponential distribution is given by:

$$F(t) = 1 - e^{-\lambda t}$$

as the cumulative distribution function exists in explicit form, by means of the inverse transformation method:

$$e^{-\lambda t} = 1 - F(t)$$

$$\ln(e^{-\lambda t}) = \ln(1-F(t))$$

$$-\lambda t = \ln(1-F(t))$$

$$t = -\frac{1}{\lambda}\ln(1-F(t)) = -\frac{1}{\lambda}\ln(1-U)$$

where U is a uniform random number. The term (1-U) has value between 1 and 0, and as the values of U are uniformly distributed so are the values of (1-U). Then, it is possible to simplify the above expression:

$$t = -\frac{1}{\lambda}\ln U$$

Although the appearance of the formula is very simple, the computation time may be very long as the logarithm is obtained by means of a series that converge slowly (Gordon, 1978). Therefore, in order to save computer time it is possible to substitute the logarithm function by a table search, which on the other hand means more storage capacity and some loss of accuracy.

The elements of the table will be the logarithms of values between 0 and 1. The number of elements will depend upon the required accuracy, taking into account that the more elements the table has, the longer the search. According to the expression for a linear interpolation:

$$\ln U = \ln U_1 + (U-U_1)\ \text{slope}_{12}$$

where

$$\text{slope}_{12} = \frac{\ln U_2 - \ln U_1}{U_2 - U_1}$$

As an example Table 7 is given.

Given a random number U, the respective interval (U_1, U_2) is obtained by means of a table search and thereafter the intermediate value is calculated by means of the interpolation expression given above. It is worth noting that the value of slope_{12} is already tabulated and therefore it is not necessary to recalculate it. The following example illustrates how to apply the method.

TABLE 7. Table for the generation of the exponential distribution.

N	U	-lnU	slope	N	U	-lnU	slope
1	0	∞	23.0	8	0.7	0.36	1.4
2	0.1	2.30	7.0	9	0.75	0.29	1.4
3	0.2	1.60	4.0	10	0.8	0.22	1.2
4	0.3	1.20	2.8	11	0.85	0.16	1.0
5	0.4	0.92	2.3	12	0.9	0.11	1.2
6	0.5	0.69	1.8	13	0.95	0.05	1.0
7	0.6	0.51	1.5	14	1.00	0	

Source: calculations made by the author.

$$U = 0.63 \text{ , then } U_1 = 0.6 \text{ and } U_2 = 0.7$$

$$\text{slope}_{12} = \frac{\ln 0.7 - \ln 0.6}{0.7 - 0.6} = \frac{-0.36 + 0.51}{0.1} = 1.5$$

$$\ln 0.63 = \ln 0.6 + (0.63 - 0.6) \times (1.5)$$

$$\ln 0.63 = -0.51 + 0.045 = -0.465 \quad \text{(real value -0.462)}$$

One method of avoiding the search consists in using a table with number of elements equal to either 10 or a multiple of 10 for equal intervals of probability. Consequently, for a table of N elements and given a random number U, its logarithm is calculated by means of the following steps:

$$M = N \times U \qquad \text{(integer value)}$$

$$DIF = U - (M/N)$$

$$\ln U = \text{Ln } (M+1) + DIF \times SLOPE\ (M+1)$$

This method has been widely used throughout this research with satisfactory results.

Normal distribution. Although its cumulative distribution function cannot be expressed in simple mathematical terms, there is a way of generating normal variates using n uniform random numbers U_i and applying the following formula:

$$R = \frac{\sum_{i=1}^{N} U_i - (n/2)}{(n/12)^{\frac{1}{2}}}$$

where the range $6 \leqslant n \leqslant 12$ is sufficient condition, and the distribution of the sequence of numbers generated approaches a normal distribution with mean zero and standard deviation equal one. A test of randomness for a aequence of N numbers is given by the following expression:

$$\frac{1}{N} \sum_{i=1}^{N} R_i \simeq 0$$

and

$$\frac{1}{N} \sum_{i=1}^{N} R_i^2 \simeq 1$$

The formula mentioned above is very easy to program but has the disadvantage of using too many uniform random numbers. So, another useful method consists in generating two independent uniform random numbers in order to apply the following equations, which give two random variates from a standard normal distribution:

$$R_1 = (-2 \ln U_1)^{\frac{1}{2}} \operatorname{Cos} 2\pi U_2$$

$$R_2 = (-2 \ln U_1)^{\frac{1}{2}} \operatorname{Sin} 2\pi U_2$$

Another useful method that gives a suitable approximation is to utilise the table of the cumulative normal distribution in order to make a search. However, only with practical experimentation is it possible to determine the trade-offs between the different methods.

Gamma distribution. A method to generate gamma variates is given by the following expression:

$$t = -\frac{1}{\alpha} (\ln \prod_{i=1}^{K} U_i)$$

where K and α are parameters of the distribution and U_i is a uniform random number.

As the effect of the variability on the simulated system is reproduced by sampling from the appropriate probability distributions, one question that arises concerns the merit in fitting a theoretical probability distribution to data obtained empirically. It is reasonable to think that the use of raw empirical data implies a simulation of a past situation, while as was shown in this section, sometimes in the generation of random variates it is better to apply table search techniques rather than to use theoretical distributions due to the inefficiency of some subroutines to evaluate the necessary functions.

On the other hand, if theoretical functions are used it is much easier to change the parameters of the distribution in order to make a sensitivity analysis, for instance. Therefore, it seems to be that there is no particular usefulness in fitting a function to empirical data, and the analyst, as a matter of judgement, should decide this especially when the fitted function is of a simple form.

D. <u>Time keeping techniques</u>. As the purpose of a simulation model is to study the behaviour of a system with the passage of time it is essential to guarantee that all the events are taken into account in chronological order. There are two basic methods for updating clock time: fixed time and next event.

<u>Fixed time method</u>. It consists in choosing an appropriate time cell dt (1 second, say) and to initialise the time to start the simulation, t. Thereafter, time is increased in dt (time cell)

and theevents that occur during this period are examined and dealt with (it could be that no events occur during a time period). This continues until the period to be simulated is completed.

Next event method. The clock is advanced to the time at which the next event occurs, so the system is viewed as proceeding from one event to another until a prescribed sequence of events is completed.

Fixed time methods are efficient from the computational point of view when events occur on a fairly regular basis and in the study of systems whose events are not very well known, specially during the initial phases of the study of the system. On the other hand, there is a loss of information about the behaviour of the system when the size of the interval is inappropriate, eg. where events occur at one end of the interval.

Next event methods are efficient when events occur unevenly in time. They do not require a decision about the size of the time increment which can be probematical. They can save computer time when the system is static for long periods of time. However, the approach can be difficult to apply to complex systems in which there are many events occurring simultaneously.

E. Other aspects. In order to implement simulation studies on a digital computer it is necessary to outline the technology available. Computer languages are generally classified in two groups: general purpose languages (FORTRAN, COBOL, ALGOL, etc) and special purpose or simulation languages (GPSS, DYNAMO,etc).

Due to the diversity of languages available it would be almost impossible to determine the best one for a particular application and in general the analyst chooses a language known by him, taking into account the hardware and software configuration of the computer where the model is to be run.

However, it is important to point out that general purpose languages usually require longer programming times while the restrictions on formats of outputs are minimal. On the other hand, simulation languages provide reduced flexibility for modelling and longer running times, although they reduce the programming time.

The usefulness of simulation experiments for solving problems that are too expensive for experimental solution or too complicated for analytical treatment is undeniable. Simulation experiments can be conducted for a great variety of purposes such as determining how good a particular design of a system performs against some specific criteria, or comparing different designs in the way they carry out a specific function, etc.

However, the complexity of a simulation model and even the decision whether or not to use this technique rests very much on the consideration of costs and possible applications. Taking into account that the validity of a simulation is measured by the extent to which it satisfies its design objectives, the analyst should decide at the outset of the experiment the variables to be used. This must be done bearing in mind that too few variables may produce a poor representation of the system, while too many may cause problems of computer storage capacity.

A simulation model is an experiment that is said to be "run" rather than"solved", which means that the output obtained is for a particular set of variables and parameters used in the experiment. This is because simulation is a methodology rather than a theory of problem solving. So in determining which of the many factors are the most relevant in affecting the system performance it is necessary to run the model with different set of values in order to make a sensitivity analysis.

The need to run the model several times arises from the fact that simulation experiments do not provide functional relationships between the output and the independent variables and parameters of the system. The consequence of this is that an optimum solution of the model can only be found experimentally.

Finally, two significant reasons for using simulation in public transport operations are mentioned by J.J.Browne (1967): the first is the variability of the demand for transportation and the second is the need for detailed evaluations of queueing processes. In order to evaluate waiting times and queue lengths in a dynamic situation where demand is permanently fluctuating, to work with averages may not be sufficiently accurate and it may be necessary to determine the corresponding distributions.

The following sections will show the application of the concepts previously discussed to the different situations arising in modelling a public transport system.

2. A public transport system: two different points of view.

An urban bus system affects different sectors of the community and these sectors can be classified into three main groups: passengers, bus operators and society as a whole. If the bus system is to be improved the impacts upon such groups should be carefully assessed taking into consideration that what is good for one sector may not be as good for another, as every affected group will regard the system according to its own interests. For instance, the users will look for a reliable and comfortable service; the operators will try to maximise their profits; and the community will probably prefer a system that causes least disruption to traffic.

Because of their size, relatively low power/weight ratio, poor manoeuvrability, and because they have to stop frequently buses are more likely to cause congestion than most other vehicles on the road. Similarly, traffic in general may affect the performance of the entire transportation system. Thus, traffic can be seen as a direct consequence of the interrelation between supply of and demand for transportation and the use of the public transport system in a more rational way may tend to alleviate its undesirable effects.

On the other hand, additional effects on non-users and the environment include physical impacts such as aesthetics, noise and air pollution. As consideration of these factors is beyond the scope of this study, the analysis is limited to the first two groups, users and operators, and their appraisal of the system.

A basic task in any public transport planning process is to understand how users perceive the bus service since it may provide insight into the manner in which they use the system. The operator's side is important as his ability to use his resources will affect not only himself but other sectors of the community. Furthermore, the variability of several factors affecting the service provided by the operator as well as the variability of the demand may tend to cause some degree of uncertainty about the way the system will be finally utilized.

The purpose here is to analyse the bus system from two different points of view, the users' and the operators', emphasizing those aspects which were used in the design of the proposed model. Where a contribution was available from one of the alternative models which have been developed recently this is noted, and some of these models are discussed in greater detail later in this chapter.

A. Relevant aspects for the users. Many transportation studies have considered the trip-makers' socioeconomic background to be the most important determinant of their choice of transport mode, i.e. family income and car ownership. However, more recent studies direct more attention to the importance of the level of service offered by the transport system. A general measure of quality of service is given in terms of level of service indicators and comfort and safety factors.

Level of service indicators.

i) Availability, which has two facets:

- accessibility, in terms of distance to get to bus stop from trip origin and then from bus stop to trip destination

- frequency of service, in terms of time.

ii) Total travel time: the total door-to-door travel time is composed of four items: walking, waiting, riding and transfer times. Relative weights of these time intervals vary since passengers perceive them differently.

iii) Reliability: most commonly defined in terms of average waiting time of passengers at bus stops as users are more concerned with this uncertain time than any other aspect of the service (Chapman, 1976)

iv) Convenience: related to the route network, since passengers generally dislike transferences and consider a person/seat ratio which exceeds 1.0 as undesirable.

v) User cost: the fare to be paid.

Comfort and safety factors. These relate to subjective qualities which can never be completely available in a bus system, although licensing regulations usually set minimum standards. Whereas improvements in level of service may actually motivate potential travellers, it is unlikely that improved comfort and safety will have a significant impact on patronage (Alter, 1976).

From the passenger's point of view, perhaps the most important factors to be considered are how long it will take to get his final destination and how much it will cost so that he perceives the level of service in terms of cost in time and money of the various trips he is likely to make.

In the context of buses one frequently used measure of level of service that is intended to cover the broad range of features of a system is generalised costs. Most transportation studies have used a simplified approach owing to the difficulties in allotting values to some elements as comfort and convenience. For example it is possible to express a bus trip in terms of money:

$$C = C_w t_w + C_{wt} t_{wt} + C_r t_r + F(d) + C_t t$$

where: t_w, t_{wt} and t_r are walking, waiting and riding times respectively;

C_w, C_{wt}, and C_r are the respective costing times;

$F(d)$ is the fare to be paid if d is the distance trevelled;

C_t costing factor for transfers made;

t number of transfers

1) Walking time. Public transport does not generally offer a door-to-door service and consequently major components of journey time are taken up in access to or egress from the network. Walking may be considered a feeder mode to the bus system but it may also compete with it of course.

Assuming a constant walking speed, walking time can be regarded as proportional to bus stop spacing and to route spacing. In other words, the more bus stops and routes which are available at a particular zone, the less the time people spend walking. If the bus stops are evenly spaced at m stops per kilometre, u is the average walking speed and D is the depth of the catchment area in metres, then the average walking time is given by Chapman et al (1976) as:

$$AWT = \frac{1}{2u} \quad (D + 1000/2m)$$

This assumes that the network of streets is rectangular.

An alternative way of calculating the minimum walking distance from random points to a bus stop for a grid or radial network is given in the Bradford Bus Study (R. Travers Morgan and partners, 1976) by means of the following expression:

$$WK = \frac{L + S}{4} + \frac{1}{3D}$$

where WK is the walking distance in kilometres

L average spacing of streets intercepting the bus route in Kms

S average spacing of bus stops in Kms

D route density in Km/Km^2

However, passengers do not always use the route which passes closest to their origin. They sometimes walk further in order to widen the choice of routes and to reduce the waiting time or simply to reduce the fare. This suggests that all bus stops are not to be treated equally as some of them provide more accessibility in terms of routes available.

2) Waiting time. The average waiting time of passengers at bus stops is a function of the bus frequency, bus capacity and traffic conditions, but there is no evidence that suggests how, if at all, waiting time may be affected by the spacing of bus stops. It is often assumed that the average waiting time is half the bus service interval and this is based on the grounds that: i) passengers arrive at bus stops at random; ii) passengers can get onthe first bus that comes;and iii) vehicles arrive regularly.

Holroyd and Scraggs (1966) studying the average waiting time in Central London considered the two first assumptions to be acceptable while they did not find the third justifiable.

If h_i represents the bus headway during a period of study, where i varies from 1 to n, being n the number of intervals observed during the whole period. If it is assumed that passengers arrive at random, the probability that a passenger will arrive during a length of time h_i is given by:

$$\frac{h_i}{\sum_{i=1}^{n} h_i} \qquad \text{where} \quad \sum_{i=1}^{n} h_i \text{ is the length of period}$$

Given that a passenger arrives during an interval h_i and assuming that he can take the first bus that arrives, then the expected waiting time is $h_i/2$ and the average waiting time during the study interval is given by the summation:

$$AWT = \sum_{i=1}^{n} \left[\frac{h_i}{\sum_{i=1}^{n} h_i} \times \frac{h_i}{2} \right] = \frac{\sum_{i=1}^{n} h_i^2}{2 \sum_{i=1}^{n} h_i}$$

The general formula of the average waiting time may be expressed in terms of the mean and variance of the service interval h as:

$$\bar{w} = AWT = \frac{\bar{h}}{2}\left(1 + \frac{\mathrm{var}\ h}{\bar{h}^2}\right)$$

This equation is inapplicable in multi-route stops where some passengers have the possibility of using more than one service to do their journey. If buses arrive regularly, var h = 0 and then clearly $\bar{w} = \bar{h}/2$. If buses arrive in bunches of n regular intervals of length $n\bar{h}$, then var h = $(n-1)\bar{h}^2$ and $\bar{w} = n\bar{h}/2$. If buses arrive at random, varh = $\bar{h}^2$ and $\bar{w} = \bar{h}$. Then, the more regular the service the lower is the average waiting time and viceversa.

In the Central London study it was found that var h = $A\bar{h}^2/(A + \bar{h}^2)$, where A = 35 and then:

$$\bar{w} = \frac{2A + \bar{h}^2}{2A + 2\bar{h}^2} \times \bar{h}$$

It was found that when service interval is below about three minutes, $\bar{w} \simeq \bar{h}$, which means that there is a tendency for both passengers and buses to arrive at random. For service intervals between 3 and 12 minutes there is a tendency for the buses to arrive more regularly and then $\bar{w} \simeq 1/2\bar{h}$.

O'Flaherty and Mangan (1970) consider that the expression used by Holroyd and Scraggs (1966) cannot be applied during peak hours as the assumption that passengers arrive randomly is probably not valid. They obtained some data in Leeds during the peak period and fitted a straight line regression to the observations:

$$\bar{w} = 1.79 + 0.14\,\bar{h} \qquad\qquad r = 0.63$$

It is worth noting that the slope 0.14 is less than slopes for equations relating random arrivals to random bus arrivals $\bar{w} = \bar{h}$, and regular bus arrivals $\bar{w} = \bar{h}/2$. This suggests that passengers do not arrive randomly.

Seddon and Day (1974) found with good results a polynomial expression to calculate the average waiting time in Manchester. This was of the form:

$$AWT = a + b\frac{\sum h_i^2}{\sum h_i} - c\left(\frac{\sum h_i^2}{\sum h_i}\right)^2$$

where a, b and c are constants.

This empirical expression is not applicable for multi-route stops either and in such cases it is necessary to estimate the proportions of passengers using each service by means of a survey.

They concluded that for low headways (10-12 min) the arrival of passengers at bus stops could be regarded as random. But for longer headways people try to time their arrival to catch a specific bus.

Passenger waiting time at bus stops is an important measure of the level of service provided. Passengers usually consider the time spent walking or waiting more inconvenient than that spent on the bus. Of course, the problem of evaluating the waiting time becomes more complex when more than one route passes through a specific stop, when routes multiply and traverse the same path for parts of their lengths and when a direct journey can be made by more than one route.

It is important to conclude from the previous considerations that given a random pattern of arrival of passengers at bus stops and a bus frequency, the average waiting time will vary basically with regularity of service, which suggests the importance of achieving high standards of regularity in the whole transportation system.

3) Riding time. The time passengers spend riding in a bus may be regarded as composed of three elements: bus running time, bus stop time and time due to delays. There is some evidence that suggests that people in general prefer to spend their time more on board a bus than walking or waiting for a service. This is the reason why the Bradford Bus Study found that walking and waiting times could reasonably be weighted by 1.75 and 2.25 respectively in order to be compared with riding time.

The average riding time spent by a passenger on a bus could be seen as a function of the distance between origin and destination, bus stop spacing, bus route spacing, running speed of buses, delays experienced at bus stops and traffic conditions on the route. Time spent by the user on the system tends to be considered in relative rather than in absolute value as that figure is compared to the time making the trip by car instead of by bus (Hobeika, 1975).

4) Fare. There are two main types of bus fares: those dependent on the distance to be travelled and flat fares. The former offer the most direct relationship between service and price though a minimum charge is imposed. Flat fares provide some advantages for the operator of the service and those who make longer journeys.

The average price elasticity of demand for urban bus trips is -0.3 with a range -0.1 to-0.6 (Bly, 1976), which means that a fare increase of 10% will cause passenger trips to decline by 3%. Similarly, a reduction in the price of transit service will not attract as many new riders as is sometimes thought. On the other hand, there is some evidence that fare increases in a fare-stage system are likely to have more effect on the total passenger-kilometres than on the total number of trips because some people tend to reduce the increase effect by alighting earlier (Bly,1976).

Finally, since generalised costs are considered to be a key determinant of passenger demand it means that any measure which reduces walking, waiting, riding and fare costs is a potential service improvement from the user's standpoint. However, it is essential to take into account that a public transport system operates on a balance of costs and services and by manipulating those elements some specific objectives may be met. Therefore, the purpose of the following section is to present the other side of the coin: the operator of the system.

B. <u>The operators' view</u>. Given certain constraints a public transport operator is in almost total control of the manner in which he uses his resources in order to provide a service. He may be concerned with the amount of output he can obtain from a set of inputs, and a particular combination of inputs would determine the output/input ratio which in turns define the efficiency of the system. In this way, efficiency refers basically to the operator's ability to utilize his resources in the provision of a service.

Although the main objective of the operator of the system is to maximise profits or to seek to cover its costs, it may not be sufficient to guarantee the provision of a good and reliable service to travellers. Therefore, the operator should consider not only the efficiency of the system but also its effectiveness in providing mobility according to some particular conditions and necessities.

Within this framework and taking into consideration the conflicting objectives of his task, the operator may define his goals and policy with respect to the operation of the system. He can for instance:

i) try to maximise the level of service keeping costs constant;

ii) try to minimise his costs of operation keeping the level of service constant; or

iii) maximise the level of service for a pre-determined level of profit or loss.

In any case he can play with different factors which are under his control in order to achieve his specific goals: bus route network, bus route spacing, bus stop spacing, frequency of service, bus size, fares, measures of traffic engineering, and so on.

The absence of a defined criterion to evaluate a bus route network has resulted in traditional analysis being very subjective. Some points to be considered in the design of a route network are:

- Number of journeys which require transfer between services should be minimised;

- Routes should be easy to understand, implying that they should be as short as possible. Short routes are to be preferred because long and complex ones may be difficult to operate at reasonable levels of schedule adherence;

- The route system should facilitate transfers with other modes of transport;

- The system should not have too many routes so as to facilitate understanding and to have an adequate frequency of service;

- Routes should be as direct as possible;

- Routes should cover operating area well; and

- Routes should serve as many traffic generators as possible, and specially try to have a traffic generator at both ends of the line.

Distances between bus stops may affect the user's travel time in different ways as there is a trade-off between walking and riding times: close bus stops imply less walking and more riding time and viceversa. However, from the operational standpoint extra bus stops will lead to increases in operating costs due to longer journeys. An expression to determine the number of bus stops per kilometre is given by Chapman et al (1976):

$$M_{opt} = \sqrt{22/L}$$

where M_{opt} is the number of stops per kilometre

L is the length of journey in kilometres.

According to the previous expression the longer the journey the greater the optimum distance between stops as the longer walking time represents only a small proportion of the total journey time. Another measure related to stops is the possible inclusion of limited stops services which could lead to the provision of faster journeys for longer distance passengers and a better utilisation of the resources at the expense of disbenefits to some other passengers.

With respect to the bus route spacing Chapman et al (Sept -1977) developed a procedure to find the optimum spacing. To do so they consider a corridor of travel length of L kms of such width that it takes P minutes to cross it on foot. Then, for R equally spaced routes the average walking time in minutes is given by P/4R if people walk direct to the nearest route. On the other hand, there will be M/2LR buses per route per hour available in each dierction if there are M bus-kilometres per hour in the corridor, which means an average headway of 120LR/M minutes. Therefore, assuming random arrival of passengers at stops and regular service, then the average waiting time is given by 60LR/M minutes. Thus, the number of routes in the corridor which minimises the sum of walking and waiting time is given by the following expression:

$$R_{opt} = \sqrt{PM / 240L}$$

It could be shown that the number of routes which makes walking and waiting times equal is also given by the previous expression. In a similar way, given some cost constraints it is possible to make trade-offs between the frequency of service and the system accessibility, as a denser network may imply decreased bus frequencies and viceversa. The immediate effect of an increase in bus frequency is the reduction of the mean waiting times and hence of the generalised costs of travel by bus for the passenger. It usually also leads to an increase in patronage, revenue and costs of operation. On the other hand, the Bradford Bus Study found that a reduction of bus frequencies during peak periods would yield net savings to the operator exceeding almost 5 fold the disbenefits to the users due to increases in waiting time.

The size of buses can have enormous repercussions on the whole system: if large buses are used then fewer will be needed for a given load and therefore fares and road congestion can be reduced. However, if small buses are used it is possible to improve the coverage and frequency of service reducing in this way walking and waiting times.

Owing to the feeble elasticity of transport demand there is limited advantage to the operator in reducing the fare. However, there is some evidence that suggests that low fares may persuade people who walk or cycle to switch to the bus system to make short journeys during the peak periods and they may also encourage more shopping, social and pleasure trips that would normally take place during off-peak periods (T.R.R.L., 1974).

Increases in the fares charged has been one of the main components considered by the operator with the view to restoring the profitability of the undertaking. Although this policy has led in general to a drop in patronage in most cases it has produced an increase in revenue (Bly, 1976).

Costs are probably the main factor for the bus operator and for this reason he is especially concerned with operating speeds because they affect the stock of the equipment, labour costs, fuel maintenance, and so on. In other words they affect the inputs to the system and therefore the attraction of passengers to the service. Within this context the operator could try to speed up the fare collection process, for instance, or to join efforts with transport planners and traffic engineers with the view to making more efficient the operation of the system.

The bus operator working in conjunction with the transport planner could improve the conditions of travel by using bus priority techniques such as bus lanes that could increase bus running times reducing in this way the basic equipment and drivers requirements for bus operations. However, this kind of measures are to be taken cautiously as a recent study in California found that separate bus lanes would actually increase rather than decrease total person delay as the time saved by bus passengers is much less than the time lost by car users who are actually denied the use of the relatively empty bus lane (Myers, 1967).

On the other hand, it is known that transport demand is ruled by cyclic fluctuations and as a result it becomes sometimes necessary to provide a larger amount of output in order to meet the peak demand, and usually this extra capacity is underemployed during the rest of the day. Then, other obvious possibility of improving operating efficiency is to balance, as much as possible, peak demand with greater off-peak demands.

It could be said that any improvement in the elements of supply such as those mentioned before could normally deteriorate the financial situation of the undertakings if there is no guarantee of an increase in ridership. However, the bus operator could try to use his resources in a more productive manner in such a way that he could use saved inputs in areas of more need. For instance, by estimating the future demand for public transport and establishing the shortcomings in the actual system to meet this future demand it would be possible to achieve an adequate and better match between supply and demand, increasing in this way the quality of service and decreasing the operating costs. An improvement in the quality of the service by these means will keep pace with the rising standard of living of the population and could attract, with low capital investment, more people to the service.

Perhaps the most important requirement for a public transport system is reliability since this is reflected in the waiting times incurred by passengers at bus stops. With an unreliable service people have to set off earlier than otherwise in order to avoid the possibility of being late, increasing consequently their travel time. Unfortunately bus services tend to be unreliable. This is mainly due to variations in: passenger arrival pattern, boarding and alighting times, inter-stop travel time, penalties for stopping, variables associated with crews,and so on. In order to improve service reliability it is advisable to reduce as far as possible variations associated with these sources and to allow for those variations which cannot be removed completely (Chapman et al, Aug-1977)

i)<u>Passenger arrival patterns at bus stops</u>. The usual way of describing an arrival pattern is in terms of the inter-arrival time between successive arrivals. For an arrival pattern with no variability the inter-arrival time is given by a constant while for arrivals that vary stochastically it is necessary to define the respective probability function.

With respect to arrival patterns one can distinguish between the arrival of passengers at stops and their possible association with expected departure of buses and the demand throughout the day which may have more definite patterns and whose changes may be more easily predicted. The first problem will be considered here.

According to Coe and Jackson (1977) passenger arrivals fall into three categories: random arrivals, timed arrivals and a third type to include those people who see a bus coming and run to the bus stop arriving more or less coincidentially with it.

As noted earlier, it is possible to assume that passengers arrive randomly for those services with low frequencies (less than 12 minutes) when there is no evident association between arrival times and expected departure of buses from stops, or between arrivals of other people.

Random passenger arrivals are usually described by means of a Poisson distribution whose probability function gives the probability of n arrivals during a time interval t, given a mean rate of arrivals per unit time interval λ. By generating a uniform random number and using the Poisson cumulative distribution it is possible to obtain the number of arrivals during the time period t, and by repeating this operation several times it is possible to simulate arrivals during a period of study T.

However, if what is required is the exact time of arrival and not the total number of arrivals per unit of time, then it is necessary to work with the exponential distribution which is regarded as the continuous analogue of the Poisson. The inter-arrival times between passengers can be generated by using the expression obtained before:

$$t = -\frac{1}{\lambda} \ln U$$

where U is a uniform random number.

In order to illustrate the scope of these techniques a complete example has been developed in which the inter-arrival times between passengers are generated using the previous expression for $\lambda = 0.021$ arrivals/second (taken from Table 1) The inter-arrival times and the respective arrival times are given in Table 8 for one hour of simulation. the series of random numbers was generated by means of a pocket calculator.

Table 8 was generated assuming that passengers arrive according to a Poisson distribution and it is possible to check the validity of this hypothesis by assuming a time interval t of 1 minute so that the 60 time periods can be classified according to the respective number of arrivals which they contain. This is shown in Table 9. For a confidence level of 0.90 and two degrees of freedom $\chi^2_T = 4.61 > 0.622$ which means that the sample is consistent and can be regarded as being Poisson distributed.

ii) Boarding and alighting times. Chapman (1975) considers that the time spent by a bus at a bus stop can be regarded as composed of two elements. One is the time associated with opening and closing doors, and the other is the boarding and alighting times. Times taken by passengers to alight and to board when a bus has arrived at a bus stop will depend on the bus design (simultaneous or consecutive process) and on the ability of driver and passengers.

For one-door buses:

$$T = C + \sum_{i=1}^{m} a_i + \sum_{j=1}^{n} b_j$$

For two-door buses:

$$T = C + \max \left[\sum_{i=1}^{m} a_i , \sum_{j=1}^{n} b_j \right]$$

where:

T is the stop time at bus stop

C is a constant

a_i is the time to alight by ith passenger; and

TABLE 8. Example of arrivals and inter-arrival times for one hour of simulation. (in seconds)

Arrival No	Inter-arrival time	Arrival time	Arrival No	Inter-arrival time	Arrival time	Arrival No	Inter-arrival time	Arrival time
1	2	2	27	188	1426	52	15	2501
2	83	85	28	29	1455	53	2	2503
3	114	199	29	16	1471	54	20	2523
4	4	203	30	20	1491	55	5	2528
5	62	265	31	26	1517	56	8	2536
6	199	464	32	141	1658	57	20	2556
7	40	504	33	24	1682	58	16	2572
8	7	511	34	53	1735	59	9	2581
9	49	560	35	57	1792	60	29	2610
10	54	614	36	75	1867	61	99	2709
11	4	618	37	60	1927	62	77	2786
12	41	659	38	44	1971	63	3	2789
13	72	731	39	72	2043	64	5	2794
14	2	733	40	4	2047	65	42	2836
15	19	752	41	14	2061	66	150	2986
16	29	781	42	8	2069	67	58	3044
17	6	787	43	53	2122	68	8	3052
18	184	971	44	62	2184	69	34	3086
19	24	995	45	0	2184	70	51	3137
20	60	1055	46	16	2200	71	11	3148
21	46	1101	47	19	2219	72	44	3192
22	31	1132	48	77	2296	73	48	3240
23	16	1148	49	113	2409	74	2	3242
24	24	1172	50	2	2411	75	29	3271
25	21	1193	51	75	2486	76	30	3301
26	45	1238						

Source: Calculations made by the author.

b_j is the time to board by jth passenger.

Sometimes it is assumed that all passengers take about the same time to board, without taking into account the number of passengers in the queue and their position in it. Consequently, if boarding time is more important than alighting time, then:

$$T = C + Bn$$

where: B = average value of b_j in other words is the marginal boarding time

n = number of passengers boarding.

TABLE 9. Observed and expected frequencies of data gathered in Table 8.

Arrivals	Observed		Expected	
0	19		17	
1	19		21	
2	13		14	
3	6	9	6	8
4	2		2	
5	1		0	
	60		60	

Source: Calculations made by the author.

For a one-man operated service surveyed in Newcastle upon Tyne, Chapman et al (1977) obtained the stop time of a bus T, as a function of the number of passengers boarding n_b by means of the following expression:

$$T = 3.5 + 4,1\ n_b \qquad \text{T in seconds}$$

They also found tremendous variations in the times taken to board a bus, for instance the time taken for four people varies between 5 and 40 seconds. These variations are explained by the agility of people and their preparedness to tender their fare, by the need for change and by the skilfulness of the driver in issuing tickets. By means of a similar analysis the following expression to find total alighting time was found:

$$T = 4.7 + 1.4\, n_a$$

Kraft and Deutschman (1977) analysed some distributions of passenger service time by means of photographic techniques and then simulated, assuming an Erlang function. For this purpose, the means and variances of the service-time distribution for each successive person boarding a bus were calculated in order to determine the mathematical function that represented the distribution. In this sense the boarding process at a bus stop is regarded as a single-server queueing phenomenon.

The boarding time is represented by a negative exponential function, which is a special case of the Erlang function:

$$p(g \geqslant t) = \left[\sum_{i=0}^{K-1} K(t-\Upsilon) \,/\, (\bar{t}-\Upsilon)^{\,i}\; e/i\,! \right]^{-K(t-\Upsilon)/(\bar{t}-\Upsilon)}$$

where:

$p(g \geqslant t)$ is the probability that time g is greater than or equal to time t;

K is a positive integer which was found to be equal to the number of doors of vehicles;

t is any service time;

$\bar{t}$ is the average service time; and

Υ is the minimum service time, which is approximately half the average service time $\bar{t}$.

Then, for a vehicle with one door, the boarding time is represented by the following function:

$$p(g \geqslant t) = e^{-(t-\Upsilon)/(\bar{t}-\Upsilon)}$$

Newell and Potts (1964) suggest that the pairing of buses is mainly due to the variation in time taken to load passengers. When a bus is delayed, more than the usual number of people will be at stops and loading will take longer which results in the bus becoming further and further behind schedule. On the other side, the next bus will find less passengers to pick up which means shorter loading times and a gain on schedule until it eventually catches the first bus. Similarly, the following buses will tend to pair. A ratio K which reflects the strength of the tendency of buses to pair is given by:

$$K = \frac{\text{passenger arrival rate}}{\text{passenger loading rate}}$$

Potts and Tamlin (1964) found experimental support for the suggestion that variation in loading times is the main cause of bus pairing while inter-stop travel variation and variation of passenger arrival rates only tend to conceal that effect.

iii) Inter-stop travel variation. Inter-stop travel variation could be regarded as being a function of the technical characteristics of the equipment, the traffic conditions affecting the bus and driver's ability. However, it could be said that between two points there is a minimal travel time which is not possible to reduce due either to technical characteristics of the vehicle or to traffic regulations.

A multiple linear regression made for links on a Newcastle upon Tyne route (Chapman et al, Aug-1977) shows the standard deviation of link travel time S_t, as a function of the mean travel time $\bar{t}$ in seconds and the length of the links l in kilometres:

$$S_t = 9.8 + 0.38\ \bar{t} - 43\ l \qquad \text{in seconds}$$

This relationship suggests that for links of equal length S_t will increase with increases in $\bar{t}$, while for links where $\bar{t}$ is equal S_t will decrease for longer links.

Given a pair of bus stops, travel time between them will be equal to the average travel time plus or minus a quantity which could be regarded as random. Within this framework. some models have treated the variations in travel time t by means of statistical distributions-mainly the normal and the gamma (Jenkins, 1976). As it is expected that buses running closely may find similar traffic conditions then some of the models also include a correlation factor to take into account this consideration, although its inclusion remains unresolved. Jackson (1972) and Gerrard and Brook (1972) developed correlation factors for their simulation models and they are presented in the next section.

The models that have assumed the variations between buses in their average speed between two nodes to be normally and independently distributed have had to impose arbitrary bounds on t in order to agree with the observations. Other models have assumed that variations in t follow a gamma distribution defined by two paremeters t_o and p:

$$f(t) = p^2\ (t-t_o)\ e^{-\ p(t-t_o)} \qquad , \ \bar{t} = \frac{2}{p} + t_o$$

where t_o is the minimumpossible travel time.

iv) <u>Penalty for stopping</u>. It is also necessary to take into account the extra time incurred because of stopping at bus stops. If a bus is running from stop I to stop J and has to stop at both stops, then the penalty for stopping is the extra time that this bus takes to cover the distance in comparison with the time this same bus would take if no stops were made. This extra time is given by the following expression obtained by Chapman et al(Sept-1977).

$$\text{penalty for stopping} = \frac{v}{2}\left(\frac{1}{a} + \frac{1}{d}\right)$$

for v = average cruising speed

a = average acceleration rate

d = average deceleration rate

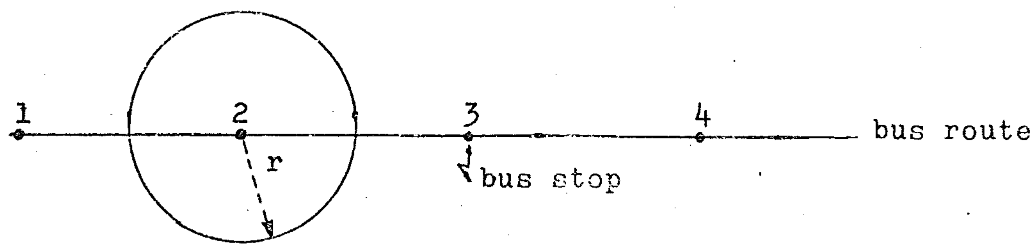
Finally, it may be said that there are some other factors related to crews affecting service reliability such as indiscipline, impunctuality, etc that also tend to make the service more irregular; however, they are beyond the scope of this study.

3. Important aspects of selected models in public transport planning.

The purpose of this section is to describe the most relevant aspects of some selected models which have been developed recently. The idea is not to summarize the theoretical basis of these models but to point out the most useful facets for the development of the proposed model which is to be explained later on. These models range from deterministic (analytical models) to simulation; however, no attempt at classification is made as some of them can be regarded as a mixture.

A. Optimum bus stop spacing. Lesley (1976) studied the factors affecting bus stop spacing from two different approaches : i) passenger generalised cost and ii) total community cost in providing a bus service. A bus route with equidistant bus stops is considered in the analysis (see Figure 8). The catchment area surrounding each bus stop is assumed to be homogeneous and g trips/unit area/unit time are generated.

FIGURE 8. Diagrammatic representation of a bus stop catchment area.



i) Passenger generalised cost.

The passenger generalised cost is assumed to be composed of three elements: riding time, T; excess time, E, or walking and waiting time and; fare, F.

Riding time. The movements of a bus between two bus stops are depicted in Figure 9, assuming constant maximum speed, acceleration and deceleration. Then, the total start to start time can be expressed by:

$$\tau = t_1 + t_2 + t_3 + t_4$$

where: t_1 = acceleration stage after leaving stop,

t_2 = period of uniform speed,

t_3 = deceleration stage arriving at bus stop, and

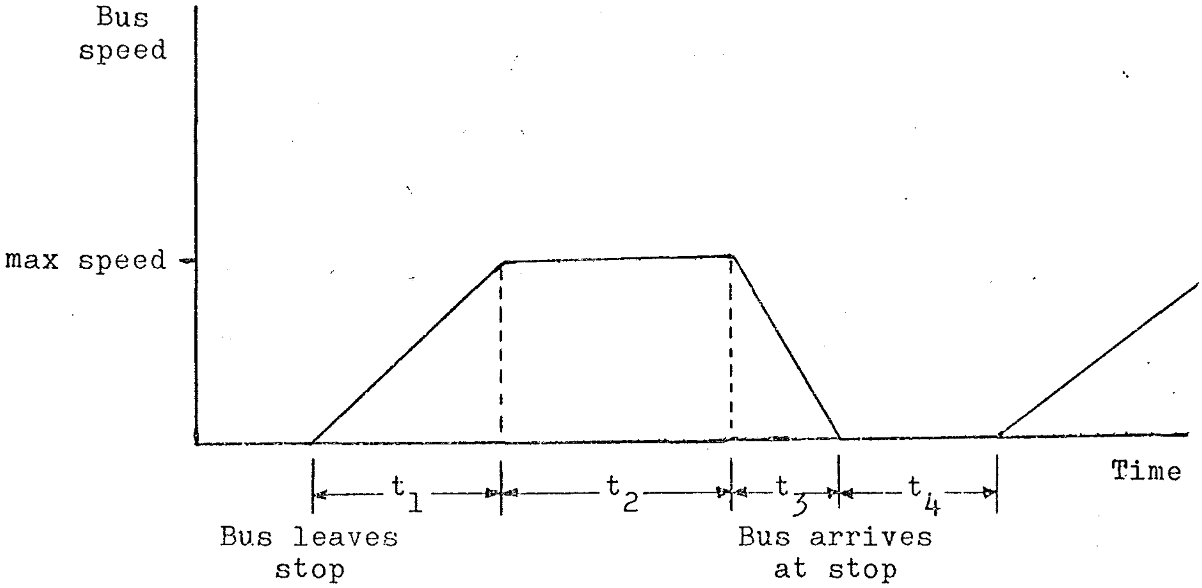
t_4 = boarding and alighting time.

The average operating speed is then $v_1 = 2r/\tau$, and if l is the average trip length, the average riding time is given by $T = l/v_1$.

Buses accelerate at rate α and decelerate at rate β, and the maximum speed is v. Then, t_1 and t_3 are given by:

$$t_1 = v/\alpha \qquad t_3 = v/\beta$$

FIGURE 9. Description of bus movements between two stops.



The distance travelled by bus at maximum speed is s_2:

$$s_2 = 2r - (s_1 + s_3) = 2r - \frac{v^2}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)$$

Then t_2 will be:

$$t_2 = \frac{2r}{v} - \frac{v}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right)$$

If f is the frequency of service, the number of passengers boarding will be:

$$n = f g \pi r^2$$

As it is considered that boarding time is greater than alighting time, time at bus stop can be given by:

$$t_4 = C_1 + C_2 n = C_1 + C_2 f g \pi r^2$$

where: C_1 = time lost, constant at each stop

C_2 = incremental boarding time per passenger.

Therefore, the average riding time of passengers is:

$$T = \frac{q}{2r}\left[\frac{v}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)+\frac{2r}{v}+C_1+C_2 fg\pi r^2\right]$$

Walking and waiting time, E. The average walking distance at each end of a bus journey is $2r/3$ because the catchment areas were assumed to be homogeneous. If w is the walking speed, the walking time is $2(2r/3w)$. The average waiting time was approximated by f/2. As it is considered that people prefer to be in the bus than walking or waiting, the previous values were weighted using a factor of 2, and then the excess time is given by:

$$E = \frac{8r}{3w} + f$$

Fare, F. The generalised cost G' of the average journey will then be:

$$G' = e(T + E) + F$$

where e is the monetary value of time.

$$G' = \frac{eq}{2r}\left[\frac{v}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)+\frac{2r}{v}+C_1+C_2 fg\pi r^2\right]+\frac{8er}{3w}+ef+F$$

By differentiating G' partially with respect to r it is possible to obtain an expression of the optimum bus stop spacing: $\frac{dG'}{dr} = 0$

$$(1) \qquad r = \sqrt{\frac{\left[\frac{qv}{4}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)+\frac{qC_1}{2}\right]}{\left[\frac{qfg\pi C_2}{2}+\frac{8}{3w}\right]}}$$

2r = optimum bus stop spacing

ii) Community cost of bus service.

It was assumed that the total community cost was constituted by the passenger total travel time in monetary terms plus the total operating cost of the service. In other words:

$$H = n_2 e\ (T+E) + D$$

where: H = total annual community cost of providing a bus service,

D = annual operating cocts,

n_2 = total number of journeys made per year on the bus route.

$$n_2 = 3600\ \frac{\text{sec}}{\text{hour}}\ \left(\frac{L}{2r}\right) \times (g\pi r^2) \times p = 1800\, Lpg\pi r$$

where L = total length of bus route

p = number of hours per year.

Now, $D = D_1 + D_2 N$

where: D_1 = fixed cost

D_2 = incremental cost per bus

N = total number of buses on route, which is given by:

$$N = \frac{2L}{v_1 f} \quad \text{and then} \quad D = D_1 + \frac{2\, D_2\, L}{v_1\, f}$$

Therefore, the total annual community costs are given by the following expression:

$$H = 1800\ pLg\pi re \left\{\frac{q}{2r}\left[\frac{v}{2}\left(\frac{1}{\alpha}\ \frac{1}{\beta}\right) + \frac{2v}{r} + C_1 + C_2 fg\pi r^2\right] + \frac{8r}{3w} + f\right\}$$
$$+ D_1 + \frac{2D_2 L}{2fr}\left\{\frac{v}{2}\left(\frac{1}{\alpha}\ \frac{1}{\beta}\right) + \frac{2r}{v} + C_1 + C_2\, fg\pi r^2\right\}$$

From $\frac{dH}{dr} = 0$ the following expression was derived:

$$(2) \qquad 0 = r^3\left\{C_2\, fg\pi + \frac{16}{3w}\right\} 1800\, eg\pi p + r^2\left\{\frac{qpe}{v} + \frac{C_2 D_2}{1800} + pef\right\}$$
$$1800\, g\pi - \frac{D_2}{f}\left\{\frac{v}{2}\left(\frac{1}{\alpha}\ \frac{1}{\beta}\right) + C_1\right\}$$

Using typical values and the expression (1), it was found that the optimum stop spacings were 92 metres and 196 metres for peak and off-peak periods respectively. By means of a sensitivity analysis it was also found that the most important variables affecting the functions were g, C_2 and f in the peak period, while v was in the off-peak period. Using expression (2), the sensitivity analysis found f and g to be the most important variables in the peak period, while g, v, e and p were in the off-peak period.

B. Optimisation of a simple bus network. This model, developed by Bly and Oldfield (1974) determines the passenger generalised cost of travel in terms of walking, waiting and riding time and the bus fare. The bus service is defined in terms of the average spacing between routes, the average density of buses per unit area and the average spacing of bus stops. The optimum bus service is that which minimises the total generalised cost of travel, C.

The average generalised cost of travel per passenger C, is given by the following expression:

$$C = C_w t_w + C_{wt} t_{wt} + C_r t_r + F$$

where: C_w, C_{wt}, and C_r = perceived cost of walking, waiting and riding respectively,

t_w, t_{wt}, and t_r = average passenger walking, waiting and riding time respectively,

F = bus fare in monetary terms.

The model expresses t_w, t_{wt}, t_r and F in terms of the different variables that affect a bus service. Within this set of variables, the spacing between routes, the density of buses per area and the bus stop spacings are considered to define the bus service. Then, optimum values for these variables can be obtained by differentiating C with respect to each variable and setting the derivative equal to zero.

Average walking time to and from bus stops. Here it is assumed that the density of trip origins is constant and that the pattern of travel is isotropic. The bus route network is square, of spacing L Km. The walk to the bus stop can only be made in directions parallel to the route network.

If a passenger always uses the route closest to his origin, the average walking distance to the route will be L/6. Here it is assumed that the route selected is not necessarily the nearest one, but the one which will take the passenger to his specific destination, and then the average walking distance to the route will be L/4. It is considered that the distance walked parallel to the route will be on average a quarter of the bus stop spacing. Then, the walking distance from origin to bus stop will be $\ell/4 + fL/4$, where ℓ is the average bus stop spacing and f is a constant factor. If V_w is the average walking speed, then the average walking time spent at both ends of the route will be:

$$t_w = \frac{1}{2V_w}(\ell + fL)$$

Average waiting time at bus stops. The average waiting time of passengers at bus stops is proportional to the average bus headway H:

$$t_{wt} = KH \qquad \text{where K is an appropriate coefficient}$$

If V_b is the mean speed of buses along all routes including stopped and layover times, the average separation of the buses will be $V_b H$. The sum of the separations between all buses in the in the area A, NV_bH is equal to twice the total route length in the area. (Where N is the number of buses available in the area).

Considering one cell of side L, the total length of route is 4L, and there are A/L^2 cells within the area. Therefore, the total route length will be $1/2\ (4L \times A/L^2)$, where the factor 1/2 is included because each route borders two cells. Then, from the previous reasoning:

$$NV_bH = 2\ (2A/L)$$

$$H = \frac{4}{MLV_b}$$

where M = N/A, density of buses per unit area.

But the expression to calculate the average waiting time does not take into account the bus capacity and then a factor studied by Oldfield was included:

$$t_{wt} = \frac{4K}{MLV_b}\left(1 + \frac{X}{Y - P/S}\right)$$

where: X and Y are constants which depend on the considered situation,

P is the average occupancy of a bus,

S is the bus capacity.

If j is the number of trips generated per unit area per unit time and J is the average distance travelled per passenger, then:

$$j\,J = \left(\frac{NP}{A}\right) V_b \quad , \quad \text{therefore}$$

$$(3) \qquad P = \frac{j\,J}{M\,V_b}$$

The average distance travelled in bus per passenger J, is a function of the bus stop spacing ℓ and the route spacing L, because the further the passenger walks the shorter the distance travelled by bus. If J' is the mean airline distance from origin to destination and r is an appropriate factor such as rJ' is the distance actually travelled on the road network, the average distance travelled by bus is given by:

$$J = rJ' - q \times \text{average walking distance}$$

$$J = rJ' - q/2(\ell + fL) \tag{4}$$

where q is an appropriate constant.

On the other hand, the mean running speed V_b depends on the free running speed of the bus U_b, the delay at bus stops and the layover time at termini. Therefore, the time that a bus takes to cover a distance D could be expressed in the following way:

$$\frac{D}{V_b} = \frac{D}{U_b} + \frac{D\,a}{\ell} + BD$$

where: D/ℓ is the number of bus stops in distance D,

a is the time at bus stop excluding boarding and alighting time,

B is boarding time per distance.

Thus,

$$\frac{1}{V_b} = (1 + \alpha)\left[\frac{1}{U_b} + \frac{a}{\ell} + B\right] \tag{5}$$

where the layover time is expressed as a factor α of the total time. Therefore, the average waiting time at bus stops is given by the following expression:

$$t_{wt} = \frac{4K}{MLV_b}\left(1 + \frac{X}{Y - P/S}\right)$$

where: V is given by equation (5)

B is given by equation (3), and

J is given by equation (4)

<u>Average riding time.</u>

$$t_r = \frac{J}{V_b} = J\ (1 + \alpha)\left[\frac{1}{U_b} + \frac{a}{\ell} + B\right]$$

<u>Bus fare.</u> If b is the mean cost per unit time of providing a bus and a crew, Mb will be the cost of providing the M buses in the area which originates J passengers per time. Then, the average bus fare is given by:

$$F = (1 + x)\ \frac{Mb}{j}$$

where x makes the percentage of profit to the company.

In this way, the average generalised cost of travel per passenger can be obtained by the following equation:

$$C = \frac{C_w}{2\,V_w}(\ \ell + fL) + \frac{4C_{wt}\,K}{MLV_b}\ (\ 1 + \frac{X}{Y-P/S}\ ) + \frac{C_r\,J}{V_b} + (\ 1+x)\ \frac{Mb}{j}$$

In order to obtain the optimum values of the average spacing between routes L, the average density of buses per unit area M, and the average spacing of bus stops ℓ, the previous equation is differentiated with respect to each parameter and the derivatives are set to zero. After some approximations it was obtained:

$$\ell_{opt} = \left\{\frac{2\,a}{\frac{C_w}{V_w} - \frac{q\,C_r}{V_b}}\left[\frac{4(1+\alpha)\,K\,C_{wt}}{LM}\ (\ 1 + \frac{X}{Y-P/S}\ ) + C_r\,J\right]\right\}^{\frac{1}{2}}$$

And the optimum values of L and M are obtained by trial-and-error and subsequent iteration from the simultaneous solution of:

$$(MY - D)^2 \left(\frac{1}{M^2} - EL\right) + XY = 0$$

and

$$L^2 - \frac{G}{M} = 0$$

where:

$$D=\frac{jJ}{V_b S} \quad ; \quad E=\frac{(1+x)bV_b}{4C_{wt}Kj} \quad ; \quad G= \frac{8C_{wt}K\,(1+ X /[Y - P/S])}{fV_b \left(\frac{C_w}{V_w} - \frac{qC_r}{V_b} \right)}$$

This model also considers situations where the travel demand has peak and an appropriate expression is derived to take into account such cases.

C. Model for Wallasey's bus service. Lampkin and Saalmans (1967) carried out an operational research study in order to replan Wallasey's bus service which was losing about 5% of its traffic each year. The following factors were considered for possible alteration: the route network, the frequencies of service, the timetables, the bus schedules and the crew schedules.

As it was important to achieve a reasonable economic performance without forgetting the quality of service offered to passengers, it was necessary to develop an appropriate measure of service. Total travel time was chosen for this purpose.

After some simplifications of the problem, an algorithm was developed to perform two specific tasks:1) to choose a set of routes to cover the area under study, and 2) to allocate bus frequencies to the routes. The main purpose was to design a bus system which gave the best total travel time for the passengers, taking into account economic restrictions.

If, Z = total travel time = $Z(f_1 , f_2 , f_3 \ldots\ldots\ldots f_N)$

where: N = number of routes, and

f_i = frequency on the ith route, in minutes.

Then, the approximate number of buses needed on route i to maintain a frequency f_i is given by:

$$e_i(f_i) = \min\left\{ \left\lceil 0.3 + \frac{R_i}{f_i} \right\rceil , \left\lceil \frac{R_i}{f_i} \right\rceil \right\}$$

where: e_i = number of buses,

R_i = round trip time for route i (including layover time),

$\lceil R_i/f_i \rceil$ = smallest integer greater than or equal to R_i/f_i.

The choice of frequencies was obtained from:

$$\min \left\{ Z(f_1, f_2, \ldots\ldots. f_N) \right\}$$

subject to $\displaystyle\sum_{i=1}^{N} e_i(f_i) \leq F$

where F is the total fleet size according to the economic restrictions.

In the process of designing the network of routes, since the bus frequencies were unknown the following properties were assumed as design criteria:

- In journeys with considerable demand, transfers should be avoided as much as possible;
- Routes should be reasonably direct;
- Routes should meet to facilitate transfers; and
- There should not be too many routes.

The input to the algorithm consists basically of a matrix of demands between nodes, a matrix of shortest distances between nodes, a list of possible termini and for each node the list of nodes directly connected to it.

The heuristic algorithm to choose a route network consists of three main steps:

i) Producing an initial skeleton route of four nodes. The selection of this initial skeleton is based on a complete combination of four-node sets in which the first and the last nodes are termini.

ii) Inserting nodes into the skeleton in order to complete the route. To consider initially if a node is acceptable to be inserted between two nodes of the skeleton route, the distance between the two nodes (of the skeleton route), via the inserted one should be less than 1.5 times the direct distance. The model also includes test to avoid the route crossing itself or back-tracking.

iii) The demands satisfied by the new route are eliminated from the demand matrix. The previous steps are repeated until this matrix contains no significant demands.

An objective function has been included in order to guarantee that the route network chosen had the properties established above. This function contains criteria to judge the skeleton routes and the process of insertion of new modes.

A basic assumption is made to calculate the total travel time for the system and is that buses arrive at stops independently of buses on other routes. In other words, the expected waiting time of a passenger at a bus stop is half the inter-arrival time if there is only one route to be taken.

The model involves the calculation of a matrix T where each element T_{ij} is the average travel time from node i to node j. As the demand matrix D is an input to the algorithm, total travel time is given by:

$$Z = \sum_i \sum_j T_{ij} D_{ij}$$

In order to calculate each element of the matrix T, it is necessary to consider three different situations:

i) There is at least one route to travel directly from node i to node j and by reasons of time it is not worthwhile to walk instead of catching a bus. The average travel time is given by:

T_{ij} = average waiting time + average bus time

It can be demonstrated by induction that if t_1, t_2,.... t_N are the inter-arrival times of the N routes that join node i and node j, the average time to the first bus is:

$$\bar{t}_N = t_1 \left[\frac{1}{2} + \sum_{r=1}^{N-1} \frac{(-1)^r t_1^r}{(r+1)(r+2)} \sum_{j_1=2}^{N-r+1} \sum_{j_2=j_1+1}^{N-r+2} \cdots \sum_{j_r=j_{r-1}+1} \frac{1}{t_{j_1} t_{j_2} \cdots t_{j_r}} \right]$$

where t_1 is the smallest inter-arrival time. If there is only one route to travel between i and j, it can be shown that $\bar{t} = 1/2\ t$, as it was expected.

If γ_1, γ_2, ... γ_N are the corresponding travel times between the two points for the N available routes, then it is necessary to estimate the proportion of passengers using each route. This is given by the expression:

$$P_i = \frac{t_1}{t_i} - \frac{1}{t_i} \sum_{r=1}^{N-1} \frac{(-t_1)^{r+1}}{r+1} \sum_{\substack{j_1=1 \\ j_1 \neq i}}^{N-r+1} \sum_{\substack{j_2=j_1+1 \\ j_2 \neq i}}^{N-r+2} \cdots \sum_{\substack{j_r=j_{r-1}+1 \\ j_r \neq i}} \frac{1}{t_{j_1} t_{j_2} \cdots t_{j_N}}$$

Then, the average bus time is given by:

$$\sum_{i=1}^{N} P_i \gamma_i$$

It is worthwhile to notice that if the bus times are assumed to be equal there is no need to estimate the proportions of passengers using each route, and the average bus time will be γ.

ii) There is at least one route to travel directly from node i to node j, but it may be sometimes better to walk. If:

w = walking time,

γ = average bus time, assumed constant in this situation,

t = time to the next arrival

Then a passenger that arrives at a stop and consults the time-table will wait if $t + \gamma \leq w$, and will walk if $t + \gamma > w$. The proportions of passengers riding and walking will be given by:

$$P_{rides} = \int_0^{w-\gamma} \phi_N (t)\, dt$$

$$P_{walks} = \int_{w-\gamma}^{t_1} \phi_N (t)\, dt$$

whre $\phi_N (t)$ is the probability density function of the time to the first arrival, and t_1 is the smallest inter-arrival time. $\phi_N (t)$ is given by:

$$\phi_N (t) = \sum_{r=1}^{N} r(-t)^{r-1} \sum_{j_1=1}^{N-r+1} \sum_{j_2=j_1+1}^{N-r+2} \cdots\cdots \sum_{j_r=j_{r-1}+1}^{N} \frac{1}{t_{j_1} t_{j_2} \cdots t_{j_N}}$$

The average travelling time is expressed by:

$$\bar{t} = \int_0^{w-\gamma} [\gamma + t]\, \phi (t)dt + \int_{w-\gamma}^{t_1} W\, \phi (t)dt$$

In the general case of N routes, the average journey time is given by:

$$T_{ij} = [\gamma - W]\, \phi_N(W - \gamma) + W + \psi'_N(W - \gamma)$$

where $\phi_N(t)$ and $\psi_N(t)$ are mathematical expressions verified by induction.

iii) No bus route joins node i to node j. At this precise moment the matrix T contains waiting and riding times where bus route is available, and walking or waiting and riding times when there is any route available but walking should be considered. In the remaining cells of the matrix there is no route to travel between node i and node j and these are filled with the corresponding walking times.

Then, a procedure developed by Murchland[1] is used in order to look at each (i,j) cell of the matrix and examine whether there is a k for which the journey times $i \rightarrow k$, $k \rightarrow J$ are less than the journey $i \rightarrow j$.

In order to consider the limited capacity of the buses, maximum utilization factors were estimated for each route. To calculate these factors the arc with maximum number of passengers is selected and this number is divided by the capacity of the route per hour. An initial solution was produced and by means of random perturbations new values of f were considered in order to improve the solution.

After solving the problem of routes and frequencies, this model also considers timetables and bus schedules.

D. TRANSEPT.[2] This model was developed for evaluating short term changes in bus route networks, but also takes into account changes in public transport demand because it considers new passengers attracted due to improved service and passengers switching to another mode of transport or to another route within the service. On the other hand, in the long term the model can evaluate changes in the road network and in the land use.

It is considered that a bus network provides benefits to several groups of people, particularly to the bus operator and to the user, and also contributes to the reduction of traffic congestion.

Two submodels are considered to be the basis of the model. One that predicts the choice of transportation mode and the other that assigns passengers to routes. These submodels are based on the concept of generalised cost. In order to evaluate a bus route network in a particular time the model comprises five main steps:

i) Demand for bus journeys is estimated. Demand for travel is considered on a zone-to-zone basis and all trips in an area are assumed to end at the zone centroid. The definition of both zone centroids and bus stops by their grid co-ordinates allows an easy calculation of the walking times between them.

The description of the bus network is based on the inventory of bus stops. Each bus stop has a reference number and its location is defined by means of grid co-ordinates. As many routes run along identical sequences of bus stops, a link is defined as a group of bus stops between which no routes start, end, join or diverge.

An important assumption made in this model is that travelling from one part of the bus network to another, all passengers will use the same path, with minimum generalised cost.

Conventional highway network models find the shortest paths between pair of nodes, meanwhile in TRANSEPT, minimum paths are found between pairs of links. In order to reduce computation the model only allows interchange between direct journeys, at link ends. A modified version of Floyd's algorithm was used.

Public transport does not generally offer a door-to-door service so that walking plays an important role as a feeder to the bus system. In this sense, generalised costs take into account the components of a bus trip: walking time, waiting time, riding time, fare, interchange, and so on. The generalised cost function has the following form (Daly, 1973):

$$Z_j = a_{oj} + \sum_{i=1}^{5} a_i x_{ij} \qquad \text{(for mode of transport j)}$$

where: x_{1j} = walking time

x_{2j} = waiting time

x_{3j} = fare

x_{4j} = riding time

x_{5j} = interchanges

a_{oj} and a_i = appropriate parameters.

The model considers two other modes of transport in addition to travel by bus: walking and travel by private car. TRANSEPT uses an n-dimensional logit model to estimate the modal split. Considering bus and private car only, the proportion of trips by bus P_b for a particular origin-destination interchange is given by:

$$P_b = \frac{e^{-\lambda Z_b}}{e^{-(\lambda Z_c + \delta)} + e^{-\lambda Z_b}}$$

where Z_b and Z_c = generalised cost of travel by bus and car

λ = calibrated parameter

δ = modal bias in favour of the car.

ii) Passengers are loaded onto buses. Passengers are assigned to available routes by means of a public transport assignment model, which takes into account the capacity of buses and hence the waiting time at bus stops. As there is a competition between routes for passengers, and between passengers for seats, the model reproduces this dynamic relationship and finds a stable solution.

iii) Overloading of buses is eliminated. When there is not enough seating capacity, the model is able to allocate additional buses in order to cope with the extra demand.

iv) The network is evaluated. The main objective of TRANSEPT is to evaluate bus route networks, and these are evaluated from four different points of view: operational, financial, user benefit and modal split.

The operational level predicts the loadings for every route and link of the network, allowing the routes to be assessed in terms of capacity. The financial level predicts revenues by route according to particular fare policies. The model also determines the user benefit in terms of generalised costs when bus networks are compared. The modal split is analysed and the economic benefits from changes can be assessed.

v) The service is improved and the recommendations of the study are implemented.

E. Partners in Management Model. The purpose of this model, which was explained by Bullock (1970), was to analyse a wide range of aspects of a bus route such as the waiting times at bus stops for different bus schedules, routing, control strategies, design of buses (capacity, fare collection, etc) and variation of bus stop patterns. This model relies basically on simulation techniques.

Passengers arrivals at bus stops are generally taken to be random according to a Poisson distribution, but in this model no time of day variations are considered, in other words passenger arrival rates remain constant during the run of the operations.

At the time of arrival of a bus at a bus stop, the elapsed time since the previous bus gives the parameter to calculate the number of passengers arriving at the stop. This value is drawn at random from a Poisson distribution.

By using an appropriate origin-destination matrix for the bus stop, each passenger is assigned to a destination stop further down the route. The O-D data are most useful expressed in matrix form, in which the individual elements are the partial probabilities.

The number of passengers boarding will depend on the places available on the bus in question. If c is the capacity of the bus and there are r_a passengers on board of which a are wishing to alight, then the places to boarding passengers will be: $(c - r_a + a)$.

The times taken to board and alight will depend on the number of passengers entering and leaving, and the total stopped time will be the sum of these two times if the bus has a single door, or the maximum of these quantities for buses with two doors.

The departure time from a bus stop is calculated as the time of arrival plus the stopped time. In order to calculate the arrival time at the next stop, the inter-stop travel time is drawn at random from the assumed time distribution. Such distribution will almost be skew, with some minimum travel time below which it is not possible to go due to vehicle characteristics.

In traffic free conditions there is a mean time which a bus should take to travel between successive stops. This mean will depend on the distance to be travelled, the type of terrain and the time of day.

The input tó this model consists mainly of the sequence of bus stops, the probability distribution of bus running times, the distribution of passengers arrivals, boarding and alighting times, and origin and destination matrix.

The output includes a record of arrivals and departures of buses from bus stops, time spent running and stopped, passenger loading in each sector of the route, queue length at stops and number of passengers unable to board because of full buses.

F. Evaluation of bus control strategies by simulation. This model, which simulates the movements of buses around a single route, was developed by Bly and Jackson (1974) in order to analyse a specific bus route in Bristol. The program was written in Control and Simulation Language (CSL).

Passengers are assumed to arrive randomly at each bus stop which has its own mean time between passenger arrivals. The model then takes this time as the mean value of a negative exponential distribution, which is used to select randomly the time intervals between successive passenger arrivals.

A matrix associated with each bus contains the number of passengers destined for each stop. The destination of each passenger is chosen randomly according to the observed distribution of destinations appropriate to each stop.

The time taken for one passenger to board a bus is also selected randomly from a boarding time distribution. The total boarding time is obtained by multiplying the selected time by the total number of passengers boarding. In a similar way, the alighting time is selected and the total stopped time of a two-door bus at a stop is given by:

$$T = C + \text{Max}\ (N_b t_b, N_a t_a)$$

where: C = dead time,

N_b, N_a = total number of passengers boarding and alighting,

t_b, t_a = boarding and alighting times

The time taken by a bus to travel from one stop to the next one on a route is obtained using a normal distribution from which the inverse of the running times are selected randomly. This calculation can include variations with time of day. In order to avoid the selection of unrealistic times, the process of choosing is repeated if the time obtained does not fall within some specific limits

When a bus arrives at a stop the model calculates the average passenger arrival rate over the period since the last passenger boarded the previous bus and this average value is used in the new process of generation. The mean passenger inter-arrival rate is constrained to vary relative to the base values according to a Gaussian curve. The selection of the running time is made in a similar way. On the other hand, a timetable is held in a matrix in order to be consulted each time a bus arrives at a timing point. Each row of this matrix holds the departure time.

As some sections of the route can be shared with other services, and a fraction of the passengers may have different alternatives or choices, the model represents these other routes by describing extra buses which are considered in less detail. No record is kept of the running of these buses, nor of the movement of their passengers.

The input to the model includes route details (stops, available buses, bus capacity, etc), running time data, passenger arrival data and variation parameters. The output consists mainly of passenger waiting times in histograms, bus occupancy information, bus headways, departures from schedules at timing points, passenger total travel time, total bus stopped time and information about the implementation of different control devices.

G. Simulation model of a bus network[3]. The main purpose of this model is to test alternative schedules and route structures in order to provide a better service to the public. The model concentrates more on the study of route networks than on individual routes, and was written in Algol.

The model assumes that passengers arrive at bus stops according to a Poisson distribution. A sample number of arrivals is obtained from this distribution and for a small interval dt, and these are summed to obtain a sample number of arrivals over any particular period.

The variation of passenger arrival patterns throughout a period is considered by defining the arrival rates at any period as a function of five parameters (see Figure 10) (Jackson, 1972):

i) mean arrival rate through the period: λ

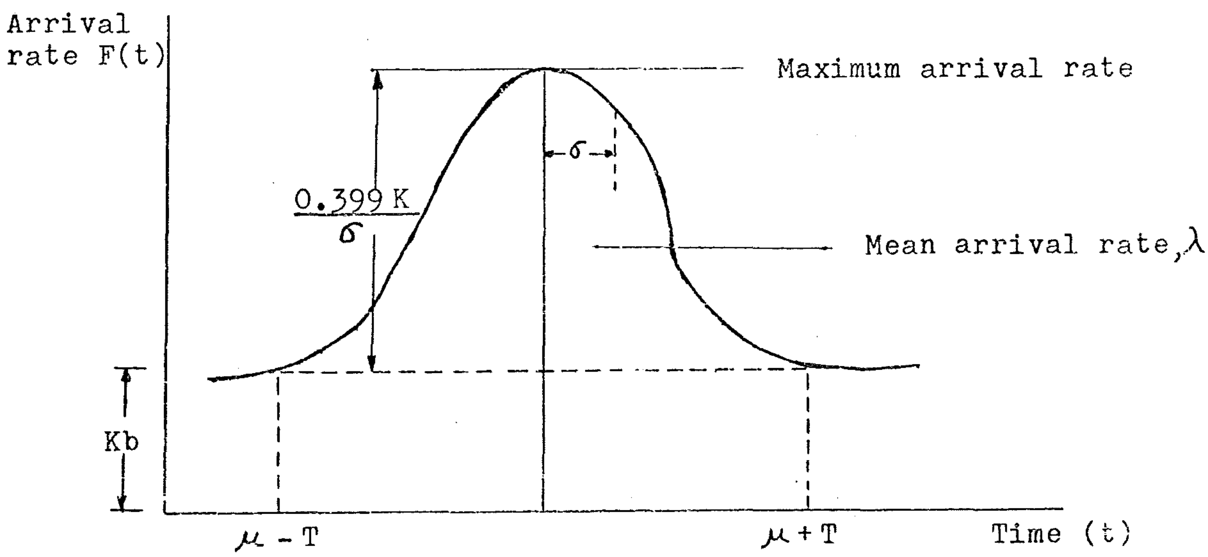
ii) measure of spread of any peaking effect: σ

iii) time of the peak: μ

iv) length of the period: 2T

v) maximum arrival rate/mean arrival rate: γ

FIGURE 10. Variation of passenger arrival rate.



The function F(t) is defined by the addition of a curve of a normal distribution and a constant in the following way:

$$F(t) = \text{arrival rate as a function of time} = K\left[f(t) + b\right]$$

where K is a scaling factor and b is to be defined. The normal distribution is given by:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}$$

If the area under f(t) is the unity, the area under F(t) is given by: K (1 + 2Tb) and this area is equal to $\lambda \times 2T$. Then,

$$\lambda = \frac{K}{2T}\,(1 + 2Tb)$$

$$\gamma = \frac{\text{maximum observed arrival rate}}{\lambda} = \frac{K\left(\frac{0.399}{\sigma} + b\right)}{K\left(\frac{1}{2T} + b\right)}$$

and then $$b = \frac{\left(\frac{0.399}{\sigma} - \frac{\gamma}{2T}\right)}{\gamma - 1}$$

and having b, F(t) is easily calculated. Then, the arrival rate at any time is approximated by KF(t).

The model also assumes that the origin and destination stops of passengers are independent of the service provided. Only direct trips are considered here; in other words, the origin-destination information is obtained for pair of stops which have a direct bus service. This means that transfers will be registered only if passengers transferring appear twice in the data and these movements are not easily identifiable.

Where it is possible to travel between two pair of stops using different bus services, a passenger will be assigned to the first vehicle with place available that arrives at the stop.

As soon as passengers arrive at bus stops they are classified into types according to their destination. If the bus arriving at the stop cannot accommodate all the passengers desiring to board, passengers to be carried will be selected according to the classification mentioned before and to a probability proportional to the time since the last arrival of a passenger who was able to board a bus for each destination.

Journey times between nodes are determined by sampling from distributions whose means depend on the individual link, time of day, sampled journey time of previous bus and on elapsed time since that bus.

It is assumed that there is some relationships between the running times of two consecutive buses travelling along the same section of the route. This relationship tries to express that when there are two buses running closely or with small headway it is likely that they will experience similar conditions with respect to traffic, weather, etc, and therefore their travelling times will be probably similar. In order to take into account this possible correlation, the travel time over the section for the second bus will be:

$$\text{travel time} = R(t_2) + e(t_2)$$

whre: $R(t_2)$ = mean running time at time t_2, given by a curve,

$e(t_2) = qe(t_1) + (1-q)S(t_2)$, where

$e(t)$ = estimated deviation from the mean running time for a bus at time t.

$S(t)$ = sample drawn from a normal distribution with mean zero and standard deviation $\sigma(t)=\min\left\{a(t)/3, b(t)/3\right\}$

$a(t) = M - R(t)$

$b(t) = R(t) - m$

M, m = maximum and minimum observed running times.

The values of q are chosen as follows:

$$q = 1 - h/2R(t_1) \qquad \text{if } h < R(t_1),$$

or

$$q = R(t_1)/2h \qquad \text{if } h \geq R(t_1)$$

where $h = t_2 - t_1$ = headway.

The model considers vehicle overtaking, and in order to avoid this the values of S(t) are restricted to be greater than a specific value.

The input to the model consists of the route network, the passenger travel patterns which is specified by means of an origin-destination matrix, the running times between stops and the passenger arrival rates. Time of day variations can be considered using histograms or analytic expressions.

A wide variety of output information can be printed: global values of average passenger waiting time, distributions of waiting times for passengers on each route, record of bus departures from timing points, etc.

H. <u>IBM simulation model</u>. This model, which is a discrete stochastic simulation model, was developed by Gerrard and Brook (1972) for investigation of scheduling and timetabling strategies for a group of routes in an urban corridor. For flexibility of input and output the model was written in FORTRAN.

Passenger arrival at bus stops is assumed to be Poisson, with arrival rate dependent on the time of the day. The arrival rate $\alpha(t)$ for a stop varies over the peak hour period according to some profile. If the interval (t_1, t_2) represents the arrival times of two consecutive buses, then the expected number of arrivals during that period is given by:

$$A = \int_{t_1}^{t_2} \alpha(t)\,dt$$

Then, the total number of passengers is generated using a Poisson distribution with mean A, or which is the same with a constant arrival rate α_e over the same period:

$$\alpha_e = \frac{A}{(t_2 - t_1)}$$

The inter-arrival times of passengers are generated and accumulated until the time exceeds (t_2-t_1) and then the number of passengers in the new queue segment will be equal to the number of inter-arrival times generated.

After generating the total number of passengers, they have to be classified by passenger type according to the route choice available at the stop in question. If there are two routes available, e.g. 1 and 2, one type of passengers will use route 1, another will use route 2, and passengers in the third group will catch either 1 or 2. Each bus stop is assumed to have a characteristic value of λ and μ, loading and unloading indices respectively, for a given pattern of movement.

In order to generate the passenger type it is necessary to calculate the probability that a passenger arriving at bus stop ith will be of the kth type. This probability is given by:

$$P_{ik} = \frac{\sum_{j \in G_{ik}} \mu_j}{\sum_{j \in G_i} \mu_j}$$

where: G_i = set of all destination stops available from stop i,

G_{ik} = subset of G_i available to passengers of type k

Having selected the P_{ik}, the cumulative distribution is formed in order to generate the type for each passenger.

A queue is considered to be formed by one or more segments and each segment consists of a number of passengers classified in exclusive types following the method explained above. Each segment of a queue is processed in sequence, but the arrival time of each passenger is not recorded within each segment. When a bus arrives at a stop and there are not enough places for people waiting to board, the model selects the passenger according to the order of segments.

The number of passengers alighting at a stop is generated using a binomial distribution. The probability that any given passenger on board a bus of route r will wish to alight at stop j is given by:

$$q_{jr} = \frac{\mu_j}{\mu_j + \sum_{w} \mu_l}$$

where w = remaining stops on schedule.
Then, the probability of m passengers wishing to alight at stop j is given by:

$$p(m) = \frac{m\,!}{M\,!\,(M-m)\,!}\,(q_{jr})^m\,(1-q_{jr})^{M-m}$$

where M = number of passengers on board at time of arrival.

The process of unloading shows that the model assumes the independence of passenger origin and destination. In this sense, the data collection was considerably reduced because only passenger loading and unloading is required.

The pattern of passenger movement between bus stops is expressed as an O-D matrix and the number of passengers travelling from stop i to stop j is assumed to be given by $\lambda_i \mu_j$, where:

λ_i = loading index at origin stop, and

μ_j = unloading index for destination stop.

In order to calculate the values of μ_j and λ_i, the following two sets of equations are used:

$$R_i = \lambda_i \sum_{i+1}^{n} \mu_j \qquad i = 1,2,\ldots.n$$

$$C_j = \mu_j \sum_{1}^{j-1} \lambda i \qquad j = 1,2,\ldots.n$$

where R_i and C_j are the number of passengers loading and unloading respectively at stop i and j for the period under consideration. The previous system of equations can be solved by fixing the value of μ_n.

One of the main problems in modelling the bus movements between bus stops is the way to represent the traffic effects on bus journey times. For each link of the bus network a mathematical expression is assumed to calculate the main daily trend of journey time.

As it is expected that buses making the same journey at similar points in time would have similar journey times, a model is assumed to represent this correlation:

$$Y = DX + \sqrt{(1 - D^2)}\ \varepsilon$$

whre: Y = the excess of the travel time of a bus over the mean travel time,

X = excess time, with respect to the mean, for the previous bus to travel along the same link,

D = e^{-at}, with values between 0 and 1,

a = constant

t = headway or time interval between successive buses,

ε = independent variable: excess time, with respect to the mean, for the bus in question.

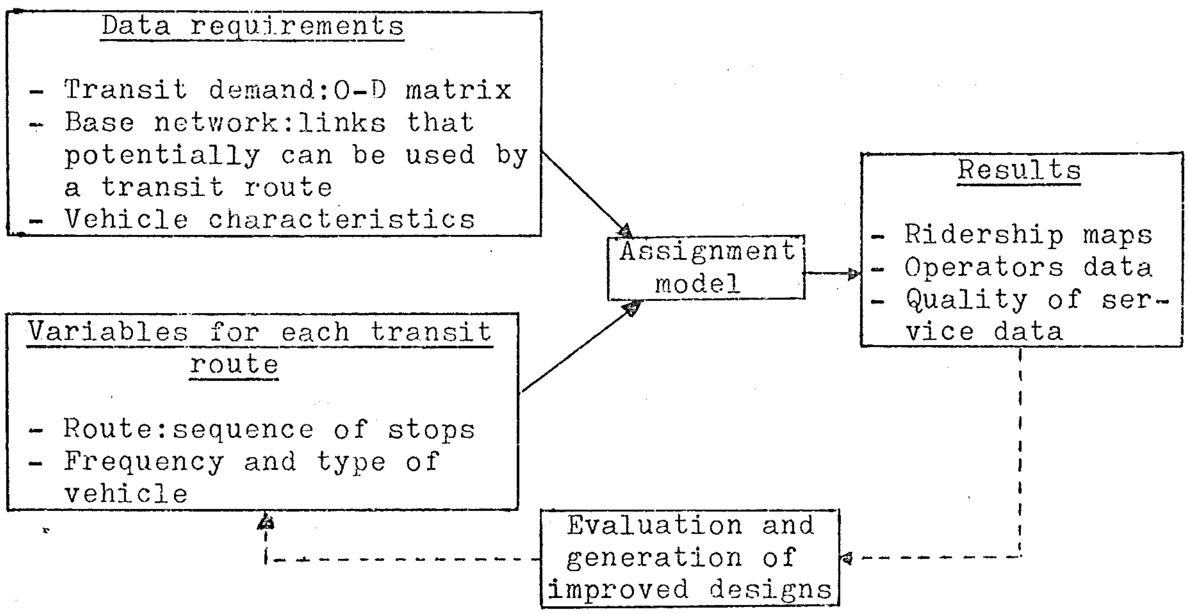
The input to the model includes passenger demand, traffic conditions, route structure, type of vehicles, frequencies of service and criteria for assessing service performance. The output consists of information about levels of service (passenger waiting times), headways at stops, departure from schedule at stops, journey times and total number of passengers loading and unloading at stops.

I. Interactive graphics system for transit route optimisation. A model developed by Rapp and Robert-Grandpierre (1974) optimises the route structure of an urban transit system. The model assigns the potential transit trips to the network in order to predict the effects on the routing structure by means of statistical measures on travel times, rolling stock use and operating costs.

This model is a tool for transit operators to improve the matching of demand and supply through network and route optimisation, increasing the level of service and decreasing the operating costs. It is considered that transport planners who are familiar with the problem can find almost an optimal solution if they have a tool that allows them to generate, evaluate and improve alternative designs. The model was designed mainly for finding short-term improvements in the transport system: bus, street, car and subway routes (see Figure 11).

The assignment algorithm is based on the hypothesis that travellers probably choose the path with the least resistance or total travel times, but that all other nonbacktracking paths have a certain likelihood of being used.

FIGURE 11. General diagram of interactive graphics model.



Dial's [4] multipath assignment model was designed for highway trip assignment and therefore requires a homogeneous network. For public transport two difficulties arise:

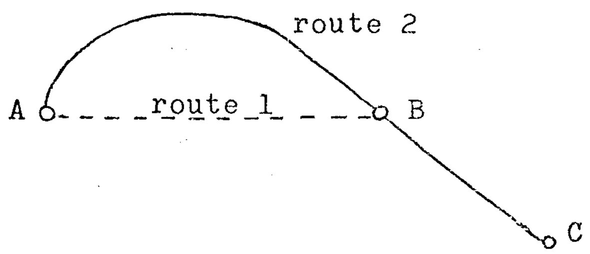
i) total path resistance = walking time + waiting time + riding time + transfer time (from origin to final destination.)

ii) the principle of optimality, which is the base of Dial's model is not straight forwardly applicable in public transport. According to this principle if minimum path between A and B uses route 1 (see Figure 12), the minimum path between A and C must be A-B-C using route 1. But in the real world, the path A-B-C- using route 2 could be better because it is not necessary to make a transfer at B.

By applying some stated rules that add dummy links to the transit network, this is transformed into a homogeneous highway network, and then it is possible to apply the Dial's multipath traffic assignment model.

One limitation of this model is that transit trip assign-

FIGURE 12. Three-bus-stop public transport network.



ment ignores any effects of limited capacity on passenger path selection. However, the model has been applied with success in the public transport system of Lausanne and in the tramway and bus network of Basel. The results are displayed adequately on a computer screen.

J. Other models. The models presented above are by no means the only ones to have been developed in the public transport area. Some other authors who have treated similar aspects of the problem in a different way are worth mentioning.

OMNIBUS (Reeves, 1970) is a suite of computer programs to assist in public transport systems. This model also works in terms of generalised costs to find shortest paths through the network and then assigns traffic flows to the links of the network. A public transport network is assumed to consist of nodes, centroids and links. Passengers leave and join the system at centroids and transfer to and from other routes at nodes. The model determines paths from centroid to centroid and to do so all possible paths between each pair of centroids are considered in order to form the optimum set.

Pasley and Sudgen (1972) apply analytical models to the rationalisation of public transport services in Teesside. This study includes a program for flows assignment and carries out the system evaluation at two levels: cost benefit analysis and financial evaluation. Holm (1973) developed a model to allocate buses to a route network in an optimum way. His model assigns passengers to the public transport system and also includes a modal choice process. Lillienberg (1973) describes a model which has been applied in practical planning in several cities in Sweden and Denmark. Erlander and Scheéle[5] describe an analytical model for optimum bus frequencies.

Clearly, many authors have studied the public transport planning process using a wide variety of methodologies. So, it is worth emphasising that the topic is not new and that the model presented in the next chapter does not deal with untreated aspects of the subject. However, as will be shown, the proposed model tackles the problem in a different way and the use of several linked submodels may be regarded as a real contribution to the analysis.

NOTES TO CHAPTER I

1) Murchland ,J.D. (1965). A new method of finding all elementary paths in a complete directed graph. L.S.E., T.N.T., 22, London School of Economics Report.

2) This summary was extracted from several papers: Local Government Operational Research Unit (1973), Daly (1973), Last and Daly (1975) and Last and Leak (1976).

3) The summary of this model was extracted from several papers and from a PhD dissertation. See Jackson H.S.(1972), Jackson H. S. (1973) and Jackson H.S. and Wren (1973).

4) Dial R.B., A multipath traffic assignment model. Highway Research Record N^{o}369, 1971, 199-210

5) Erlander, S. and S. Scheéle (1974) A mathematical programming model for bus traffic in a network. In Transportation and traffic theory (ed. Buckley), A.H.& A.W. Reed, Sidney, 581-605.

CHAPTER II

PROPOSED PUBLIC TRANSPORT ASSIGNMENT MODEL

The purpose of this chapter is to present in detail the different parts of the proposed model which has been developed for planning and evaluation of public transport operations in urban areas. As explained in the Introduction above it is assumed that the travel demand has been established so that the problem may be reduced basically to the assignment of this demand to the existing transport facilities. The model employs a modular approach and consists of several linked submodels which deal with the different aspects of the assignment process. Many of the considerations which were involved in the design of the model are related to the theory which was discussed in the previous chapter and therefore no special emphasis has been put on them here.

The public transport assignment problem can be divided into three different stages: i) the determination of individual options or ways to make a trip[1], in such a way that they constitute reasonable alternatives for travellers; ii) process of loading passengers onto selected options; and iii) process of loading passengers onto routes and vehicles. The first stage constitutes the basis of the whole process, although the assignment itself is done in the second and third steps.

When a person wants to travel from A to B using a public transport system, he may have one of two alternatives: either i) the nearest bus stops to A and B are connected by a direct route service; or ii) the two bus stops are not connected so that the person must make at least one transfer.

Of course this does not mean that a passenger making this particular trip has to use the nearest bus stops to A and B; he could, if he has a good knowledge of the bus network, have a different set of options using for instance the walking mode as a feeder to the bus system. For example, if he does not have a direct service from A to B, he could walk to a further bus stop in order to get a direct service to B.

In any case, the main problem is how to identify individual options or travel paths between any two points of a network. Although there are still some deficiencies in public transport assignment knowledge, there is little doubt that people tend to travel in such a way as to minimise their total travel costs. But the problem is how to express the concept of minimum costs when, besides the distance to be travelled, it is also necessary to consider walking time, waiting time at bus stops, fare and penalties for transferring. In other words, travel by public transport involves the user in at least two different types of expenditure: money and time. As it was explained in the previous chapter, total expenditure is usually referred to as generalised cost, C.

Taking into account these considerations, the first part of this chapter deals with the development of an algorithm to find alternatives for travelling between any two points of a network.

Having identified the different ways of making a trip, the second part of the model is concerned with the manner in which people choose their option. If transport is in general considered to be an undesirable necessity, then it is possible to assume that travellers select their option so as to minimise their perceived cost. This would mean that the more expensive the option the less is the probability of its being chosen. The second part of this chapter will deal with the problem of loading trips onto selected travelling options taking into account these considerations.

When there is only one route available to make a journey, a user may be assumed to catch the first bus that comes, subject to seat availability. If there is more than one service on the route, then once again the user will take the first bus that comes regardless of the service. The loading of passengers onto routes and buses is the subject of the third part of this chapter.

1. Option selection process.[2]

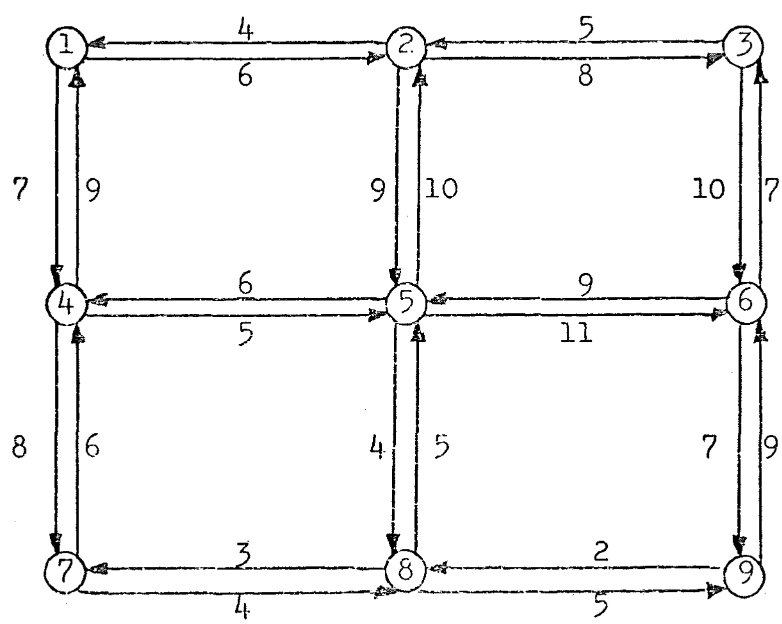
This part of the chapter deals with the development of a model which identifies the different alternatives for making a trip given some characteristics of the public transport network. It is divided into two sections. The first is concerned with the main approaches to finding minimum paths in networks, while the second treats the same problem but applied specifically to public transport networks.

The first section includes a matrix model, called PATH, that calculates shortest paths for simple networks while registering the sequence of nodes to be followed. The second section describes a method in matrix notation called ROUTE-1 developed from PATH, this time to deal with public transport networks. Furthermore, ROUTE-1 serves as basis to design a more complex model, ROUTE-2, which considers the walking mode as a feeder to the transport system and registers alternative options for making a trip besides the optimum one.

A. Algorithms to find minimum paths in networks. The models which are described here may be classified into two different groups: The first group comprises some of the main non-matrix algorithms to find shortest paths in networks, while the second one is composed of the matrix notation models. The Moore algorithm plays a leading role in the first group, whereas the traditional matrix approach which is the basis of the proposed model is considered in the second group.

The following simplified network has been used extensively throughout this work to check many programing details of the various submodels. It is introduced here to illustrate how the different minimum paths algorithms work and will be referred to again later, so it is important for the reader to become familiar with its characteristics. Figure 13 shows a network in which the figures on the links represent travel time in minutes between pairs of nodes.

FIGURE 13. Sample network.



a. Non-matrix notation methods. Given a network of N nodes, and if minimum paths are to be found from home node H to all other nodes, then the idea is to build a tree of (N-1) directed links in such a way that the paths described by the branches constitute the shortest ones.

E.F. Moore (1957)[3] who was working in the field of telecommunications developed a model that received his name and was capable of finding shortest paths in networks. The method consists initially in going from H to every connecting or adjacent node in order to record the respective travel time, then the closest node to H is selected and the corresponding link is placed in the tree. Thereafter, all connected nodes to nodes already reached are involved, and the process is repeated until all nodes have been selected.

If the Moore algorithm is to be applied to obtain minimum paths from home node 1, then nodes 2 and 4 are initially checked and the first one being closest to home node is selected. Then node 2 has two connecting nodes, 3 and 5 , and this time node 4 is selected with a distance of 7 to home node. This process is repeated until all nodes are reached. The minimum tree from node 1 is shown in Figure 14.

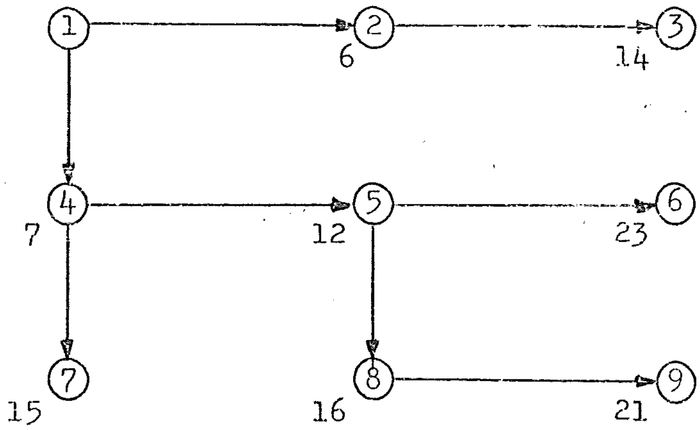
In mathematical terms the problem consists in finding a label $[I, p(J)]$ for each node J of a link (I,J), where I is a node already labelled and p(J) is the shortest distance from home node to J. Initially home node is labelled $[-, 0]$ and the next node to be labelled is found by determining the J for which:

$$p(J) = \min_{I \in X} \left[p(I) + \min_{J \in \bar{X}} L(I,J) \right]$$

where: X = set of labelled nodes in the network, and

$\bar{X}$ = set of unlabelled nodes in the network.

FIGURE 14. Minimum tree from node 1 .



The process finishes when all the nodes in the network have their corresponding label (Blunden, 1971). For the network under study the initial conditions and the first two iterations are shown below:

i) $X = \{1\}$

$\bar{X} = \{2,3,4,5,6,7,8,9\}$

$$p(J) = \min_{I=1} \left\{ p(I) + \min_{J=2,3,4,5,6,7,8,9} L(I,J) \right\}$$

$$= \min \left\{ p(1) + \min \; L(1,2), L(1,4) \right\}$$

$$= \min \left\{ 0 + \min \; [6,7] \right\} = 6 \qquad \text{for } J = 2$$

Therefore, node (2) is labelled [1,6] and the process is repeated after updating sets X and $\bar{X}$

ii) $X = \{1,2\}$ $\qquad \bar{X} = \{3,4,5,6,7,8,9\}$

$$p(J) = \min_{I=1,2} \left\{ p(I) + \min_{J=3,4,5,6,7,8,9} L(I,J) \right\}$$

$$= \min \begin{Bmatrix} p(1) + \min \; L(1,4) \\ p(2) + \min \; L(2,3), L(2,5) \end{Bmatrix}$$

$$= \min \begin{Bmatrix} 0 + 7 \\ 6 + \min\,[8,9] \end{Bmatrix}$$

$$= \min \begin{Bmatrix} 7 \\ 14 \end{Bmatrix} = 7 \qquad \text{for } J = 4$$

and then node (4) is labelled [1,7]. By working through the process it is possible to complete the labels of the rest of the nodes and it is easy to see that the results coincide with those given in Figure 14.

The Moore algorithm was used by R. Dial (1967) to develop a transit pathfinder algorithm which will be discussed in the second section of this chapter. The Moore algorithm was broken down by Dial into three steps:

i) The elements of the Cumulative-time Table from home node to node J, CUM(J) are set arbitrarily high except for home-node cell which is set to zero. The Link-sequencing Table, TABLE(CT,K) contains the J-node of link whose cumulative time is CT. TREE(J) contains the I-node of the final link in the shortest path between home node and node J. The first link to be entered into TABLE (0,1) is link 1-1 which is a dummy link.

ii) After a link has been selected, all other links connected are examined in order to determine their possible entry to the Link-sequencing Table TABLE(CT,K). The cumulative time CT, of each new link under test is found by adding its travel time to the current cumulative time. If the calculated CT is less than CUM(J) the link is placed in TABLE(CT,K) while in TREE(J) is placed the corresponding node.

iii) CT is incremented until a new J-node is found in TABLE(CT,K), then the corresponding link is selected and step ii) is repeated. This process goes on until N-1 links have been selected, where N is the number of nodes in the network. Then the answer to the problem is found in TREE(J), and the respective cumulative times are given in CUM(J). For the example under study results are shown in Tables 10 and 11.

TABLE 10. Cumulative times via minimum path.

J	CUM(J)	TREE(J)
1	0	home node
2	~~∞~~, 6	1
3	~~∞~~, 14	2
4	~~∞~~, 7	1
5	~~∞~~, ~~15~~, 12	~~2~~, 4
6	~~∞~~, 23	5
7	~~∞~~, 15	4
8	~~∞~~, 16	5
9	~~∞~~, 21	8

Source: Calculations made by the author.

TABLE 11. Link-sequencing Table.

CT	TABLE(CT,K)	CT	TABLE(CT,K)
0	~~1~~	13	
1		14	~~3~~
2		15	~~5~~, ~~7~~
3		16	
4		17	
5		18	
6	~~2~~	19	
7	~~4~~	20	
8		21	~~9~~
9		22	
10		23	~~6~~
11		24	
12	~~5~~	25	

Source: Calculations made by the author.

As defined by J.A.George and H.G. Daellenbach (1978), dynamic programing is a computational method that breaks up a complex problem into a sequence of easier subproblems by means of a recursive relation which can be evaluated by stages. According to R.P. Potts (1978) this technique can be applied to find the shortest route between two nodes and this formulation is convenient as a basis for the model validation.

The minimum path between nodes (1) and (2) could be found by applying dynamic programing (Hollingdale, 1978) in the following way:

Let $g_n(1\text{-}9)$ denote the length of the shortest route from (1) to (9) passing through (N-1) links. To determine $g_4(1\text{-}9)$ the principle applies as follows:

$$g_4(1\text{-}9) = \min \begin{Bmatrix} g_1(1\text{-}2) + g_3(2\text{-}9) \\ g_1(1\text{-}4) + g_3(4\text{-}9) \end{Bmatrix}$$

where $g_3(2\text{-}9)$ and $g_3(4\text{-}9)$ are defined in the same way:

$$g_3(2\text{-}9) = \min \begin{Bmatrix} g_1(2\text{-}3) + g_2(3\text{-}9) \\ g_1(2\text{-}5) + g_2(5\text{-}9) \end{Bmatrix}$$

$$g_3(4\text{-}9) = \min \begin{Bmatrix} g_1(4\text{-}5) + g_2(5\text{-}9) \\ g_1(4\text{-}7) + g_2(7\text{-}9) \end{Bmatrix}$$

$$g_2(3\text{-}9) = 17$$

$$g_2(5\text{-}9) = \min \begin{Bmatrix} g_1(5\text{-}6) + g_1(6\text{-}9) \\ g_1(5\text{-}8) + g_1(8\text{-}9) \end{Bmatrix} = \min \begin{Bmatrix} 11 + 7 \\ 4 + 5 \end{Bmatrix} = 9$$

(via (8))

$$g_2(7\text{-}9) = 9$$

Consequently, having calculated the two-link paths it is straight to determine the three-link distance and hence the four-link distance which is the desired one.

$$g_3(2\text{-}9) = \min \begin{Bmatrix} 8 + 17 \\ 9 + 9 \end{Bmatrix} = 18 \qquad (\text{ via } 5)$$

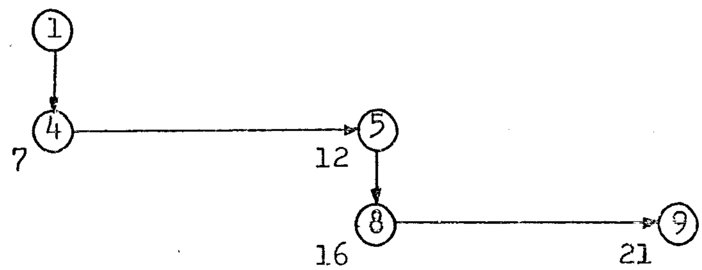
$$g_3(4\text{-}9) = \min \begin{Bmatrix} 5 + 9 \\ 8 + 9 \end{Bmatrix} = 14 \qquad (\text{ via } 5)$$

and then,

$$g_4(1\text{-}9) = \min \begin{Bmatrix} 6 + 18 \\ 7 + 14 \end{Bmatrix} = 21 \qquad (\text{ via } 4)$$

Therefore, the shortest path from 1 to 9 (see Figure 15) is equal to the one obtained by the Moore algorithm.

FIGURE 15. Shortest path from 1 to 9.



Although the examples given above are easily solved, that does not mean that the precise verification of these algorithms is straight-forward. However, such an exercise is beyond the scope of this study.

b. Matrix notation methods. Any network may be represented as a matrix in which each cell entry is used to record a relationship between a pair of nodes and this relationship could be, for instance, the travel time between nodes. Therefore, a matrix can take information in a concise way and in an acceptable and convenient form for mathematical and computing manipulation. The respective matrix for the matrix given in Figure 13 is shown in Table 12 below. This matrix is called L_1 , Structure Matrix, and each element or cell $L_1(I,J)$ represents travel time between I and J. This matrix will be used to explain the traditional matrix method to find minimum paths.

TABLE 12. Matrix L_1.

	1	2	3	4	5	6	7	8	9
1	0	6	∞	7	∞	∞	∞	∞	∞
2	4	0	8	∞	9	∞	∞	∞	∞
3	∞	5	0	∞	∞	10	∞	∞	∞
4	9	∞	∞	0	5	∞	8	∞	∞
5	∞	10	∞	6	0	11	∞	4	∞
6	∞	∞	7	∞	9	0	∞	∞	7
7	∞	∞	∞	6	∞	∞	0	4	∞
8	∞	∞	∞	∞	5	∞	3	0	5
9	∞	∞	∞	∞	∞	9	∞	2	∞

Source: Based on Figure 13.

As may be seen in Figure 13, from node 1 to node 3 there is no link, in other words these nodes are not connected and therefore the corresponding cell in Table 12 is set to infinity, ∞ . However, it is possible to find a path to go from 1 to 3 making an indirect connection. For instance, to go first from 1 to 2 and then from 2 to 3 , connection that makes a total travel time of 6 + 8 = 14. In order to inspect all possible connections between those nodes using at the beginning a maximum of two links, row 1 of the Structure Matrix is faced with column 3 of the same matrix, as it is shown in Table 13.

Therefore, making a connection which includes two links, the best path to go from 1 to 3 is via 2 . Thus, by using matrix L_1 which contains the minimum direct travel time between pair of nodes, it is possible to find a new matrix L_2 with shortest distances, taking into account both direct connections and indirect paths of up to two links. Matrix L_2 is given in Table 14. In mathematical terms, each entry cell of the new matrix L_m is given by the following expression:

$$(6) \qquad \underset{\substack{I=1,N \\ J=1,N}}{L_m(I,J)} = \underset{K=1,N}{\min} \left[L_{m-1}(I,K) + L_{m-1}(K,J) \right]$$

where N = number of nodes in the network.

TABLE 13. Two-link connections 1 → 3

1 → 1	0 + ∞	1 → 3	= ∞
2	6 + 8	2	= 14 — minimum
3	∞ + 0	3	= ∞
4	7 + ∞	4	= ∞
5	∞ + ∞	5	= ∞
6	∞ + 7	6	= ∞
7	∞ + ∞	7	= ∞
8	∞ + ∞	8	= ∞
9	∞ + ∞	9	= ∞

Source: Calculations made by the author.

It is worth noticing, for instance, that a path to go from 1 to 6 has to have at least three links; that is the reason why $L_2(1,6)$ still remains equal to ∞ . By following the same procedure, matrix L_3 can be found (see Table 15). This matrix contains the minimum travel times between pair of nodes, taking into account both direct connections and connections with up to four links.

It could be shown that when $L_{m+1} = L_m$, then matrix L_m contains the shortest path between any pair of nodes (Taaffe and Gauthier, 1973). In this example, it could be demonstrated that $L_3 = L_4$. Therefore, bearing in mind that this method does

TABLE 14. Matrix L_2.

	1	2	3	4	5	6	7	8	9
1	0	6	14	7	12	∞	15	∞	∞
2	4	0	8	15	9	18	∞	13	∞
3	9	5	0	∞	14	10	∞	∞	17
4	9	15	∞	0	5	16	8	9	∞
5	14	10	18	6	0	11	7	4	9
6	∞	12	7	15	9	0	∞	9	7
7	15	∞	∞	6	9	∞	0	4	9
8	∞	15	∞	9	5	14	3	0	5
9	∞	∞	16	∞	7	9	5	2	0

Source: Calculations made by the author.

TABLE 15. Matrix L_3.

	1	2	3	4	5	6	7	8	9
1	0	6	14	7	12	23	15	16	21
2	4	0	8	11	9	18	16	13	18
3	9	5	0	16	14	10	21	18	17
4	9	15	23	0	5	16	8	9	14
5	14	10	18	6	0	11	7	4	9
6	16	12	7	15	9	0	12	9	7
7	15	19	25	6	9	18	0	4	9
8	18	15	21	9	5	14	3	0	5
9	20	17	16	11	7	9	5	2	0

Source: Calculations made by the author.

not record the paths to be followed between nodes, this matrix approach proves to be much simpler and easier to program than the methods explained in the first section.

The previous method requires two matrices of N x N, where

N is the number of nodes in the network. This approach has a serious disadvantage in that it does not register the corresponding sequence of nodes that constitutes a path. According to some authors (i.e. MacDougall, 1976) to do so would be a matter of a trial-and-error procedure which could require a considerable amount of computer storage. Another deficiency is that the method could be too slow. Therefore, in order to overcome, at least partially, the above disadvantages an algorithm named PATH is proposed. This model constitutes a simplified version of the procedure explained earlier.

In contrast with the traditional matrix approach, PATH requires only one array of N x N, is capable of obtaining minimum paths in a simpler way, while at the same time registering them in a convenient and practical manner for retrieval. In this case the N x N matrix records distance in time and is called Structure Matrix D. For the network given in Figure 13, matrix D is equal to matrix L_1 (see Table 12). Given D, the basic expression to find shortest paths is as follows:

$$(7) \qquad \underset{\substack{I=1,N \\ J=1,N \\ I \neq J}}{D(I,J)} = \min \left[D(I,J), \left\{ \min_{\substack{K=1,N \\ K \neq J \text{ and } K \neq I}} D(I,K) + D(K,J) \right\} \right]$$

N = number of nodes in the network

D = Structure Matrix of dimension N x N

Given I and J, the expression $\left\{ D(I,K) + D(K,J) \right\}$ is evaluated varying K from 1, 1 by 1, until N and the minimum value selected is compared to the actual value of D(I,J). If this figure is greater than the one obtained then the cell value is updated. It is precisely here where the difference between expressions (6) and (7) lies, as in the first method the corresponding cell is not updated. Instead, the respective value is kept in a different matrix.

If matrix D is updated going row by row, then it can be seen that D(1,1), D(1,2) and D(1,4) remain unchanged, while D(1,3) = 14 and D(1,5) = 12 . Up to now the two methods coincide in their results, but when a new value for D(1,6) is calculated the difference between them begins to appear:

$$D(1,6) = \min \left\{ D(1,6),\ \min_{\substack{K=1,9 \\ K \neq 6 \text{ and } K \neq 1}} \left(D(1,K) + D(K,6) \right) \right\}$$

$$D(1,6) = \min \left\{ \infty,\ \min\ (6+\infty, 14+10, 7+\infty, 12+11, \infty+\infty, \infty+\infty, \infty+9) \right\}$$

$$D(1,6) = \min \left\{ \infty, 23 \right\} = 23$$

It is clear that in the first method the value $L_2(1,6)$ remains at ∞ because during its calculation values of $L_1(1,3)$ and $L_1(1,5)$ are still not updated. In the new approach cells are being continuously updated and the solution tends to converge more quickly.

Following this method, a new matrix D is obtained and it can be seen that this matrix is closer than L_2 to the final solution, which in this case is obtained in the next iteration. The new matrix D is given in Table 16, while the optimum matrix is, of course, equal to L_3 (Table 15). It can be seen that convergence to an optimum solution is faster with the new approach and in some cases its use could lead to savings in complete iterations. With an example studied by Ackoff (1968) of a network of 11 nodes and solved with the traditional method it was necessary to do two iterations before reaching the final solution, while it can be proved that by using PATH only one iteration is necessary.

TABLE 16. Matrix D before final solution (first iteration).

	1	2	3	4	5	6	7	8	9
1	0	6	14	7	12	23	15	16	21
2	4	0	8	11	9	18	19	13	18
3	9	5	0	16	14	10	24	18	17
4	9	15	23	0	5	16	8	9	14
5	14	10	18	6	0	11	7	4	9
6	16	12	7	15	9	0	16	9	7
7	15	21	29	6	9	20	0	4	9
8	18	15	23	9	5	14	3	0	5
9	20	17	16	11	7	9	5	2	0

Source : Calculations made by the author.

For a network with N nodes, the shortest path from I to J can contain (N-1) links at most. With the traditional approach L_m contains the shortest paths with 2 links or fewer; L_3 with 4 links or fewer; L_4 with 8 links or fewer, and so on. This means that 2^{m-1} gives the maximum number of links for the mth iteration. Therefore, the shortest paths will be found at the mth iteration when $L_m = L_{m+1}$ or at most when $2^{m-1} \geqslant N-1$ for the first time.

For the network under study, the solution will be found at most when $2^{m-1} \geqslant N-1$, then $2^{m-1} \geqslant 8$, which means m = 4th iteration at most; and in fact it was found for m = 3. Thus, for a network with 1000 nodes, the solution will be found at most when $2^{m-1} \geqslant 999$, then m = 11th iteration at most.

With PATH, the previous concepts do not apply in the same way. Looking at Table 16 D(1,9) = 21, which happens to be the optimum path from 1 to 9 via 4 5 and 8. This means that this matrix got an optimum path with 4 links. However, D(2,7) = 19, while the optimum path is 16 with only 3 links and has not been obtained yet. This is because the optimum path from 2 → 7 is via 5, and the shortest path from 5 → 7 has not yet been obtained when calculating the whole path 2 → 7. Therefore, this depends very much on the way matrix D is updated. PATH updates row by row and from left to right.

Although there is no way to predict the number of iterations to find the final solution, it is clear that PATH converges faster than the traditional method and therefore must save computer time.

An important characteristic can be drawn from the analysis of simple networks: if a minimum path from I to J passes through K, then the first leg of that path will be the shortest path from I to K. For example, the minimum path from 1 to 9 passes through 8 and that means that the shortest path from 1 to 8 constitutes the first part of path 1 → 9.

On this basis it is possible to develop a procedure that registers paths of a network by recording the last node visited before reaching the final destination. A matrix called MT(N,N) is set to zero at the beginning of the process and each cell of this matrix is updated according to the following rule, for a minimum path from I to J via K:

$$\begin{aligned} MT(I,J) &= K & \text{if } MT(K,J) &= 0 \\ MT(I,J) &= MT(K,J) & \text{if } MT(K,J) &\neq 0 \end{aligned}$$

For the final solution of the network under study, matrix MT is given in Table 17. It is worth noting that this matrix is initialized only at the beginning of the process and therefore every cell is updated according to the rule given above.

In order to reconstruct the minimum path from I to J the following rule applies: take value MT(I,J), make K = MT(I,J) and get successive values of MT(I,K) until K = 0. The series of K values (except 0) constitute the nodes visited in the way from I to J. For example, from (1) to (9): MT(1,9) = 8, MT(1,8)= 5, MT(1,5) = 4, MT(1,4) = 0. Therefore the respective path is (1) → (4) → (5) → (8) → (9), with a total travel time of 21.

TABLE 17. Matrix MT for final solution.

	1	2	3	4	5	6	7	8	9
1	0	0	2	0	4	5	4	5	8
2	0	0	0	1	0	3	8	5	8
3	2	0	0	1	2	0	8	5	6
4	0	1	2	0	0	5	0	5	8
5	2	0	2	0	0	0	8	0	8
6	2	3	0	5	0	0	8	9	0
7	4	5	6	0	8	9	0	0	8
8	4	5	6	7	0	9	0	0	0
9	4	5	6	7	8	0	8	0	0

Source: Calculations made by the author.

The PATH algorithm requires for its programing two matrices D(N,N) and MT(N,N) for N-node networks, where MT registers the last node visited before reaching final destination. In this model the following variable is defined:

$$\text{TODIS} = \sum_{\substack{I=1,N \\ J=1,N}} D(I,J)$$

When no cell has been updated during an iteration, then

TODIS will remain unchanged and the final solution will have been found. PATH was written in FORTRAN and Flowchart 1 shows the set of operations necessary for the programing of this model. On the other hand, Flowchart 2 shows the procedure to retrieve minimum paths after the final solution has been reached. In order to do so an array MPATH(N) is required, where N is the number of nodes in the network. PATH would be worth applying when it is necessary to find minimum paths between all pair of nodes of a network as this algorithm works with the network as a whole. The whole development and results of the network under consideration are given in Appendix A.

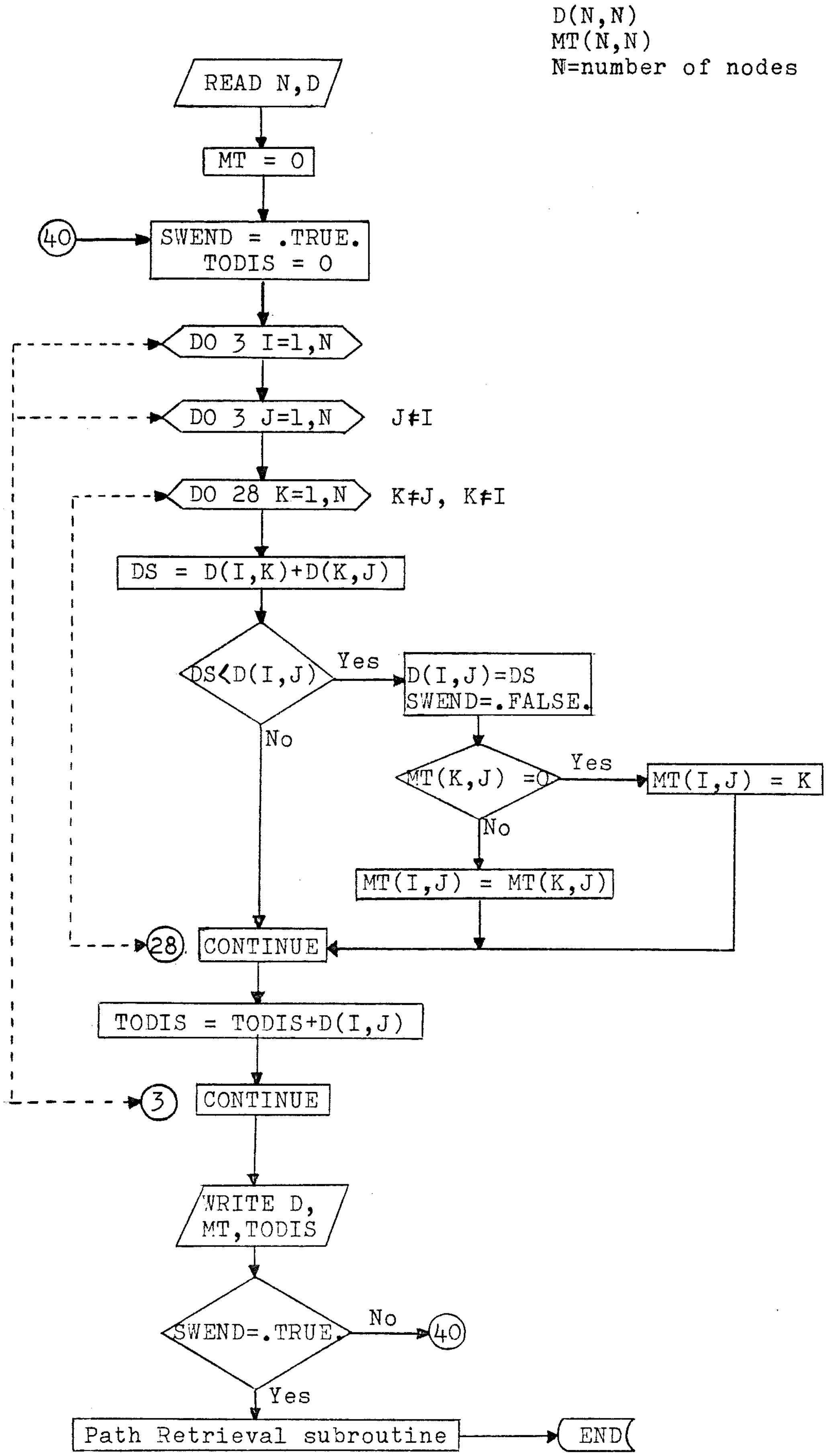
B.<u>Algorithms to find minimum paths in a public transport network</u>.

In public transport networks it is necessary to include waiting times at bus stops in addition to the link travelling times. This contrasts with simple networks and is because public transport networks comprise a system of fixed routes that follow a specific path mainly at regular intervals of time. Furthermore, the analysis of a bus network should include walking time, fares and penalties for transferring.

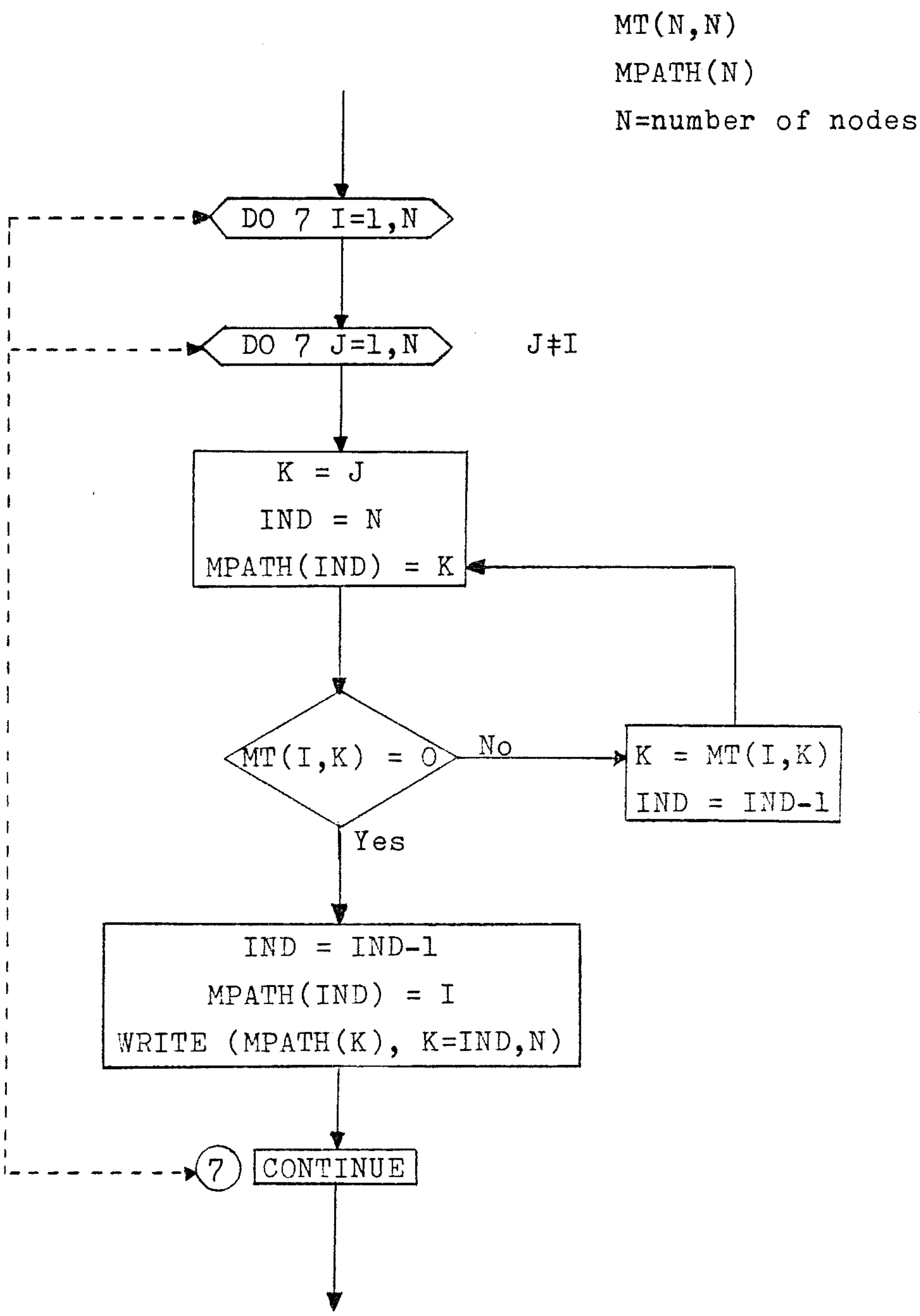
The average waiting time at bus stops (AWT) may be expressed as a function of the frequency of service. Assuming that passengers arrive randomly at bus stops, that they can get on the first bus that comes and that the service runs regularly, then the AWT can be taken as half the frequency. When there is more than one route traversing a common link, then AWT is given by the following expression:

$$(8) \qquad AWT = \frac{0.5}{\sum_{b \in S} 1/f_b}$$

FLOWCHART 1. PATH



FLOWCHART 2. PATH retrieval subroutine



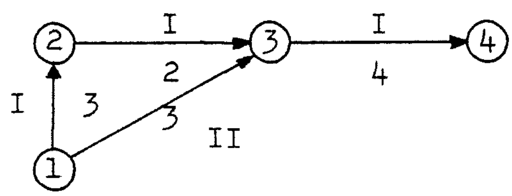
where: f_b is the frequency of service b, in time units

S is the set of routes available in a common link.

In simple networks if a minimum path from I to J passes through K, then the first part of that path is the minimum path from I to K. However, this important characteristic does not usually apply to a transit network as is shown by the following example (Figure 16). Roman numbers indicate routes while the figures on the links represent travel time in minutes. The frequencies of service are as follows:

Route	Bus/hour	Frequency (min)
I	20	3
II	3	20
III	6	10

FIGURE 16. Four-bus-stop network.

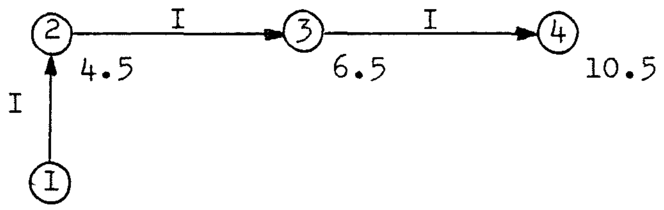


Given this information, total travel time between 1 and 3 , and 1 and 4 is computed in the following way (waiting times in brackets):

1 → 4	via 2	(3/2) + 3 + 2 + 4 = 10.5 ← shortest path
1 → 4	via 3	(20/2) + 3 + (3/2) + 4 = 18.5
1 → 3	via 2	(3/2) + 3 + 2 = 6.5 ← shortest path
1 → 3	direct	(20/2) + 3 = 13

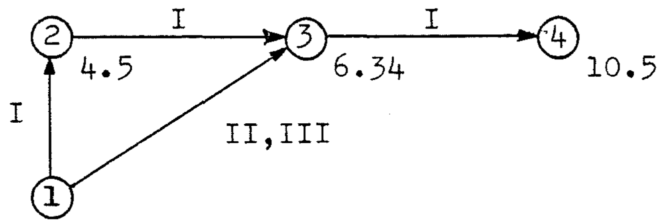
Thus, in this case, the minimum path from 1 to 4 passes through 3 via 2, using the minimum path 1 → 3 (see Figure 17).

FIGURE 17 . Minimum paths from node (1).



But if a new route III is provided to travel directly from (1) to (3) , with a frequency service of 10 minutes, there will be 9 buses per hour (3 from route II), which means an average service interval of 6.67 minutes, and then AWT will be 3.34 minutes. Therefore, it is necessary to recalculate the total travel time as follows (see Figure 18):

FIGURE 18. Minimum paths from 1 after including new route.



(1) → (4) via (2) (3/2) + 3 + 2 + 4 = 10.5 ← shortest path

(1) → (4) via (3) (3.34) + 3 + (3/2) + 4 = 11.84

(1) → (3) via (2) (3/2) + 3 + 2 = 6.5

(1) → (3) direct (3.34) + 3 = 6.34 ← shortest path

Now, the minimum path from (1) to (4) passes through (3) but does not use the minimum path from (1) to (3) . Therefore, the important feature on which path retrieval of simple networks is based cannot be applied to public transport networks. However, it could be seen that if a trip from A to B requires a transfer at C, then the minimum path from A to C constitutes the first part of that trip (Dial, 1967).

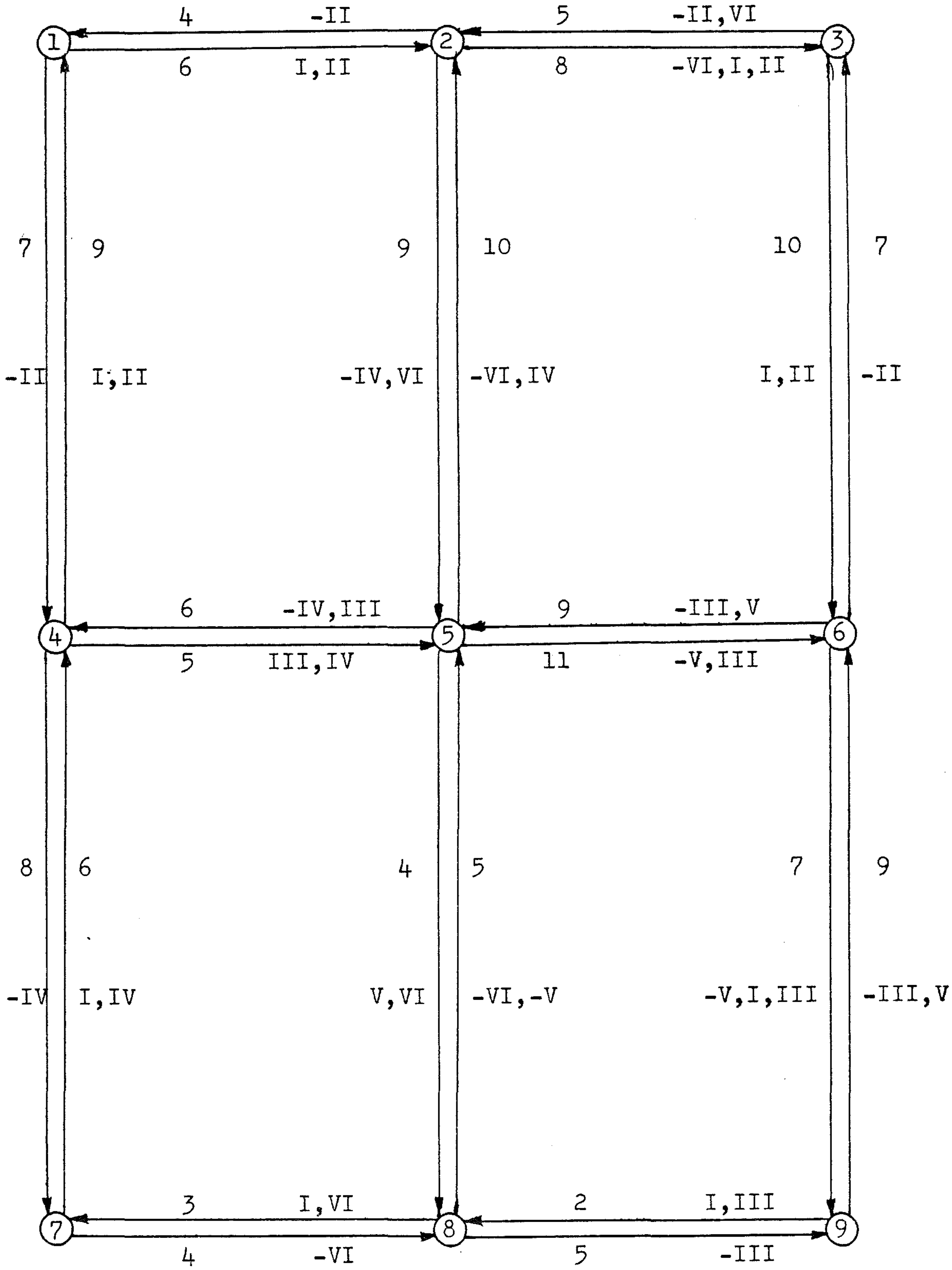
Bearing in mind the previous considerations, this section will concentrate basically on two matrix-notation models to obtain minimum paths for a public transport network: ROUTE-1 and ROUTE-2. The former was developed from PATH, a model already explained in the first section of this chapter; the latter is a further improvement of ROUTE-1. This part also includes two non-matrix algorithms, one developed by R. Dial (1967) and the other made as part of the London Transportation Study described by Lane, et al (1971).

For the purpose of exemplifying the different algorithms the bus network shown in Figure 19 is proposed. The network has nine bus stops and six routes. Figures on the links denote riding time in minutes while roman numbers indicate bus services. Frequencies of service are also given in minutes.Route IV for instance, runs in this way $7 \rightarrow 4 \rightarrow 5 \rightarrow 2$, while route -IV runs following the same path but in the other direction, both routes having a frequency of 15 minutes.

a. Non-matrix algorithms for minimum paths in public transport networks. R. Dial (1967) made an extension of the Moore algorithm in order to consider the special characteristics of minimum paths in a bus network. The manner Dial treated that particular algorithm to develop a computer program has already been considered in the first section.

Two main assumptions were considered by Dial in his path-finder algorithm: one is that link travelling times remain constant during the process, and the second is that waiting time at boarding points can be approximated by half the frequency of

FIGURE 19. Bus network example.



Route	I	II	III	IV	V	VI
Frequency(min)	10	12	8	15	10	8

of service.

This algorithm requires a network description composed of route frequencies and trunk-line links. Route frequencies are given by the average number of buses per unit of time for each route in the system. Each route has the same frequency for both directions; therefore, if the out-bound frequency is different to the in-bound one, then the two directions must be coded as separate routes. Route frequencies are used in the model to assess waiting times X(S), given by the following expression:

$$X(S) = \frac{0.5}{\sum_{K \in S} F(K)}$$

where F(K) is the frequency of route K (number of buses per unit of time) and S is the set of routes serving the link.

A trunk-line link is said to have three different components:

i) a pair of nodes A and B that denotes one-way flow between them;

ii) the travelling time between A and B; and

iii) a set of routes that traverse the link A-B in the same time.

This implies that when there is more than one level of service between two nodes (i.e. bus and underground) it is necessary to code dummy links in order to assign the routes providing the other levels of service.

As in the Moore algorithm, the minimum path time of the home node is set to zero and then this node is the first to be extracted. After a link is extracted, all links exiting from it are examined in order to see if they are going to be put in the Link-sequencing Table. Assuming that E is the extracted link and G the generated one, then the total travel time will be equal to the cumulative time plus G's link time plus a transfer penalty factor. This factor will be equal to X(G) when there are no common routes between links E and G, in other words when there is a transfer involved; otherwise, this factor will be equal to X(G) - X(E).

The column of extracted links constitutes the tree with minimum paths from home node 1. The first extracted link is 1 - 1, a dummy link with a total travel time to home node equal to zero. For the network under study the generated links are 1 - 2 and 1 - 4 and their time to home node is recorded in the Working Table (see Table 18). Then link 1 - 2 will be extracted and the whole process repeated until N links have been extracted, where N is the number of nodes in the network. Final results are given in Table 19 which is theworksheet for the example under question. Minimum paths from home node 1 are shown in Figure 20 below.

TABLE 18. Working Table.

Time slot	L.S.T.	T.S.	L.S.T.	T.S.	L.S.T.	T.S.	L.S.T.
0	1→1	9		18		27	4→7
1		10		19		28	2→7
2		11		20	2→5	29	
3		12		21	4→5	30	
4		13	1→4	22		31	
5		14		23		32	
6		15		24		33	5→6
7		16	1→3	25	2→8	34	8→9
8	1→2	17		26	1→6	35	

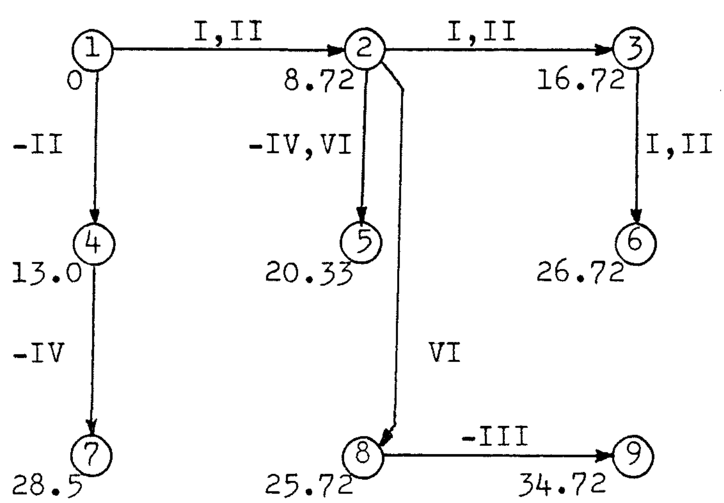
Source: Calculations made by the author.

TABLE 19. Worksheet for transit-tree-building example.

Step	Extracted link	Generated link	Time calculation
0	-	1-1(0)	0
1	1-1(0)	1-2(I,II) 1-4(-II)	0+6+(2.72)=8.72 0+7+(6)=13
2	1-2(I,II)(8.72)	1-3(I,II) 2-5(-IV,VI)	8.72+8=16.72 8.72+9+(2.61)=20.33
3	1-4(-II)(13.0)	4-5(III,IV) 4-7(-IV)	13.0+5+(2.61)=20.61 13.0+8+(7.5)=28.5
4	1-3(I,II)(16.72)	1-6(I,II)	16.72+10=26.72
5	2-5(-IV,VI)(20.33)	5-6(-V,III) 2-8(VI)	20.33+11+(2.22)=33.55 20.33+4-(2.61)+(4)=25.72
6	2-8(VI)(25.72)	2-7(VI) 8-9(-III)	25.72+3=28.72 25.72+5+(4)=34.72
7	1-6(I,II)(26.72)	1-9(I)	26.72+7-(2.72)+5=36.0
8	4-7(-IV)(28.5)	7-8(-VI)	28.5+4+(4)=36.5 > 25.72
9	8-9(-III)(34.72)	None	-

Source: Calculations made by the author.

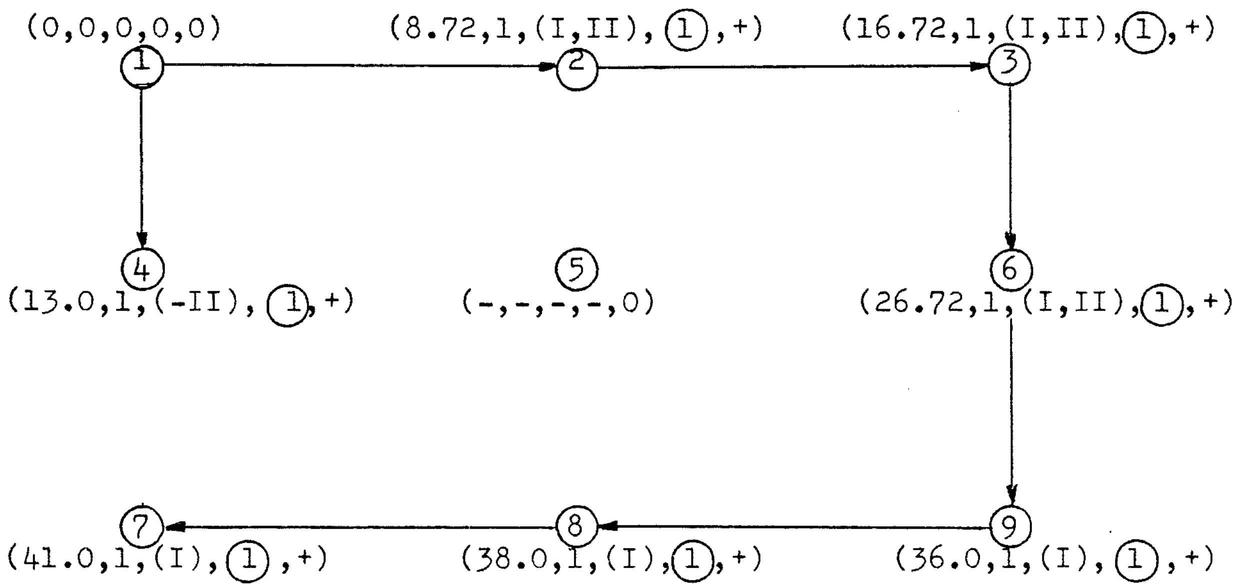
FIGURE 20. Minimum paths from 1 .



The London Transportation Study Model, which also obtains minimum paths in a transit network, has been described by Lane et al (1971). In this model every node of the network has some information that constitutes its label: 1) time taken from home node to the given node; 2) number of boardings to reach the given node; 3) bus route or routes on which arriving; 4) node at which bus was boarded; and 5) an indicator, either '0' or '+'.

At the beginning of the process, home node has a '+' indicator, while all other nodes have a '0' indicator. Immediatly after all services through home node are considered and the nodes served by these routes change their indicators from '0' to '+'. Then the label of each node affected is updated and the indicator of home node is set to '0'. For the network under study, Figure 21 shows the node labels after one boarding has been considered.

FIGURE 21. One boarding.



Later on the bus services passing through a node with a '+' indicator are examined in the same way as it was done for home node. Nodes not reached before are updated and their indicators changed to '+' sign. On the other hand, in case of nodes reached before, travel time is calculated by the new route in order to make the appropriate changes. The process is repeated until no further changes are required (see Figures 22 and 23)

FIGURE 22. Two boardings.

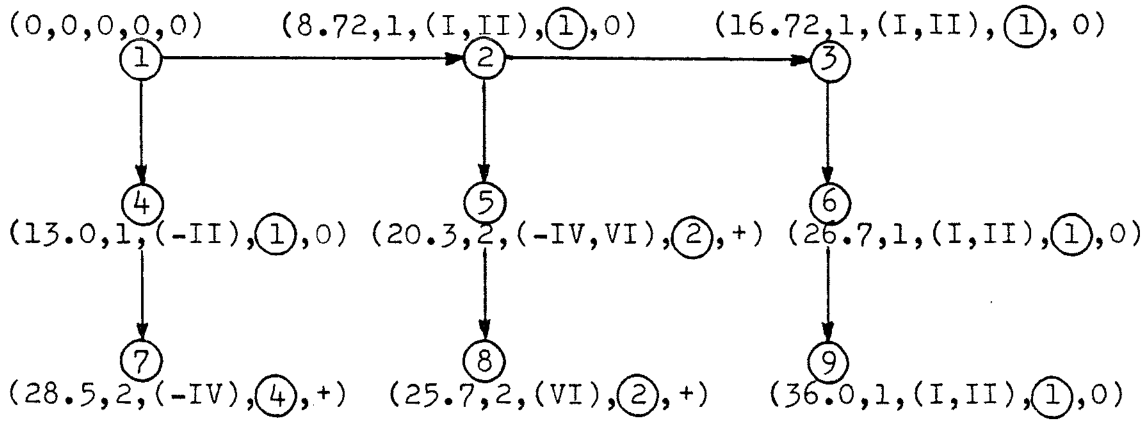
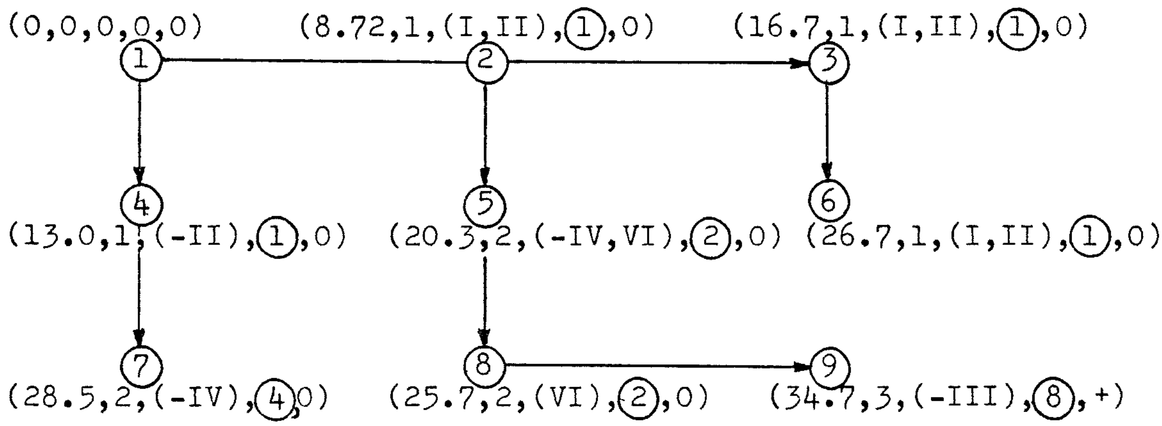


FIGURE 23. Three boardings.



b. A matrix approach to deal with public transport networks. A matrix-notation algorithm called ROUTE-1 has been developed to obtain in a computer run the minimum paths between every pair of nodes in a public transport network. PATH, algorithm which deals with simple networks, has been the basis for the design of this model. In order to illustrate the proposed algorithm the same network described in Figure 19 will be considered. This bus network has only 9 bus stops and 6 routes. Although a real one is bigger and more complicated, the main idea is to develop the mathematical approach and the computing procedure which will allow the study of real world networks.

As with the models explained before, ROUTE-1 also uses travel time as the measure of travel resistance. Here total travel time comprises riding time plus waiting time at boarding points. The basic data input to the model are given below: three parameters and three matrices are needed.

i) Three parameters:

- NUSTO, which gives the number of bus stops in the network. For the example NUSTO = 9.
- NUROU, which gives the total number of bus routes in the system. A bus route is defined as a string of bus stops numbered uniquely with two termini at both ends of the line. In a circular route one bus stop constitutes both termini. A bus route could have two different directions, and in this algorithm the direction is determined by a sign, either '+' or '-' which precedes the bus route number. On the other side, a link with no bus service is initially allocated a route called (NUROU + 1), a code which implies that there is no route available. Here NUROU = 6, therefore route VII implies no service.
- NUROC, which gives the maximum number of bus routes which share the same link in the network. In this example NUROC = 3.

ii) Three matrices:

- a matrix of distances D(NUSTO,NUSTO) in which each element

D(I,J) is the distance in time between I and J. When there is no direct link between I and J the corresponding cell D(I,J) is set to infinity or to a higher number (in this case 999). Matrix D is equal to matrix L_1 given in Table 12.

- a matrix of routes MR(NUSTO,NUSTO,NUROC) in which each element MR(I,J,N) is a bus route which joins I and J, where N =1,..... NUROC. For example, MR(1,2,1) = I; MR(1,2,2) = II, means that there are two routes between stops 1 and 2. Although matrix MR has three subscripts, for simplicity it will be used with only two to make reference to the whole set of routes. Then, MR(1,2) = I,II (see Table 20).

- a matrix of frequencies FR(NUROU) in which each element FR(M) indicates the bus frequency in minutes for bus route M (see Table 21).

TABLE 20. Matrix MR (initial values).

	1	2	3	4	5	6	7	8	9
1	-	I,II	VII	-II	VII	VII	VII	VII	VII
2	-II	-	-VI,I,II	VII	-IV,VI	VII	VII	VII	VII
3	VII	-II,VI	-	VII	VII	I,II	VII	VII	VII
4	I,II	VII	VII	-	III,IV	VII	-IV	VII	VII
5	VII	-VI,IV	VII	-IV,III	-	-V,III	VII	V,VI	VII
6	VII	VII	-II	VII	-III,V	-	VII	VII	-V,I,III
7	VII	VII	VII	I,IV	VII	VII	-	-VI	VII
8	VII	VII	VII	VII	-VI,-V	VII	I,VI	-	-III
9	VII	VII	VII	VII	VII	-III,V	VII	I,III	-

Source: Based on Figure 19.

Given matrix D, it can be transformed into a public transport network by adding the respective waiting times at the beginning of the journey. For instance, from 1 to 2 riding time is

6 minutes and this pair of nodes is served by routes I and II with bus services of 10 and 12 minutes frequency respectively. By means of expression (8) (page 118), the average waiting time is calculated as follows:

$$\text{AWT} = \frac{0.5}{1/\text{FR(I)} + 1/\text{FR(II)}} = \frac{0.5}{1/10 + 1/12} = 2.73 \text{ minutes}$$

TABLE 21. Bus frequencies, matrix FR.

Route	I	II	III	IV	V	VI
Frequency(min)	10	12	8	15	10	8

The function (or subroutine) which calculates the AWT (given the bus frequencies in minutes) is called SUM(argument), where the argument is the list of route numbers. This means that SUM(MR(1,2)) = SUM(I,II) = 2.73 minutes. Therefore, matrix D (see Table 22) is transformed by means of the following expression:

$$\begin{aligned} D(I,J) &= D(I,J) + \text{SUM}(MR(I,J)) \\ &\quad I=1,\text{NUSTO} \\ &\quad J=1,\text{NUSTO} \end{aligned}$$

$$\text{for } D(I,J) \neq 0 \quad \text{and} \quad D(I,J) \neq 999 \text{ or } \infty$$

As in PATH this method also requires a matrix MT(NUSTO, NUSTO) which will record the last transfer made from one bus service to another. This matrix is initially set to zeros and the way to update it will be explained later on. ROUTE-1 also requires two vectors or working table IR(NUROC) and MPATH(N), where the maximum value of N is given by the expression:

NUSTO + JOURNEYS * NUROC

where JOURNEYS is the highest number of services boarded for all the paths obtained.

TABLE 22. Matrix D transformed into a public transport network.

	1	2	3	4	5	6	7	8	9
1	0	8.7	999	13.0	999	999	999	999	999
2	10	0	9.6	999	11.6	999	999	999	999
3	999	7.4	0	999	999	12.7	999	999	999
4	11.7	999	999	0	7.6	999	15.5	999	999
5	999	12.6	999	8.6	0	13.2	999	6.2	999
6	999	999	13.0	999	11.2	0	999	999	8.5
7	999	999	999	9.0	999	999	0	8.0	999
8	999	999	999	999	7.2	999	5.2	0	9.0
9	999	999	999	999	999	11.2	999	4.2	0

Source: Calculations made by the author.

The fundamental expression of the PATH model is given on page 113 (expression (7)), while the rule to update matrix MT appears on page 116. However, since a public transport network differs from simple networks, the mathematical expression of the ROUTE-1 model will not be so simple as is shown below. On the other hand, it was stated before that if the minimum path from I to J requires a transfer at K, then this path would use the minimum path from I to K. This very important characteristic of public transport networks allows the model to register minimum paths by keeping the last transfer made in matrix MT.

PATH is a model that inspects paths from I to J via node K. ROUTE-1 works with the same principle in mind but taking into account the set of fixed routes comprised by the public transport system. But before giving the fundamental expression of ROUTE-1, some important aspects should be explained and this may be done by means of two hypothetical examples:

i) $\text{(I)} \xrightarrow{\text{II,IV}} \text{(K)} \qquad \text{(K)} \xrightarrow{\text{III}} \text{(J)}$

The first example shows the simplest situation in which a passenger can travel from I to K without transfer and can go from K to J also directly, but if travelling from I to J via K he should make a transfer at K since there is no common route between the two sections.

In order to determine common routes between section I-K and section K-J, a new subroutine called INTER is introduced. Given MR(I,K) = II,IV and MR(K,J) = III, this subroutine will produce IR(N), N=1,....NUROC, where each element of IR is a common route between the two sections. In other words, subroutine INTER obtains the intersection between two sets. In this example IR = Φ (empty set) which means that there are no common routes and therefore this option implies a transfer at bus stop K.

As was explained before, matrix D is transformed into a public transport matrix by adding waiting times at respective bus stops. For this example travel time (riding plus waiting time) may be given by the following expression:

$$D(I,J) = D(I,K) + D(K,J) \qquad \text{for } \text{(I)} \rightarrow \text{(J)} \text{ via } \text{(K)}$$

Furthermore, if MT is to register the last transfer of the journey, then clearly MT(I,J) = K. Matrix MR should record the last route or routes to be taken, so MR(I,J) = MR(K,J) = III.

ii) $\text{(I)} \xrightarrow{\text{II,V}} \text{(K)} \qquad \text{(K)} \xrightarrow{\text{II,IV}} \text{(M)} \xrightarrow{\text{III}} \text{(J)}$

where: MR(I,K) = II,V ; MR(K,J) = III ; MR(K,M) = II,IV

MT(I,K) = 0 ; MT(K,J) = M

In this case a passenger going from I to J via K could take route II at I, go to M where he should make a transfer in order to get route III. This means that there is a common route, II, between the two sections. But if subroutine INTER defined above is to be used to get common routes between the two sections, results will not be as predicted because $MR(I,K) \cap MR(K,J) = \Phi$. In fact what INTER does is to get common routes between MR(I,K) and MR(K,JT), where JT is the next bus stop where there is a transfer after leaving K. If there is no transfer in the second section, as in the previous example, JT = J; but here JT = M, so INTER gets common elements between MR(I,K) and MR(K,M), then IR = II

Therefore, what INTER does is to get the intersection between two sets, so in mathematical terms:

$$IR(N) = MR(I,K,N) \cap MR(J,JT,N) \qquad N = 1,NUROC$$

Total travel time from I to J via K in this case cannot be calculated in the same way as was done in the first example because D(I,K) includes the waiting time for routes II and IV, while in this journey only route II is involved. On the other hand, D(K,J) also includes route V which should not be considered. This odd situation could be overcome if travel time is calculated using the following expression:

$$D(I,J) = (D(I,K) - SUM(MR(I,K))) + (D(K,J) - SUM(MR(K,JT))) + SUM(IR)$$

If MT is to register the last transfer made and MR to register the last route or routes taken, then clearly MT(I,J) = MT(K,J) = M; and MR(I,J) = MR(K,J) = III.

Therefore,in mathematical terms, ROUTE-1 is given by the following expressions and rules:

$$(9) \quad \underset{\substack{I \neq J \\ I=1,NUSTO \\ J=1,NUSTO}}{D(I,J) = \min} \left[D(I,J), \underset{\substack{K=1,NUSTO \\ K \neq J \text{ and } K \neq I \\ D(I,K) \neq \infty \text{ and } D(K,J) \neq \infty}}{\min} \left(D(I,K) + D(K,J) + T_K \right) \right]$$

where: D is matrix of distances already transformed into a public transport matrix

$T_K = 0$ if $IR = \Phi$;

$T_K = SUM\left[IR\right] - SUM\left[MR(I,K)\right] - SUM\left[MR(K,JT)\right]$ otherwise

$SUM\left[X\right] = \sum_{b \in X} \frac{0.5}{1/FR(b)}$

FR is matrix of frequencies

MR is matrix of routes

$IR(N) = MR(I,K,N) \cap MR(K,JT,N) \qquad N=1,NUROC$

JT bus stop where next transfer is made after leaving K in the way to J. If there is no transfer then JT = J.

Rules to update MR:

$MR(I,J) = MR(K,J)$ if $IR = \Phi$ or $JT \neq J$

$MR(I,J) = IR$ otherwise

Rules to update MT:

$MT(I,J) = MT(K,J)$ if $MT(K,J) \neq 0$

$MT(I,J) = K$ if $IR = \Phi$

$MT(I,J) = MT(I,K)$ otherwise

In the updating of MT the first rule which is satisfied is the one which applies. It is worth remembering that MR always registers the last route(s) to be taken; and MT registers the last stop where a transfer ia made. Therefore, the rule mentioned above only serve to guarantee these two properties.

When matrix of distances D has no changes between two iterations is because the optimum paths have been reached; where an iteration is defined as the whole updating of D, MR and

MT. The flow diagram of ROUTE-1 is shown in Flowchart 3.

Coming back to the example under study, it can be seen, for instance, that nodes (1) and (3) are not connected which means that D(1,3) = 999 and MR(1,3) = VII. Then,a path to go from (1) to (3) would require more than one link; for example a path constituted by links (1)→(2) and (2)→(3). In this way, according to the basic data, total travel time is calculated in the following way, taking into account that a passenger doing the journey would take either route I or route II.

$$(1) \xrightarrow[6^*]{I,II} (2) \qquad (2) \xrightarrow[8^*]{-VI,I,II} (3)$$

* These distances do not include waiting time.

Total travel time = riding time (1)→(2) + riding time (2)→(3) + waiting time at (1)

TTT = 6 + 8 + SUM(I,II) = 14+ 2.7 = 16.7 minutes

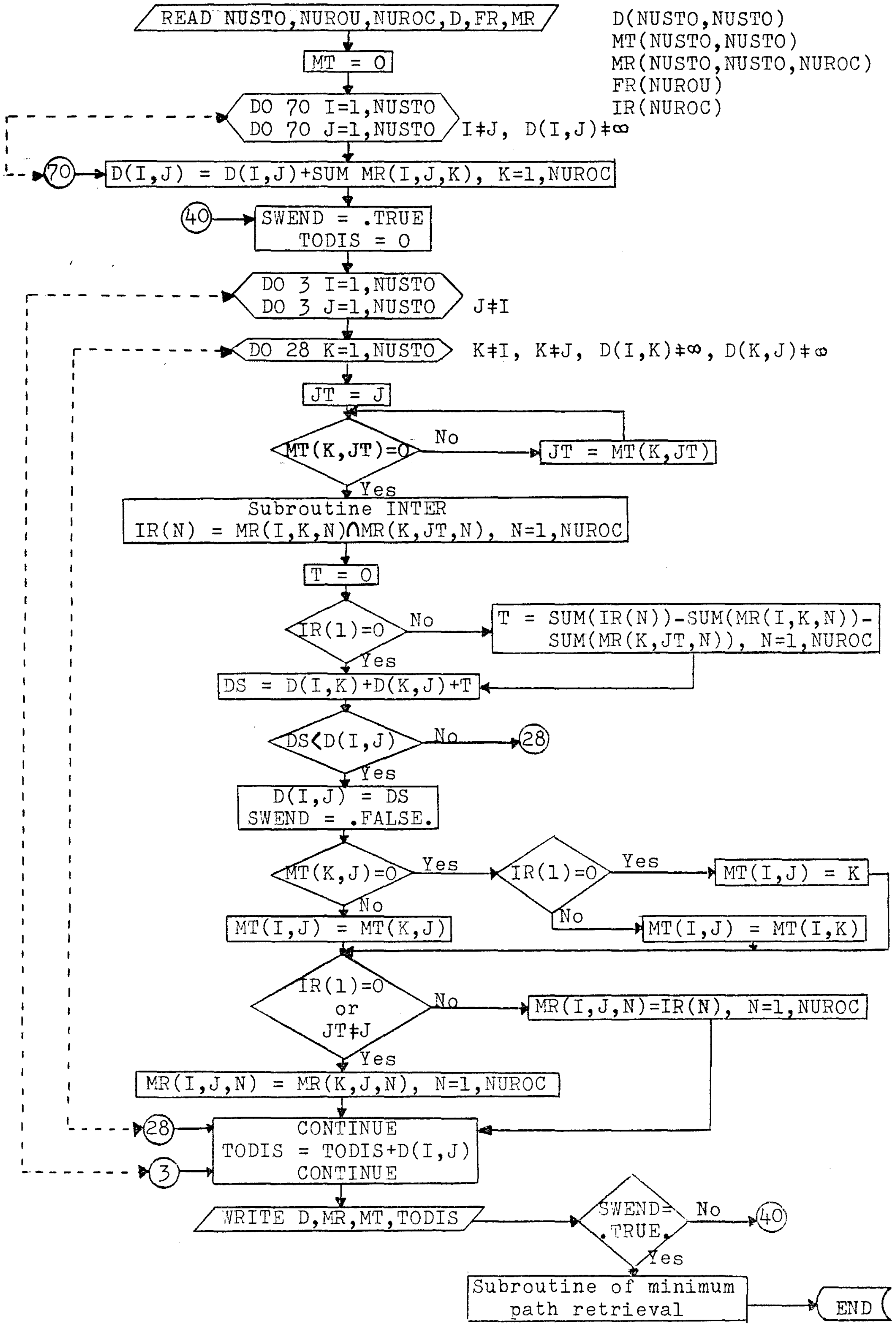
In order to inspect all the possible paths from (1) to (3) using up to two links, row 1 of matrix D (Table 22) is faced with column 3 of the same matrix, as it is suggested by expression (9). If either D(I,K) or D(K,J) are equal to ∞ the inspection is not made as it is not worth.

Therefore:

$$D(1,3) = \min \left[D(1,3), \min_{\substack{K=1,NUSTO \\ K\neq 1 \text{ and } K\neq 3}} \left(D(1,K) + D(K,3) + T_K \right) \right]$$

$$\begin{aligned} D(1,3) &= \min \left[999, (D(1,2) + D(2,3) + T_2) \right] \\ &= \min \left[999, (8.7+9.6+SUM(I,II)-SUM(I,II)-SUM(-VI,I,II) \right] \\ &= \min \left[999, (8.7 + 9.6 - 1.6) \right] = \min \left[999, (16.7) \right] = 16.7 \end{aligned}$$

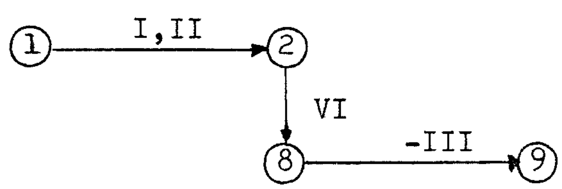
FLOWCHART 3. ROUTE-1



This value is equal to the result given above following a different approach. In this case, according to the rules to update MR and MT, MR(1,3) = IR = I,II and MT = 0. The results and the whole development of the network under study are shown in Appendix B.

In order to show the results obtained by ROUTE-1, the minimum path between (1) and (9) is considered. The minimum distance is given by D(1,9) = 34.7; last transfer is made at MT(1,9) = 8; and MR(1,9) = -III which means that from (8) to (9) this route is to be used. Now from (1) to (8) MT(1,8) = 2 and MR(1,8) = VI which means that between (2) and (8) route VI is to be used. Finally, MT(1,2) = 0 (means that there are no transfers) and MR(1,2) = I,II. Then, the optimum path from (1) to (9) is given by Figure 24.

FIGURE 24. Optimum path from (1) to (9).



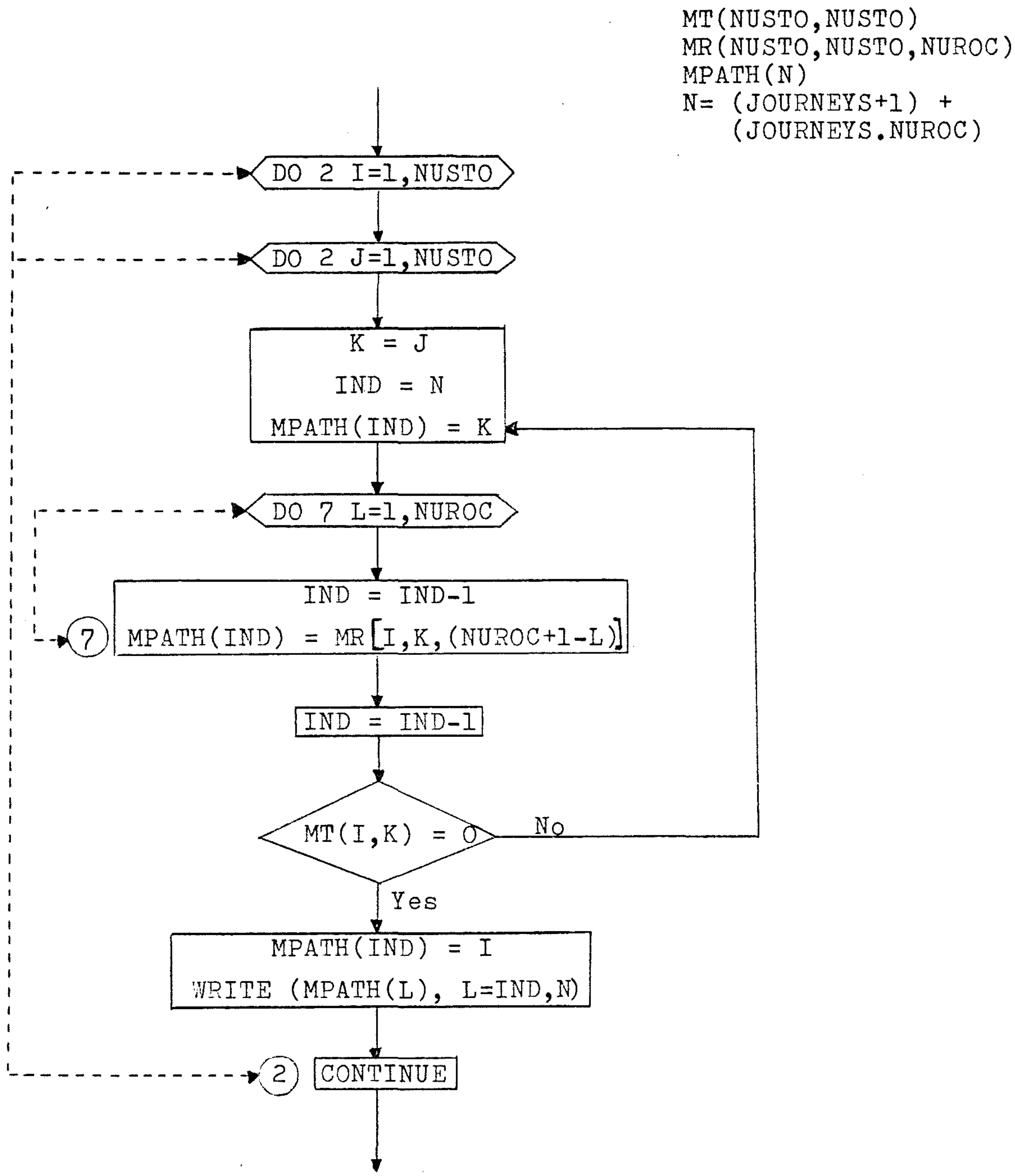
Flowchart 4 shows the way to retrieve minimum paths given matrices MT and MR; the respective paths are registered in matrix MPATH.

This simple example shows how practical ROUTE-1 is for registering paths and how easy it is to retrieve or recover some needed information. This model is capable of doing what the two-non matrix models which were described before do; however, ROUTE-1 has some disadvantages which are worth mentioning and it is worth noting that some of them also apply to the non-matrix algorithms.

i) Total travel time between two nodes is said to be

FLOWCHART 4. Subroutine for retrieval of minimum paths

(ROUTE-1)



composed of riding time plus waiting time at boarding points. This expression does not take into account the fare to be paid and its inclusion into the generalised cost expression could affect dramatically the results as is shown for instance in the minimum path between 1 and 9 (Figure 24) which contains two transfers, i.e. three boardings and three fares, with a total travel time of 34.72 minutes. However, taking bus route I it is possible to travel direct from 1 to 9 in 36.0 minutes (only one fare). The latter path would have been chosen the fares had been considered.

ii) In a realistic situation people could walk a little farther in order to catch a direct service avoiding a transfer and possibly an extra fare. Some people could also walk towards their destinations in order to reduce their fares and could also walk farther in order to widen their choice of routes. Therefore, the walking mode not only competes with the transport system but it is its natural feeder as well, giving people more freedom and saving in waiting time and expenditure of money. If the walking mode is to be considered then it is necessary to code dummy links in order to provide the respective level of service. This would mean a tremendous task in the coding of nodes and links and would also be reflected in the size of the matrices to be used.

iii) ROUTE-1 works on the basis of an all-or-nothing assignment which means that all trips between pairs of nodes are allocated to the links forming the minimum paths between them. Therefore, this technique could lead in some cases to a poor assignment as some non-optimum but reasonable alternatives to travel would not be taken into account. For instance, from 3 to

8 the optimum path consists in taking route VI for a direct journey of 22.0 minutes, while the same journey could also be made direct with route I in 24.0 minutes. However, ROUTE-1 has only registered the first option which is the optimum.

iv) In ROUTE-1 matrix MR(NUSTO,NUSTO,NUROC) is used initially as an input to the model in order to describe the routes serving every pair of nodes. In other words, this model requires a route description of every link in the network. But this approach is a computer memory consuming one and, on the other hand, in a real size network it could be too complicated and time consuming to give the routes in that way while it would be very easy to make mistakes.

v) As Flowchart 3 implies, ROUTE-1 was given in an unrefined and simple way. This means that there are several parts open to improvement. For instance, subroutine INTER is called too many times and this imposes a heavy burden on computing time.

Bearing in mind the previous considerations a new version called ROUTE-2 was developed from ROUTE-1. ROUTE-2, which is explained below, is considered to be the result of four main changes made to the first version:

i) In this model as in ROUTE-1 it is assumed that link riding times remain constant during the whole process. On the other hand, average waiting times are also taken as half the frequency of service and the respective expression is given in page 118. However, in this case total travel time is given by the following expression:

$$(10) \qquad TTT = t_r + W_{wt}\, t_{wt} + W_{wk}\, t_{wk} + F(t_r) + t\,c$$

where: TTT is total travel time in seconds,

t_r, t_{wt}, t_{wk} stand for riding, waiting and walking times in seconds respectively,

W_{wt} and W_{wk} are the weights or scale factors for waiting and walking times,

$F(t_r) = At_r + B$, where A and B are appropriate constants to give the fare in seconds,

t is the number of transfers made in a trip, in other words t is equal to the number of boardings minus 1,

c is a constant in seconds that represents an extra penalty for transferring.

Weight factors for waiting and walking times are used in order to unite the time and perceived cost elements of travel into a single measure (in this case riding time). These trade-off rates reflect the greater reluctance normally experienced by people to spend time waiting or walking rather than to spend the same time on board a bus.

For the purpose of this study it is assumed that fares rise linearly with riding time; however, where flat fares apply the constant A should be set to zero. Similarly, when no extra penalty for transferring is to be considered c is set to zero.

ii) ROUTE-2 will also deal with options including the walking mode but does not require the use of dummy links. The model works with a new matrix D2(NUSTO,NUSTO) which contains the average walking time in seconds between every pair of nodes in the network. Consideration of other levels of service other than walking requires dummy links.

iii) In ROUTE-1 matrices D, MR, and MT were used to register minimum paths only. In ROUTE-2 basically the same principle applies. However, as a matter of routine this model also inspects other alternative paths while it is searching for the optimum ones and if they are worth registering then they are kept in a special file and when the process ends the respective printout is produced.

iv) In ROUTE-2 every route of the system is described as a string of consecutive stops. Although a route can have two directions, if both directions follow the same path then it is only necessary to describe one direction and the program assumes the other one automatically. For example, for the network in Figure 19 route IV goes $7 \rightarrow 4 \rightarrow 5 \rightarrow 2$, and given this direction the model also considers the other direction $2 \rightarrow 5 \rightarrow 4 \rightarrow 7$, and this direction is named -IV.

However, if there is a route with only one direction or a route with two directions which follow different paths then this is recognized by the use of a special code. This is the case, for instance, of route I described by means of the following stops: $1 \rightarrow 2 \rightarrow 3 \rightarrow 6 \rightarrow 9 \rightarrow 8 \rightarrow 7 \rightarrow 4 \rightarrow 1$ and has only one direction.

The basic data input to the model are given below: seven parameters, three matrices and a bus route description.

i) Seven parameters:

- W_{wt} the scale factor for waiting time
- W_{wk} the scale factor for walking time
- A the slope of the linear fare function
- B the independent parameter of the fare function
- c the extra penalty for transferring in seconds

- NUSTO the total number of bus stops in the network
- DOPTIM in percentage, a figure carefully chosen. If X is the optimum total travel time between I and J, then the model will register all the inspected paths falling in the range X⟶X(DOPTIM + 1).

ii) Three matrices:

- Matrix D1(NUSTO,NUSTO) which initially gives riding times in seconds between connected pair of nodes. At the end of a computer run this matrix gives total travel times between every pair of nodes. This recorded time takes into consideration all the elements of expression(10).
- Matrix D2(NUSTO,NUSTO) gives the average walking times in seconds between every pair of nodes in the network.
- Matrix of frquencies FR(NUROU) (where NUROU is the total number of routes in the system) gives the bus frequencies in seconds. If FR(M) is given a negative value, then route M only has one direction.

iii) Bus route description: every route of the network is given as a string of consecutive bus stops.

A special subroutine called INOUT organizes the information for the model. Having the bus route description, this subroutine determines the different bus routes which serve every link of the network. For the purpose of this study a group of routes serving a link is called Route-set and INOUT also arranges the obtained Route-sets into matrix MCLASS(LC,NUROC) where LC is the maximum number of Route-sets for the whole network. It is evident that every Route-set is unique although there could be several links in the network with equal Route-sets. On the other hand, LC is a number that could be increased as iterations are passing and new Route-sets are to be taken into account.

As in ROUTE-1, this model also requires a matrix MT(NUSTO,NUSTO) (whose cells are initially set to zero) to register the last transfer made in a trip, if any. Furthermore, matrix MR(NUSTO,NUSTO) will keep the Route-set number corresponding to the last part of the journey. In other words, if MR(I,J)=X then MCLASS(X,N) for N=1,...NUROC will give the route numbers to go from I to J. On the other hand, this time NUROU (total number of routes in the system) and NUROC (maximum number of routes that share the same link in the network) are also obtained by INOUT.

Matrix D1 contains initially the riding times and it can be transformed into a public transport matrix by adding the respective waiting time and fare. For this purpose subroutine SUM, which was introduced in page 132 is also used. Therefore, matrix D1 is transformed by means of the following expression:

$$D1(I,J) = D1(I,J) + A\,D1\,(I,J) + SUM\left[MCLASS(MR(I,J),N)\right] + B$$

for I=1,NUSTO
J=1,NUSTO
N=1,NUROC
$D1(I,J) \neq 0$ and $D1(I,J) \neq \infty$

In this model every value obtained by subroutine SUM is kept in the vector KSUM(LC) in such a way that this subroutine is called only once for each Route-set. In other words, if the model requires the average waiting time of the Route-set X for the first time, then subroutine SUM is called and the respective value is registered in KSUM(X); therefore, every time the average waiting time is required the model turns to vector KSUM instead of calling subroutine SUM.

This model also requires subroutine INTER to obtain the intersection between two Route-sets. However, since MCLASS is used to register the different Route-sets, then the definition of this subroutine is slightly different. Given two Route-set numbers X and Y, INTER(X,Y) gives MCLASS(X,N)$\cap$MCLASS(Y,N) for N=1,NUROC. When the intersection is different from empty set, then INTER checks whether the given set has already been registered before in MCLASS. In negative case MCLASS is updated and the value of LC increased by one.

In order to avoid subroutine INTER being called too many times another matrix is used to register the Route-set number obtained. So that given X and Y, this matrix will register a number Z which will be either zero when the intersection is the empty set or $0 < Z \leqslant LC$ representing a Route-set kept in MCLASS(Z,N), N=1,NUROC.

Bearing in mind that INTER(X,Y) = INTER(Y,X) then instead of a bi-dimensional matrix, this model uses a vector KINT(T) to record the value of Z where T = LC(LC-1)/2 for the maximum value of LC. Given X and Y where $X > Y$, Z will be kept in KINT(f(X,Y)) where $f(X,Y) = ((X^2-3X+2/2) + Y)$.

By this procedure, when the model needs the intersection of two Route-sets for the first time subroutine INTER is called and from then on the information required is obtained from KINT giving thr respective Route-set numbers in such a way that the first one is higher than the second one.

Given I and J ROUTE-1 inspects every K from 1 to NUSTO in order to determine the best alternative. In this analysis both sections I-K and K-J are completely covered by public transport so that a(bus x bus)option is obtained and called here first option. However, apart from this ROUTE-2 analyses two further options: i) I-K walking and K-J by bus called second option and; ii) I-K by bus and K-J walking called third option. The following example illustrates the three different options for the path I-J via K:

$$I \xrightarrow{\text{V,VI}} K \qquad\qquad K \xrightarrow{\text{III}} J$$

i) <u>$\text{OPT}_k 1$:(bus x bus)</u>

$$\text{Total travel time} = \left[W_{wt} \times (\text{waiting time at I} + \text{waiting time at K})\right] + \left[(\text{riding time I-K} + \text{riding time K-J}) \times (A+1)\right] + 2B$$

ii) <u>$\text{OPT}_k 2$:(walk x bus)</u>

$$\text{TTT} = \left[W_{wk} \times \text{walking time I-K}\right] + \left[W_{wt} \times \text{waiting time at K}\right] + \left[\text{riding time K-J} \times (A+1)\right] + B$$

iii) <u>$\text{OPT}_k 3$:(bus x walk)</u>

$$\text{TTT} = \left[W_{wt} \times \text{waiting time at I}\right] + \left[\text{riding time I-K} \times (A+1)\right] + B + \left[W_{wk} \times \text{walking time K-J}\right]$$

It is worth emphasising that the main purpose of this model is to register minimum paths; within that process some non-optimum alternatives are examined and when they are worth considering they are registered in a file as a matter of routine. However, this does not mean that all the paths amongst the range (optimum) $\rightarrow$(optimum x (DOPTIM + 1)) are considered. The model analyses combinations of optimum paths and hence there is no way to examine options composed of non-optimum sections. Furthermore, there are alternatives which are eliminated because of some basic assumptions.

It is assumed that people can walk from A to B to take a bus at B on their way to C even though they can catch it at A. Some people could walk towards their destinations in order to avoid one or more fare stages and some people could also walk with the view to widening their choice of routes. However, the reverse is not allowed as it is assumed that people do not get off the bus to walk when they can cover this distance in the bus in which they are travelling.

Another particular situation which could arise from the application of this procedure is shown by means of the following example:

$$I \xrightarrow{V} K1 \xrightarrow{V} K2 \xrightarrow{V} J$$

To travel from I to J via K1 is the same than to travel from I to J via K2 and according to its basic principles the model will consider two options when in fact there is only one. In order to avoid this the subroutine in charge of decoding the options obtained was provided with the facility of detecting equal options.

In mathematical terms ROUTE-2 is given by the following expressions and rules to update matrices:

$$(11) \qquad D1(I,J) = \min \left[D1(I,J), \min_{\substack{K=1,NUSTO \\ K \neq I \text{ and } K \neq J}} \left\{ OPT_k 1, OPT_k 2, OPT_k 3 \right\} \right]$$

for $I=1,NUSTO$
$J=1,NUSTO$
$I \neq J$
I,J not connected

where: D1 matrix of distances already transformed into a public transport matrix

$OPT_k1 = D1(I,K) + D1(K,J) + T_k + FA$

$OPT_k2 = ABS(D2(I,K)) + D1(K,J)$

$OPT_k3 = D1(I,K) + ABS(D2(K,J))$

$T_k = -KSUM(NBSMR1)$ if $NBSMR1 = NBSMRX$

$T_k = KSUM(NK) - KSUM(NBSMR1) - KSUM(NBSMRX)$ if $NK \neq 0$

$T_k = 0$ if $NK = 0$

NBSMR1 = ABS(MR(I,K))

NBSMRX = ABS(MR(K,JT))

JT bus stop where next transfer is made after leaving K in the way to J. If there is no transfer then JT = J

$KSUM(X) = W_{wt}$ SUM $\left[MCLASS(X,N), N=1, NUROC\right]$

$NK = 0$ if $MCLASS(NBSMR1,N) \cap MCLASS(NBSMRX,N) = \Phi$ for N=1,NUROC

NK = respective Route-set number otherwise

FA = c when there is a transfer at K, FA = -B otherwise

D2(I,J) keeps the walking distance (in seconds) between I and J multiplied by W_{wk}

Rules to update MR (applies the first to be fulfilled)

- MR(I,J) does not change if D1(I,J) is not updated
- MR(I,J) = ABS(MR(K,J)) if either

i) OPT_k1 is the optimum and there is a transfer either at K or between K and J or;

ii) OPT_k2 is the optimum

- MR(I,J) = NK when OPT_k1 is the optimum and the previous condition does not apply
- MR(I,J) = NUWK when OPT_k3 is the optimum. Where NUWK is the Route-set allocated to the walking mode.

Furthermore, MR(I,J) = -MR(I,J) when either OPT_k2 is the optimum or when $MR(I,K) < 0$

Rules to update MT (applies the first to be fulfilled):

- MT(I,J) remains unchanged when D1(I,J) is not updated
- MT(I,J) = K when OPT_k3 is the optimum
- MT(I,J) = MT(K,J) if MT(K,J) $\neq$ 0
- MT(I,J) = MT(I,K) if there is no transfer at K
- MT(I,J) = K when there is a transfer at K. If OPT_k2 is the optimum then MT(I,J) = K + NUSTO

The matrix MT is to register the node or bus stop where last transfer is made, if any, while the matrix MR is to register the last Route-set used in the journey from I to J. The rules mentioned above will only guarantee these two important requirements.

Although with the advent of new technologies computer memory and computing time are becoming less critical for the development of complex and large models two additional steps were considered for the sake of ROUTE-2 efficiency:

ii) Given I and J it is not worth considering every K as its geographical position could be quite beyond a reasonable range (e.g. from Glasgow to London there is no point inspecting a route via Inverness). So although ROUTE-2 does not deal with the nodes geographical position it takes them into account by inspecting only those options which fall within the range D1(I,J) > D1(I,K) + D1(K,J) - MAXW, where MAXW $=\left[W_{wt}\left(\text{Maximum bus frequency in the network} / 2\right)+B\right]$. The definition of MAXW is not arbitrary as it considers the maximum negative value of T_k+FA defined for OPT_k1.

ii) If within the process of finding the shortest path between I and J the value of D1(I,J) is not updated or improved, then it is assumed that the minimum for this pair has been reached and if there are no alternative paths to the optimum then this pair is forgotten for the remaining iterations. A pair which has already reached the optimum is distinguished by giving a negative value to its respective cell in matrix D2.

It was said before that ROUTE-2 differed from ROUTE-1 in the sense that it was able to record additional options to travel between two nodes. However, in doing so that model cannot rely completely on matrices MT and MR as they can only cope with optimum options. Therefore, it was necessary to use a Work File for the process of registration and eventually the subroutine to retrieve the information became much more complex. Despite its complexity this subroutine works on the basis of the one shown in Flowchart 4 which takes into account the rules to update MT and MR; its inclusion was not considered relevant for the purpose of this research.

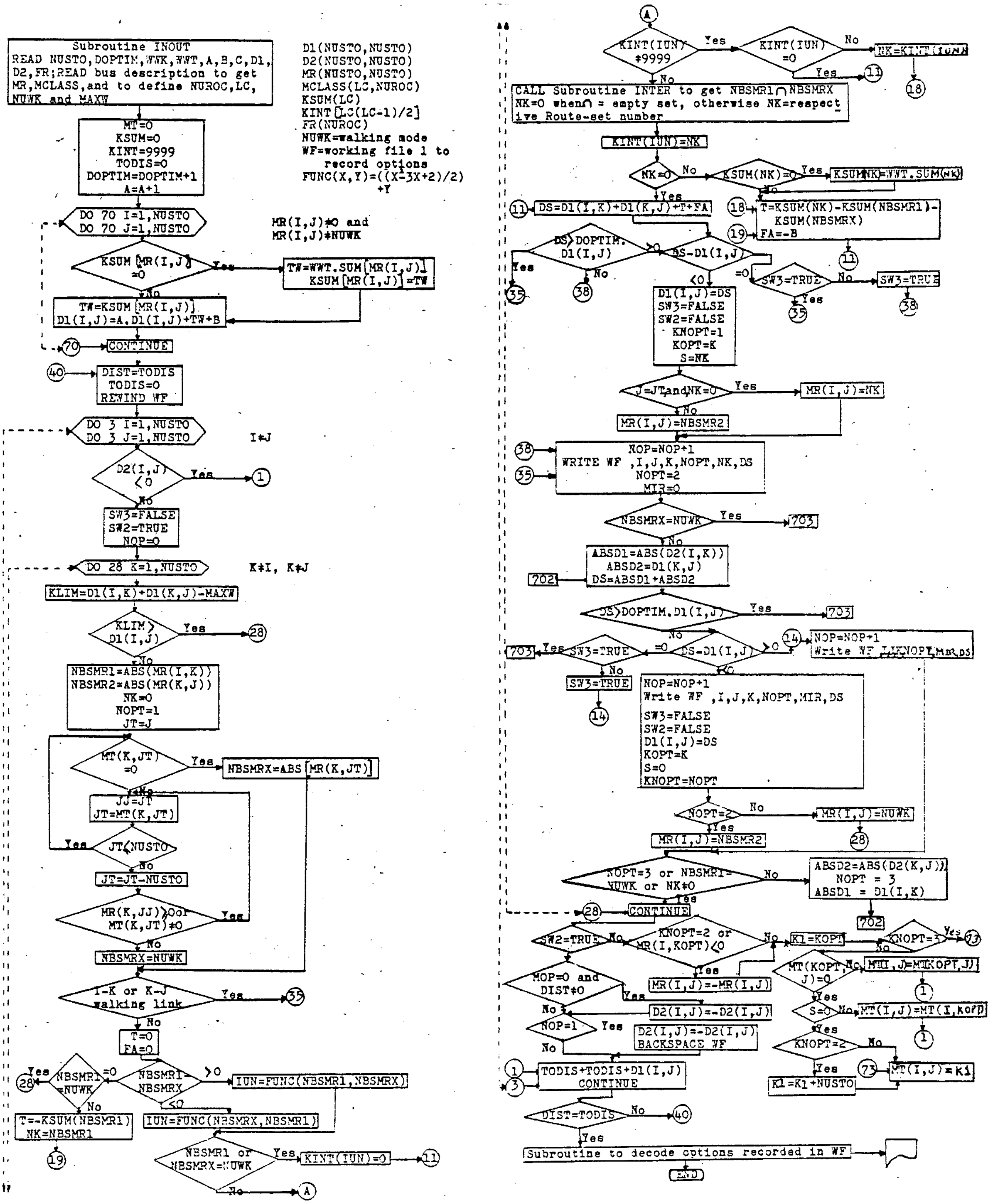
As in ROUTE-1 when matrix D1 has no further changes between two iterations is because the optimum options have been reached and the process ends by itself. Flowchart 5 shows the basis steps for the programing of ROUTE-2. From this diagram it is possible to see that ROUTE-2 is much more complex than the first version and therefore it has to work harder.

With respect to the outputs, ROUTE-2 produces two different files that are to be used as inputs to the subsequent submodels of the system:

i) a file that contains the bus route description of the transport network and the bus routes that constitute every defined Route-set;

ii) a file that keeps the different options to travel between every pair of nodes in the network. Each trip is broken

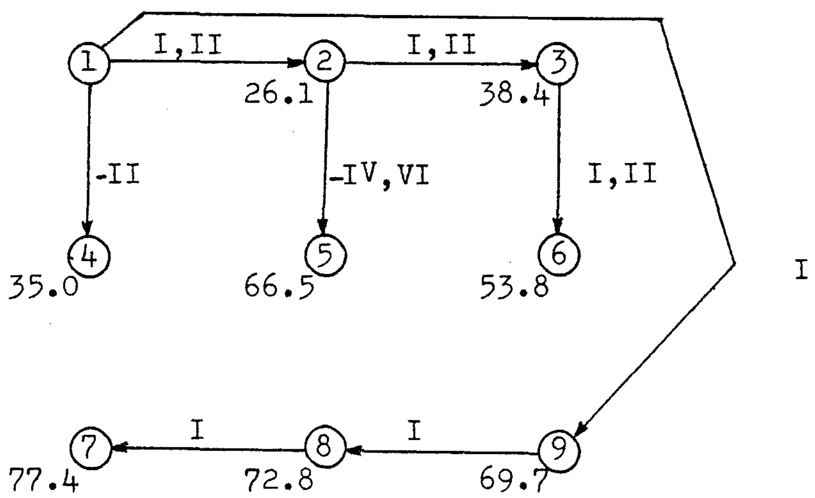
FLOWCHART 5. ROUTE-2



down into single journeys giving the respective Route-set that makes reference to the route services available to cover that distance. For every option total travel time is decomposed into its components: riding, waiting and walking times and fare. When required this model can produce the respective printout of this information.

With the view to showing how this model works Appendix C gives the complete development of the network presented in Figure 19 giving arbitrary values to the parameters. As an illustration Figure 25 shows the minimum paths from node (1) to every other node in the network and Figure 26 shows alternative options to travel from that node.

FIGURE 25. Minimum paths from node (1).



In summary ROUTE-2 constitutes a simple though precise and effective way of obtaining options to travel between any two pair of nodes in a network. It only requires as input data: the location of nodes and bus stops within the study area, walking and riding distances between connected nodes, the bus route description where routes are given as a sequence of bus stops, the bus service frequencies and the appropriate parameters to calculate generalised costs.

FIGURE 26. Alternative options from node 1.

$1 \xrightarrow{-II} 4 \xrightarrow{III,IV} 5$ 69.3

$1 \xrightarrow{-II} 4 \xrightarrow{walk} 5$ 76.3

$1 \xrightarrow{I,II} 2 \xrightarrow{VI} 7$ 80.4

$1 \xrightarrow{II} 4 \xrightarrow{-IV} 7$ 84.9

$1 \xrightarrow{I,II} 2 \xrightarrow{VI} 8$ 75.8

2. Loading of passengers onto selected options.

Having obtained the different alternatives to make a particular trip and knowing the number of people wanting to travel, then the next step is to allocate passengers to options. In other words, if there are T passengers wishing to travel from A to B and there are I options to make the trip then the problem is to assign passengers to every option, the simplest case being when there is only one option so that all users will take it.

The following example shown in Figure 27 illustrates this situation: a passenger travelling from A to B could take either route I or route II for a direct journey and this constitutes the first option; he could also take route III to C and then route IV to B (option 2); or he could walk to D and then take route V to B (option 3).

A good and reasonable assumption is to suppose that people will travel trying to minimise their travel costs, which means that the higher the costs of an option the less the number of people taking it. On the other hand, travel costs for each option are given by ROUTE-2 in terms of total travel time which is assumed to summarize the expenditure in time and money incurred by the passenger. Therefore, the total number of passengers taking option i, T_i, would be inversely proportional to the total travel time of that option and this fact is expressed by the following mathematical expression:

$$T_i = \frac{(1/TTT_i)^m}{\sum_{j=1}^{I} (1/TTT_j)^m} T \qquad (12)$$

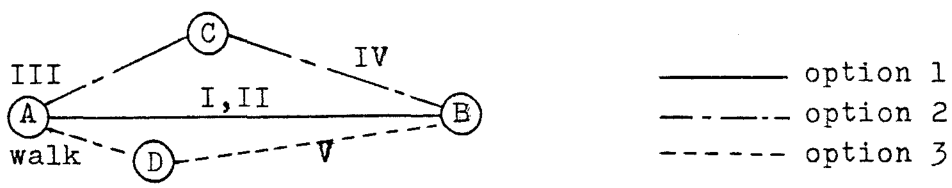
where: TTT_i is total travel time of option i

T is total number of passengers willing to travel

m constant parameter that considers passenger's sensitivity

$$\sum_{i=1}^{I} T_i = T$$

FIGURE 27. Theoretical options to travel between A and B



In charge of apportioning people amongst the different options is the submodel ASSIGN which works on the basis of the theoretical framework defined above. The input data required by this model is described below and consists mainly of three parameters, one matrix and a file produced by ROUTE-2:

- NUSTO: total number of nodes in the network

- NUWK: Route-set number assigned to walking mode

- m: parameter that expresses passenger sensitivity for expression (12) given above

- DEM(NUSTO,NUSTO): matrix of origin-destination which makes reference to door-to-door movements

- File of options to travel between any pair of nodes broken down into single journeys and including total travel time and its different components.

For the hypothetical example of Figure 27 DEM(A,B) = T gives the total number of passengers wanting to travel between these nodes. From the file of options the values of TTT_1, TTT_2 and TTT_3 are obtained and then the obtaining of T_1, T_2 and T_3 is straight forward. Having calculated the number of passengers assigned to each option the model works with single journeys and forgets the concept of option. This means that for instance, T_2 passengers using option 2 are equivalent to T_2 people taking route III from A to C plus T_2 people travelling from C to B with route IV. Similarly, T_3 passengers using option 3 are equivalent to T_3 passengers walking from A to D plus T_3 passengers travelling from D to B in route V.

Therefore, ASSIGN transforms the O-D matrix given by DEM into another matrix MOD(NUSTO,NUSTO) which contains journeys instead of trips as it records stop-to-stop movements. This means that the total number of trips given by DEM should be less (equal when there are no transfers) than the total number of journeys obtained by MOD.

Matrix MOD constitutes the main output of this model and its information is complemented by the route services available to make the respective journeys. It is worth noting here that when there are two or more Route-sets available to make a particular journey (i.e. they take different paths for a direct journey) the model treats this situations separately, but finally for the purpose of the file to be used by subsequent models it assumes that people wait at bus stops for the first bus that comes irrespective of the path followed.

It is very important to emphasize that despite the fact that the situation is simplified and that some of the information will not be used any more by subsequent models no piece of information is lost. Quite the reverse, the model arranges this information in a suitable and useful way for the analyst who may wish to know more about the transport facilities offered by the network and the manner people use them. Two different sets of data are produced by ASSIGN:

i) Information about the walking links which are used with more frequency by passengers and the origin-destinations which involve more walking according to the registered number of trips.

ii) Similar information to the one produced by ROUTE-2 with the difference that the print-out produced by ASSIGN includes the number of people apportioned to every option. See for instance Appendix C which contains the complete results for the bus network given in Figure 19. With the purpose of organizing this output in a different way a new variable called TYPE is defined as:

$$\text{TYPE} = 2^{\text{Nbuses}} + \text{Nwalks}$$

where: Nbuses is the number of journeys made by bus within a trip or option. When the number of journeys is equal to 1 then Nbuses is set to zero

Nwalks is the number of journeys made walking within an option.

The definition of TYPE implies that the higher its value the more complex the option analysed. For instance, for an option with one bus and one walk TYPE = 2, while for an option with three buses and no walks TYPE = 8 . Taking into account these concepts SORTASSIGN classifies the options according to their complexity and the number of trips. This output could be useful for the analyst in order to define policies related to transferences within the transport system for instance.

3. Loading of passengers onto selected routes and buses.

Having transformed the door-to-door matrix into a stop-to-stop matrix the next problem is to assign passengers to routes and buses. If there is only one route available to make a journey a user is supposed to catch the first bus which arrives subject to seat availability. Similarly, when there is more than one route available people are supposed to take the first bus regardless of the route number.[4]

In general terms, when there are N bus routes to make a journey, and each route runs buses every f_i $(i=1,2,\ldots N)$ minutes, where f_i is the bus frequency of route i, then the number of buses per hour per route i will be $60/f_i = bh_i$. Therefore, the probability that a bus of route i arrives at a bus stop is given by:

$$p_i = bh_i \Big/ \sum_{j=1}^{N} bh_j \qquad \text{where } bh_j = \text{buses per hour route } j$$

For two routes sharing the same section the respective probability will be p_1 and p_2 , where $p_2 = 1-p_1$. This means that if there are

P passengers wanting to make a journey, then it may be expected that Pp_1 people will take route I, while the rest $P(1-p_1)$ will take route II, approximately.

The previous expression does not take into consideration the bus capacity which could be critical in some circumstances and it does not allow to obtain useful information such as average waiting times at bus stops, total travel times, passengers not boarded because of seat unavailability and relevant factors in bus service operation. As the analytical formulation of the problem to get that information would have been too complicated it was considered that simulation would constitute an adequate way to tackle this question. It was also chosen because of the variability of transport demand and the need for evaluation of complex queueing processes.

Since the simulation of a system requires the generation of an artificial experience (which can be considered to be characteristic of the situation) then before the process itself is presented two relevant topics will be discussed: the pattern of arrival of passengers at bus stops and the model describing the time that buses take moving between stops.

A. Passenger arrival process. The pattern of arrival of passengers at bus stops may be regarded either as dependent upon the schedule of buses and the knowledge of passengers about it or as random which means that there is no evident association between the arrival times and the expected departure of buses from stops, or even between the arrival of other passengers.

The assumption of random arrivals could be fairly realistic for short bus headways (less than about ten minutes), while for longer frequencies of service people would try to time their arrival in order to catch specific buses (Jenkins, 1976). Therefore, for the purpose of this study it was assumed that people arrive randomly at bus stops. This means that an arrival can occur at any time subject only to the fact that the mean arrival rate has some given value. In more formal terms, the time of the next arrival is independent of the previous bus or passenger arrivals and the probability of an arrival in an interval Δt is proportional to Δt.

A random arrival pattern can usually be described by means of a Poisson distribution (Gordon, 1978) which states that the number of arrivals in any interval of time is independent of the number of passengers who arrived during the previous time interval. Thus, as was explained in the first chapter, by generating a uniform random number and using the cumulative distribution of the Poisson function it is possible to obtain the number of arrivals for a period of time t.

However, in order to generate the exact times of arrival of passengers at bus stops the model, called ARRIVE, uses the exponential distribution which is regarded as the continuous analogue of the Poisson distribution. This procedure will allow the model to record very easily individual waiting times. The cumulative function of the exponential distribution is given by:

$$F(t) = 1 - e^{-\lambda t}$$

where λ is the mean rate of arrival per unit of time and t is the time interval.

The problem of sampling from any distribution can be regarded as that of transforming a random uniform number by means of the inverse cumulative distribution function (Tocher, 1963). The inverse cumulative function of the exponential distribution is given by:

$$(13) \qquad t = -\frac{1}{\lambda} \ln U$$

where U is a uniform random number. Therefore, knowing the average rate of arrival and using a random number series it is possible to generate the inter-arrival times with the view to determining the exact time of arrival of passengers at bus stops.

The main inputs to ARRIVE are the stop-to-stop O-D matrix, MOD(NUSTO,NUSTO), where NUSTO is equal to the number of nodes in the network, and the respective bus routes to make the respective journeys. On the other hand, random numbers are generated by using the NAG library mounted on the ICL-2976 of the University of Glasgow.

If there are MOD(A,B) people wishing to travel from A to B, then λ = MOD(A,B)/T, where T is the period of time of the O-D matrix, value which is also to be supplied to the model. By using random numbers arrivals are generated for a time period TS which is directly related to the time the simulation is going to last. In this way ARRIVE creates a file which keeps information about every passenger who has arrived. A passenger register contains his origin, destination, time of arrival and bus routes for a direct service between origin and destination. As the simulation requires this file to be classified by bus stop origin and arrival time, submodel SORTING organizes the information accordingly.

As was suggested in the first chapter, the appearance of expression (13) is very simple but the calculation time can be high as the natural logarithm is obtained by means of a series that converges slowly. Therefore, ARRIVE was given the possibility of working with either natural logarithms given by the computer or with a table search. The latter saves computer time at the expense of more storage requirements and some loss of accuracy. The respective table can be created by submodel LOGAR.

B. <u>Inter-stop travel variation submodel</u>. One variable which is a function of the technical characteristics of the equipment, the prevailing traffic conditions and the qualities of the driver is the inter-stop travel variation. However, between two specific bus stops there is some minimal travel time which is not possible to reduce due to either the characteristics of the vehicle concerned or traffic regulations. On the other hand, buses can incur delay because of stopping due to decelerating to and accelerating from bus stops. Additional delays can occur when re-entering the traffic stream.

The variations in travel time between bus stops can be assumed to follow a gamma distribution with parameters t_o and p (Jenkins, 1976) given by the following expression:

$$f(t) = p^2(t - t_o)\ e^{-p(t - t_o)}$$

where: t_o is the minimum possible travel time

$p = 2/(\bar{t} - t_o)$

$\bar{t}$ is the mean travel time

The respective cumulative function can be obtained by the following integer:

$$F(t) = \int_{t_o}^{t} f(x)dx = 1 + e^{pt_o}\ e^{-pt}(pt_o - pt - 1)$$

and from here:

(14) $\ln\left[1 - U\right] - pt_o = \ln\left[1 + pt - pt_o\right] - pt$

Therefore, by means of a trial and error procedure, giving p, t_o and U it is possible to calculate the travel time t.

Given t_o and $\bar{t}$, submodel TRAVAR working with expression (14) estimates values of t for random numbers U between 0 - 1, creating the respective table. Thus, the whole process is reduced to a table search in which giving a random number it is straightforward the obtaining of t.

It is expected that buses between two stops will experience similar traffic conditions if they are running closely, which means that they will probably spend the same time to cover this particular distance. If buses are running farther apart they will probably face different traffic conditions and their travel times are likely to be more independent; however, this does not necessarily mean that they cannot be equal. For this reason this model includes a correlation factor developed by the IBM (Gerrard and Brook, 1972) which is a function of the bus headway h:

(15) $$D = e^{-ah}$$

where a is an appropriate constant. Having obtained a travel time for bus m, t'_m , by means of the table provided by TRAVAR, then the excess time over the mean travel time $\bar{t}$ is calculated by the following expression:

(16) $$Y = DX + \sqrt{(1 - D)^2}\, e$$

where: X = excess time for the previous bus along the same link, then $X = t_{m-1} - \bar{t}$

e = excess time for the bus in question, $e = t'_m - \bar{t}$; and then,

$$t_m = \bar{t} + Y$$

This means that if the pair of buses in question are running close together then $h \simeq 0$ and $D \simeq 1$, and therefore $t_m \simeq t'_{m-1}$. On the other hand, for a longer headway $D \simeq 0$ and then $t_m \simeq t'_m$ which means that the travel times are not correlated.

C. Simulation of public transport operations. The submodels explained in the previous sections of this chapter are used to prepare the input data to submodel SIMULA whose main functions are: i) the execution of the cycle of actions involved in carrying out the simulation of the operations of a public transport system in an urban area and; ii) the gathering of meaningful statistics during the simulation process and the preparation of an output report.

In SIMULA buses are considered to be entities in the system to be simulated. The characteristics or attributes of the buses are expressed by means of the input variables to the model and it is assumed that events are represented by the arrival of buses at stops. One special kind of stop is the route terminal where buses start and end their journeys.

The inputs to the model comprise the following information:

i) Bus route description and number of routes in the system (NUROU). Every route is described as a sequence of bus stops.

ii) Frequencies of services for each route of the system in seconds. FR(R), R=1,...NUROU.

iii) Inter-stop travel variation tables to determine bus travel time between any two pair of stops in the network.

iv) Time of arrival of passengers at every bus stop of the network giving their destinations and bus services to be taken.

As it was considered there was no iteraction between these arrivals and the events of the system it is possible to create a sequence of arrivals in preparation for the simulation (Gordon, 1978).

v) Number of buses per route BPR(R), R=1,...NUROU.

vi) Bus capacity[5].

vii) Parameter a for correlation factor given by expression (15).

viii) Crew recovery time for buses arriving at stop termini.

Having given the number of buses per route, every bus of the system is completely identified by a code number and the route service to which it belongs. This easy identification allows SIMULA to keep a record of every bus with the following data:

- Bus capacity
- Number of passengers on board
- Bus position given as a bus stop number
- Journey direction, either 1 or 2
- Number of passengers alighting at every bus stop of the route served by the bus.

Initially, the number of passengers on board for each bus is assumed to be zero. In order to complete the initial condition of the system the model assumes that for every route R,BPR(R)/2 buses are initially at stop terminal 1 ready to start journey in direction 1. Similarly, there are $\left[BPR(R) - BPR(R)/2\right]$ buses at terminal 2 to start journey in direction 2. With this information the initial state of the system is completely defined.

The passage of time in the simulation is recorded by the model by means of a number referred to as clock time which at the beginning is set to zero and subsequently indicates how many seconds[6] of simulated time have passed since the start of the simulation process. Then, the next step is to define the sequence of events.

Clearly, if there are N buses in the transport system, then at any point in time there are N events which are due to occur in the near future. SIMULA works with a special subroutine called EVENTQUEUE which keeps events organized according to their future chronological occurrence. All that remains is to establish the time of occurrence of the different events of the system. Considering the assumed initial conditions it is therefore necessary to determine the time buses will be leaving their respective terminal in order to start their journey.

For every route R at every terminal it is assumed that the first bus will be arriving at the first stop[7] of the route at time FR(R), the second bus at 2FR(R), the third at 3FR(R) and so forth. Working on these basis subroutine EVENTQUEUE organizes the different events in chronological order, which clearly means that the first event to occur will involve a bus of the route with the highest frequency of service (lowest headway between buses).

The model considers that the first event to occur will be the arrival at a bus stop of a specific bus on a specific route, and therefore it updates the clock time to equalize the relevant arrival time. Having the bus position the model checks the number of people alighting there in order to update the number of passengers on board. If there is room for more people then SIMULA reads the arrival file in order to board the passengers who have arrived before the bus departure time and who are waiting for that particular service. People are boarded on the basis of first come first served but always taking into account the bus seat availability. Considering the passengers destination the array which stores the information of the number of people alighting at each stop is updated accordingly.

Average waiting times are easily calculated by obtaining the difference between boarding times and passenger arrival times. The model also considers the passenger arrival times in order to work out the total travel time. It is worth mentioning that as soon as passengers are allocated to a bus they loose their identity and will not be treated individually any more.

Boarding time is assumed to be equal to $(p_b \times A_b)$, where p_b is the number of passengers boarded and A_b is the average boarding time. Similarly, $(p_a \times A_a)$ gives the alighting time, where p_a is the number of passengers getting off and A_a is the average alighting time. It is assumed then that the time spent at bus stops by any bus depends upon the highest time of the boarding and alighting processes. Consequently, knowing the arrival time at the stop the departure time will be completely defined.

Having the bus position and its direction, the model checks the route description in order to determine the next stop to be visited. By using the tables of the inter-stop travel variation SIMULA estimates the travel time between the bus actual position and the next stop. Thus, adding this travel time to the departure time the arrival time at next stop will be determined. In other words the time of the next event for the bus under consideration will be known. Then, the bus position is updated accordingly.

If the bus is already full when arriving at the stop or if it becomes full in the process of boarding, then the model checks the passengers' queue in order to determine the number of people left behind because of seat unavailability. Then, this bus will be given its departure time and accordingly its arrival time at next stop. The number of people on board when buses leave the stops is used in the statistical process to determine the average occupancy between stops.

Having work out the arrival time at next stop then subroutine EVENTQUEUE finds the position for this activity within the queue of events, always taking into account their chronological oder of occurrence. SIMULA is ready once again to deal with the first event in the queue after updating the clock time.

One special event which is considered differently is the arrival of a bus at a stop terminal. At these stops people get off the bus but there are no passengers boarding. Having the bus frequency FR(R) of route R, then SIMULA determines the next departure time in order to maintain a desired schedule. However, this departure time may not be the final one as it is essential to consider the time from which the bus is available for departure that is given by the arrival time at the terminal plus the respective alighting time plus crew recovery time. The latter is then compared with the schedule departure time in order to define the true departure time and gather the respective statistics. Next, the bus journey direction[8] is changed, the respective schedule is updated and the estimated departure time is assumed to be equal to the arrival time at the first bus stop of the route. Then, considering this event time, EVENTQUEUE finds the appropriate place for this event in the queue of event occurrences.

The clock time is used to guarantee that the process is simulated for as long as it is desired. However, it is clear that there is a bias in the statistics collected in a process like the one previously described due to the way buses are allocated initially. To remove this bias SIMULA allows the use of two different processes:

i) Buses can be moved through the network following exactly the same principles explained before but without loading and unloading passengers. This is a fast process which implies that the clock time continues being equal to zero and eventually no statistics are gathered.

ii) To simulate formally during a certain period of time. Stop the process and leaving passengers and buses where they are restart the simulation gathering fresh statistics from the point of restart (Gordon, 1978).

SIMULA gathers meaningful information at three different levels: stop, route and the whole transport system. The main statistics are given below:

- Average occupancy in percentage between stops, per direction and for the system;

- Average waiting time at bus stop, per direction and for the system;

- Total travel times per route and direction and for the system;

- Passengers boarding at each bus stop, per direction and for the system;

- Passengers alighting at each bus stop;

- Passengers not boarded due to bus seat unavailability at every bus stop, per direction and for the system and;

- Average differences between the expected and the actual departure time at bus termini and summary for the system.

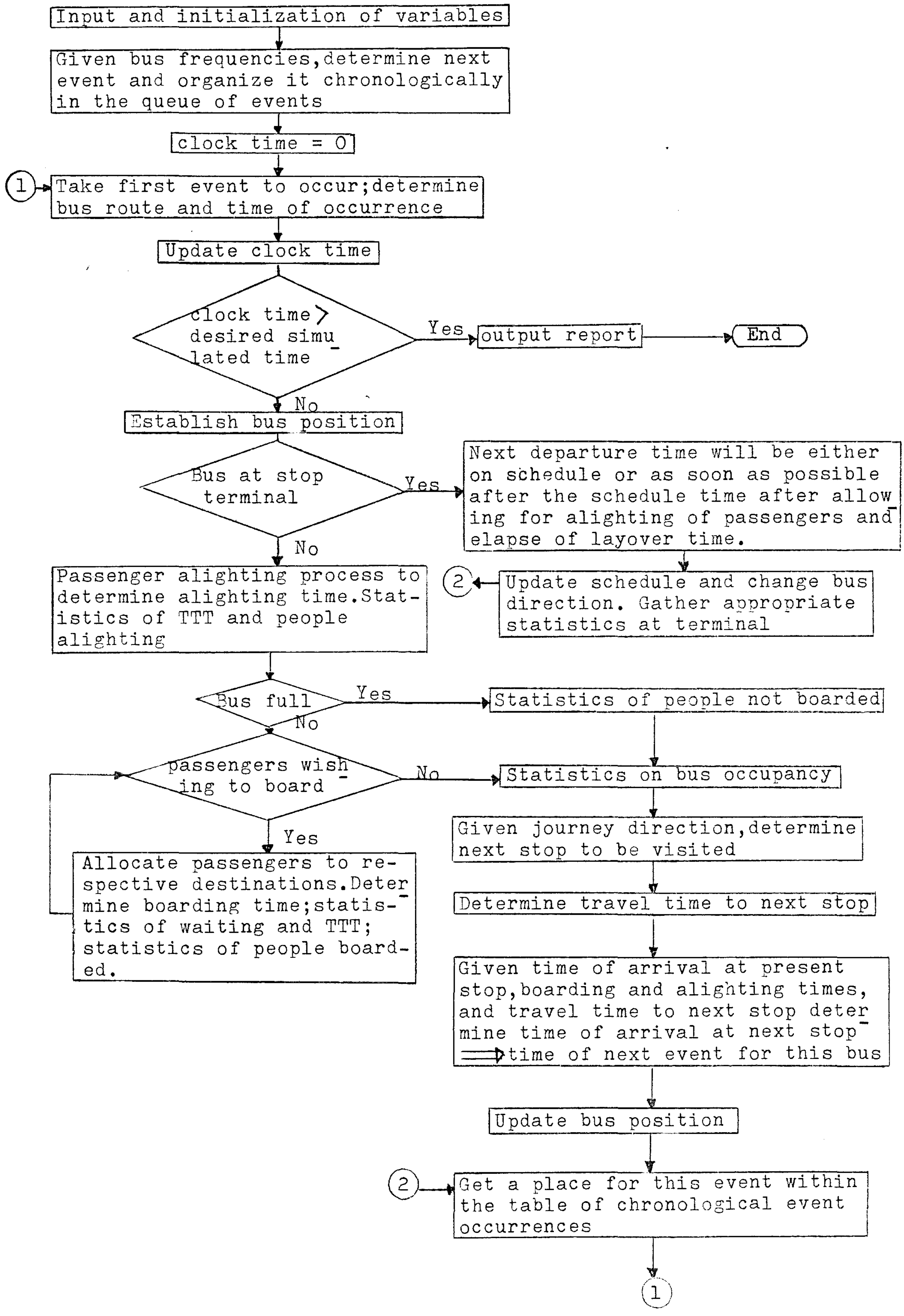
Flowchart 6 shows diagrammatically the logical relationships between the successive steps in the simulation process. SIMULA constitutes a basic model for the simulation of bus operations in urban areas. Although for the purpose of this study the model is considered to be rather adequate it is worth noting that the completion of several subroutines whose parameters are already considered in the program will allow to analyse other aspects of bus operations such as sources of irregular headway, tendencies of some buses to overtake vehicles of the same route, etc. Illustration of how SIMULA displays results will be shown in the next chapter.

In summary, the proposed public transport model attempts to describe the travel patterns of people in an urban area using a series of linked submodels and may be considered to be a description of the decision-making process which an average passenger might be expected to follow when he considers making a bus journey. Flowchart 7 shows the different submodels of the model and the way they are interrelated.[9]

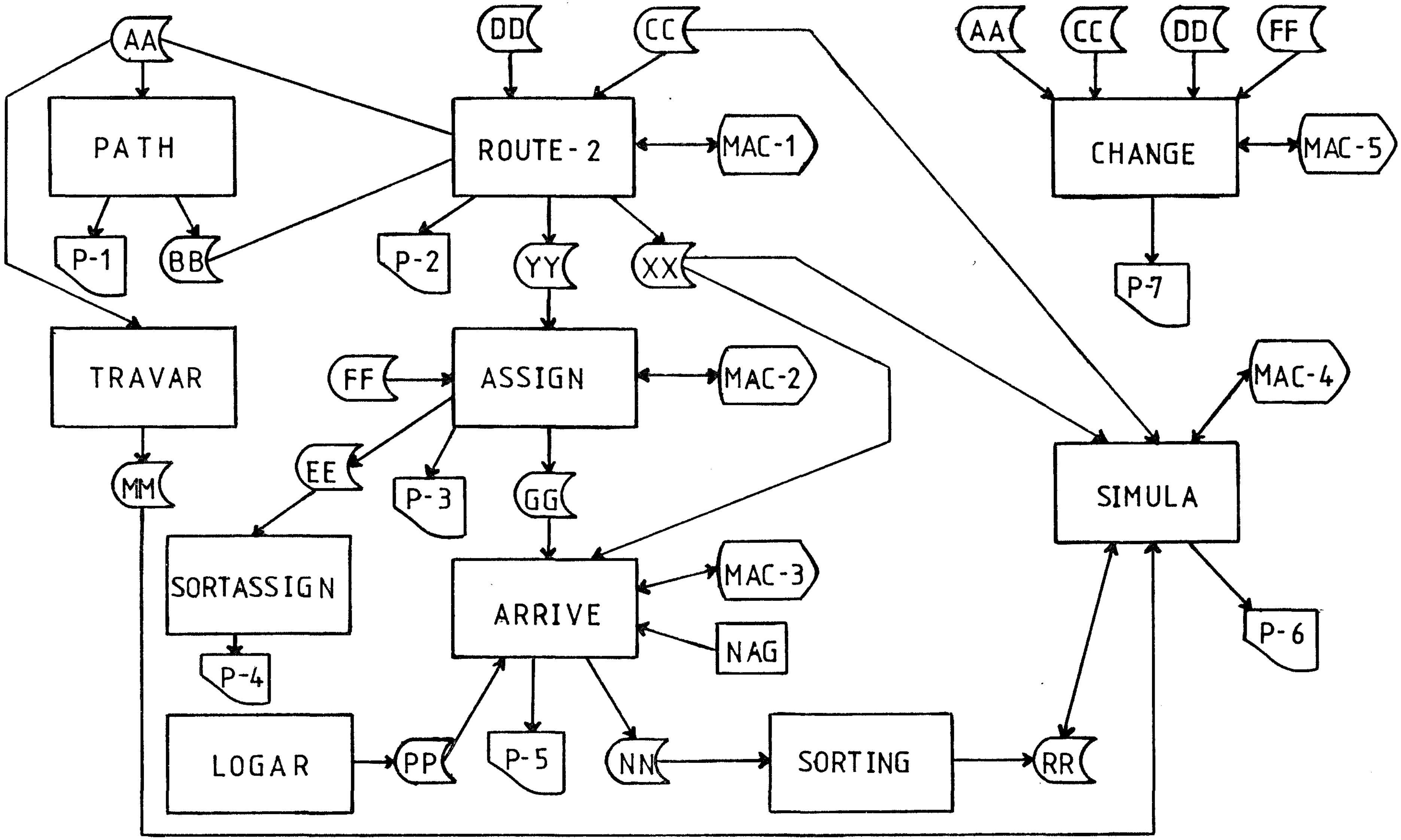
All the computer programs were written in FORTRAN with the exception of SIMULA, SORTING, and SORTASSIGN that were written in COBOL. In the programming and testing of the programs the large scale ICL-2976 computer of the University of Glasgow was used. This computer provides both batch and interactive access.

The following chapter is devoted to the application of the model in a real situation in order to see whether it can produce results consistent with the observed data.

FLOWCHART 6. Basic steps of SIMULA



Flowchart 7 - Proposed Public Transport Assignment Model+



* Note to Flowchart 7.

AA- Bus stop spacing and inter-node distance; average and maximum bus speeds; average walking speed
BB- Matrix of walking times
CC- Bus frequencies
DD- Bus route description (I)
EE- Intermediate file
FF- O-D matrix
GG- Transformed O-D matrix; matrix of Route-sets
MM- Cumulative distribution function for inter-stop travel time
NN- Passengers arrivals per O-D
PP- Table with natural logarithms between 0-1 (optional file)
RR- Passenger arrivals per bus stop
XX- Bus route description (II); route set description
YY- Parameters; options to travel between stops
NAG- Numerical Algorithms Group Library (random number generation)
MAC-1- Set of parameters such as fare, weighting factors for waiting and walking, etc.
MAC-2- m, power for assignment function
MAC-3- Length of simulation; use of file PP; etc
MAC-4- Bus capacity; number of buses per route; layover time; etc
MAC-5- Required statements to modify respective files
P-1- Printout with minimum distances and path description between every pair of nodes of a network
P-2- Printout with set of options to travel by public transport between every pair of nodes of a network
P-3- Printout similar to P-2 but includes the number of people taking every option
P-4- Printout similar to P-3 but classified according to option complexity and number of people taking every option
P-5- Statistical tests made to the process of passenger arrivals at bus stops
P-6- Summary of simulation per route and direction and for the whole system of buses
P-7- Changes made to files; approximate number of buses needed per route

NOTES TO CHAPTER II.

1) A trip refers to door-to-door movement, therefore it may comprise one or more public transport journeys and possibly journeys made by another mode of transport, for instance walking.

2) The following definitions and concepts are considered of particular relevance for a better understanding of this chapter:
An algorithm is a method of solving a problem by following a sequence of precisely stated successive steps.
A network consists of a number of nodes, each one being connected to one or more nodes by links. Two nodes are said to be connected if there is a link that joins them; and two links are connected if there is a node common to both.
A path is a sequence of distinct nodes and direct links joining them.
A cycle is a path in which the initial and final nodes coincide.
A network is said to be connected when each of its nodes is connected to every other node. A tree is a connected network without cycles.

3) Moore, E.F., The shortest path through a maze. International Symposium on the Theory of Switching Proceedings, Harvard University, April 2-5 1957.

4) It is assumed that with the exception of service frequency the routes provide a similar level of service.

5) The program allows to work with two different bus capacities per route.

6) Unit of time chosen for SIMULA.

7) Evidently the first stop position coincides with the terminal position.

8) When a bus arrives at a terminal is initially given a provisional journey direction, either 6 or 7 according to its real direction.

9) Submodel CHANGE provides facilities to make modifications to files containing basic data such as inter-node distances, bus stop spacing, bus frequencies, bus network description, origin-destination information. It also gives an initial estimate of the number of buses per route in order to run a desired frequency of service.

CHAPTER III

A CASE STUDY

Model validation or verification is considered to be one of the most critical problems in computer modelling and simulation and can be regarded in two stages: validation of model concepts and validation of model implementation. A model may be assumed to be valid if it produces results which are meaningful when properly interpreted and it measures what it is supposed to measure. However, in strict terms to validate a model means to prove the model to be true (Naylor, et al, 1968) and this question involves such a complicated set of criteria that it could hardly be said that the purpose of the work described in this chapter was to validate the model whose theoretical framework has been discussed in the previous chapter.

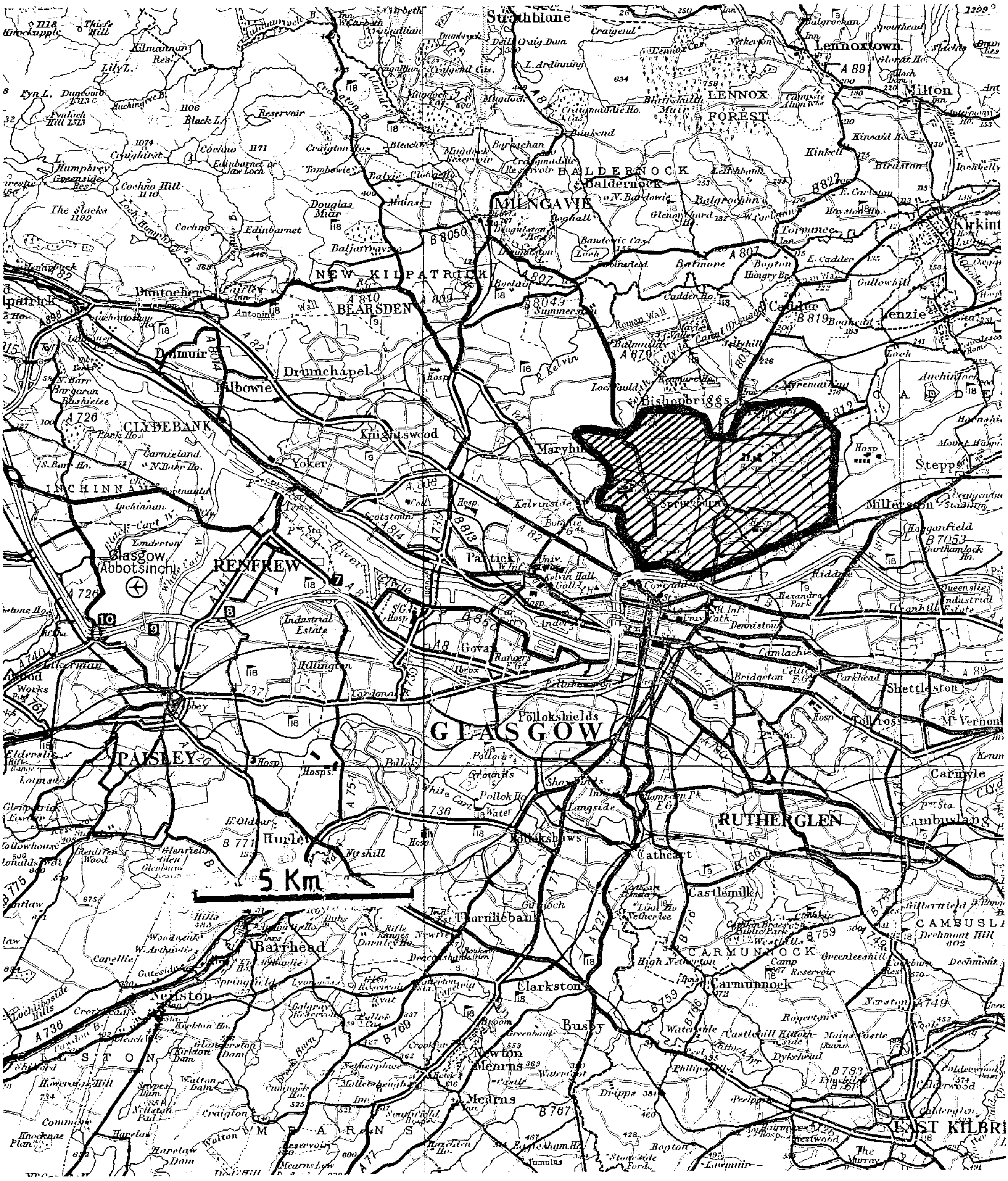
Therefore, the main objective here was to test the computer package in a real situation and to show whether it could produce logical and meaningful results. The first section of the chapter gives a brief description of the Study Area, the section summarizes the survey and assumptions used, and the third section presents the results and conclusions obtained.

1. The Study Area.

In order to test the model an area of approximately 20 square miles and a population of 82,000 was chosen. This area includes the districts of Cadder, Possilpark, Hamiltonhill, Lambhill, Milton, Possil, Cowlairs, Port Dundas, Colston, St Rollox, Springburn, Balornock, Barmulloch and parts of Ruchill and Auchinairn. Map 1 shows the Study Area in relation to Glasgow.

Map 1

Study Area in relation to City of Glasgow



Springburn, the core of the Study Area, which is $2\frac{1}{2}$ Km north of the Glasgow City Centre, grew up out of the railway boom of the mid-nineteenthcentury. It was the centre of the North British Locomotive Company, one of Glasgow's greatest engineering industries, and eventually became the most important place in Britain for locomotive design and railway construction. The closure of this industry in 1963 with the immediate loss of employment was generally regarded as a disaster for Springburn, which after being an important and busy industrial centre became a poor and deprived area with high unemployment.

With respect to the transport system the Springburn area contains 37 routes. The Scottish Bus Group operates 23 which basically radiate from the centre of the city of Glasgow. Of the 14 routes which the Greater Glasgow Passenger Transport Executive (GGPTE) operates only one (route 8) runs completely inside the area, having both termini within it. Apart from the bus network, there is an electric rail service which links Springburn with Glasgow city centre.

Most of the information needed for this study was taken from the Public Transport Survey carried out in 1975. However, in order to complete the data set required it was necessary to turn to several sources and studies made in other cities of Britain as there was no available information for Springburn.

2. <u>Data input to the study</u>.

The Springburn Public Transport Survey was carried out in April 1975 for the GGPTE as part of the Executive policy to find out whether any modification to the bus route network in Glasgow was required. Springburn was chosen by the GGPTE for study as there was a need to review transport facilities in the area because of the proposed construction of the Springburn Expressway and the future comprehensive redevelopment of the area that would include a District Shopping Development with a bus terminal.

The Springburn Survey began on the 14th of April and lasted for four weeks. One bus service was surveyed on each day of the week (except Tuesdays) and the objective was to record every journey starting within the Study Area between 0700 and 1900 hours, and when this was impracticable then the results were expanded accordingly. The data produced were:

i) Number of passengers boarding and alighting at each bus stop per hour.

ii) Number of passengers using each section of bus route per hour.

iii) Number of passenger-miles on each section of route.

iv) Origin-destination matrices per hour and purpose of trip.

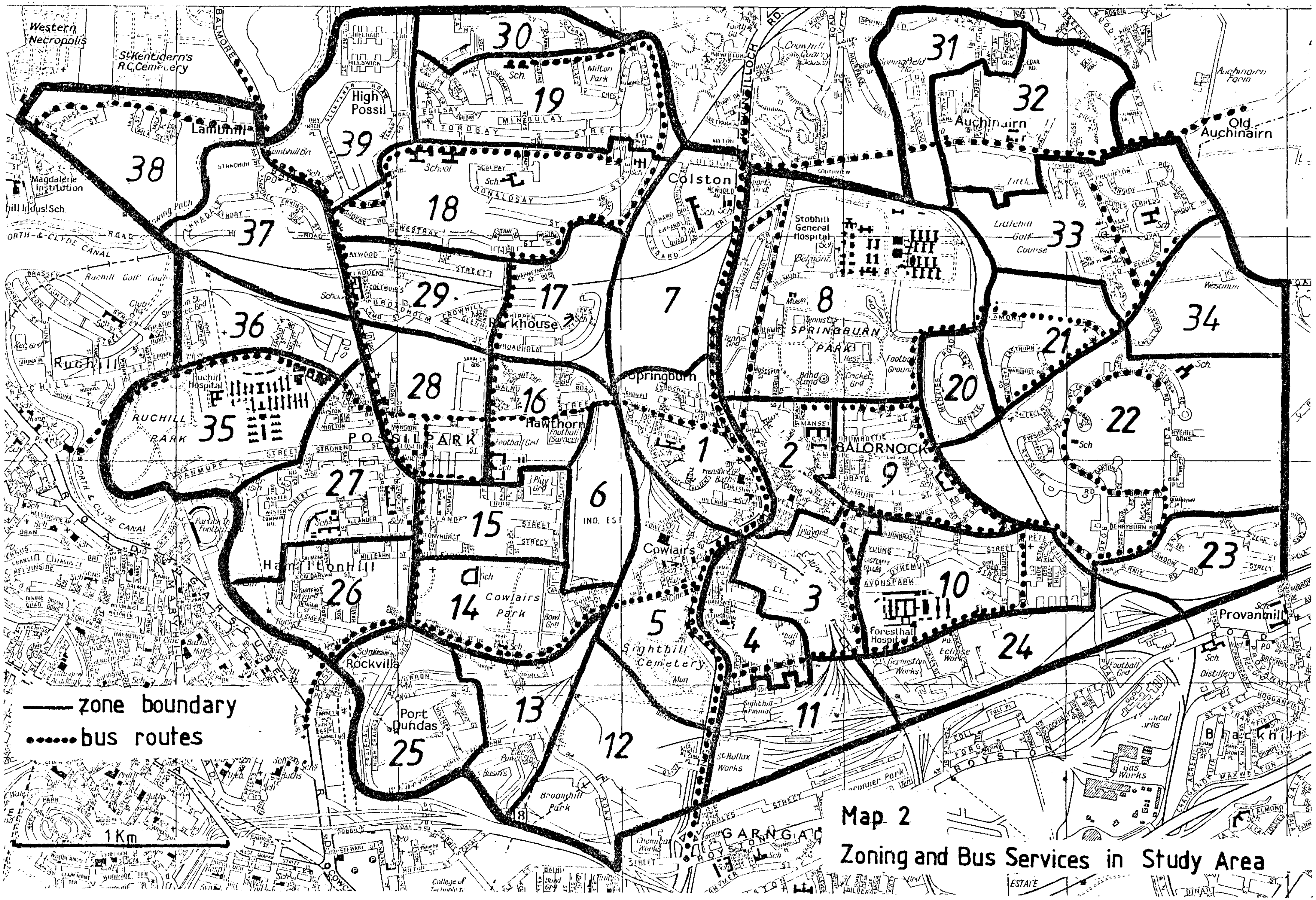
Temporary survey staff recruited for the purpose were in charge of recording the passengers movements. They issued a card to passengers boarding a bus inside the Study Area indicating the bus stop at which they had boarded by means of a code number. These cards were collected and placed in a pocket file corresponding to the alighting bus stop when passengers left the bus. Cards of passengers still on board when the bus crossed the area boundary were placed in a special pocket of the file. This data was the basis for preparing the matrices of boarding and alighting movements and eventually served to yield information on passengers travelling on each section of route and passenger-miles carried.

In oder to obtain O-D data, one card in four handed out by the survey staff had a questionnaire printed on the reverse. Since the data related only to a sample of passengers on a sample of buses in the day an expansion was necessary. The factors considered in the expansion of the O-D data took into account the total number of passengers boarding each service per hour and the total number of buses run in each hour.

For the O-D matrices the Springburn area was divided in 39 internal zones and in order to consider internal/external trips 38 additional external zones were considered outside the Study Area. Map 2 shows the internal zoning and public transport network in the area. Taking into account the 77 zones a total of 52,765 trips per day were made and of them 20,690 trips were made completely inside the area. It is important to note that the 77 x 77 matrix does not represent the total movements through the Study Area as for instance a passenger travelling from Glasgow to Springburn would not have been surveyed unless he had made a transfer within the Study Area. Furthermore, someone travelling between two external zones would not have been surveyed at all.

Given the constraints of the 77 x 77 matrix the present study was restricted to movements with both origin and destination inside the Springburn area and therefore only the matrix of 39 x 39 zones was considered. The main goal was to try to reproduce the pattern of transport inside the area between 0700 and 1900 and to compare the results obtained with the information given by the survey.

The O-D matrix constituted the transport demand within the area of study, but for the purpose of this study it was necessary to concentrate the demand at points whose geographical positions could be established easily. This step was essential



Map 2
Zoning and Bus Services in Study Area

for the bus route description and for the determination of riding and walking distances between nodes of the network. As the main intention was to design a regular size network, easy to handle and suitable for the purpose of this study, 44 nodes were considered (then NUSTO = 44) from which there were 35 bus stops and 9 nodes connected to stops by walking links. However, in this process land use considerations and the O-D matrix were also taken into account. Table 23 shows how the 39 zones were grouped and concentrated into the 44 nodes.

TABLE 23 .Grouping and concentration of zones into nodes.

Node	Zones	Node	Zones	Node	Zones
1	1	16	16	31	31,32,23
2	2	17	17	32	-
3	3	18	18	33	-
4	4,11	19	19	34	34
5	5	20	-	35	35,36
6	6,13,14	21	21	36	-
7	7	22	22	37	37,39
8	8	23	23	38	38
9	9,20	24	24	39	-
10	10	25	25,26	40	-
11	12	26	-	41	-
12	-	27	27	42	-
13	-	28	28	43	-
14	-	29	29	44	-
15	15	30	30		

Source: Grouping made by the author based on Map 2.

Then, in this way the O-D matrix of 39 zones was transformed into a 44 x 44 matrix. On the other hand, in some cases it was necessary to reassign passengers to different O-D as the ones given by the transformed matrix were unsuitable for the way the model requires its inputs. This was the case, for instance for trips with same origin and destination (travel within the same zone), or trips between two connected nodes when there was only a walking link between them. In these situations the O-D was modified accordingly taking into account the zones involved, the bus network and the location of real bus stops in the area.

As in general for the Scottish Bus Group services boarding is not permitted within the Study Area on inward journeys, nor alighting on outward journeys then these services were not considered and only the 14 routes operated by the GGPTE were taken into account.[1] In other words it was assumed that all movements inside the area were made using the GGPTE services. This is not a strong assumption as apart from the restrictions for boarding and alighting inside the area the SGB services have in general very low frequencies as is shown in Table 24.

TABLE 24. Off-peak frequencies of SGB services.* (in minutes)

Route	Freq.	Route	Freq.	Route	Freq.
14	60	175	30	191	30
20	30	177	30	192	30
27	120	180	**	193	30
170	30	181	**	194	30
171	30	182	30	195	30
172	30	183	60	196	**
173	infrequent	184	60	197	**
174	30			198	**

Source: GGPTE, Greater Glasgow Transport Map.

* At peak periods there may be in many instances a more frequent service.

** Only peak periods.

In order to get the bus frequencies of the services operated by the GGPTE, the total number of buses dispatched per route from both termini between 0700 and 1900 hours was considered.

Table 25 shows the bus frequencies in seconds for each route; both directions having the same frequency.

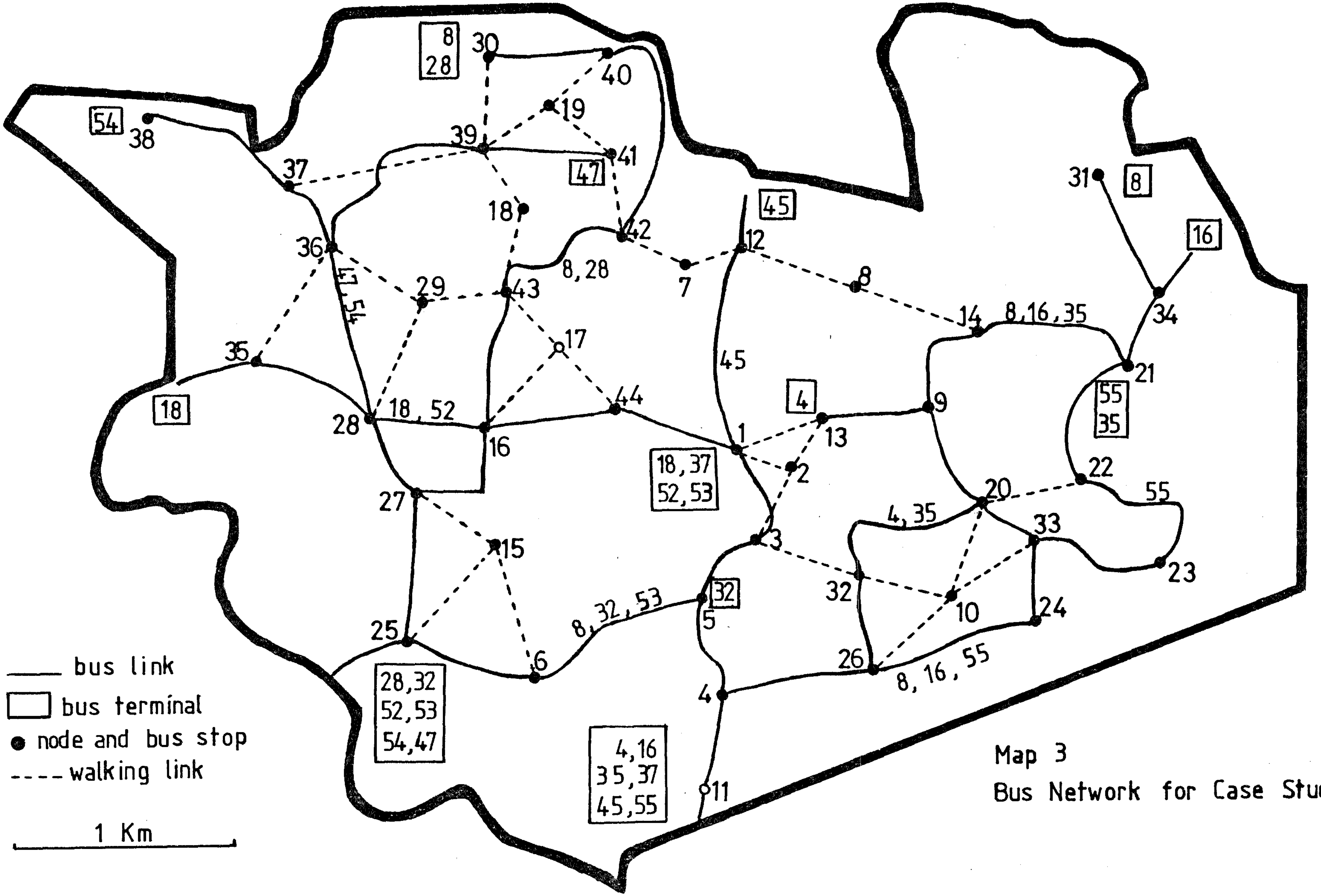
TABLE 25 . Bus frequencies for GGPTE services.

Route	Dispatched buses (0700-1900 hours)	Frequency (seconds)
4	135	640
8	131	660
16	116	745
18	229	377
28	200	432
32	145	596
35	100	864
37	217	398
45	201	430
47	208	415
52	167	517
53	180	480
54	178	485
55	128	675

Source: Springburn P.T. Survey, 1975 and author's elaboration.

Map 3 shows the network of 44 nodes and the bus routes that were considered in the case study. On the other hand, Table 26 shows the bus route description given as input to the model.

With the purpose of giving the matrix of riding distances between connected nodes, the distance in metres between them was measured, and assuming an average bus travel speed of 21.2 Km/hour (The Bradford Bus Study (BBS), 1976 and Chapman, 1976) it was transformed into seconds. For the walking distances the procedure was similar and it was assumed an average walking speed of 1 m/sec (Chapman, 1976). It is worth mentioning that walking links were basically restricted to main streets and



Map 3
Bus Network for Case Study

established paths as there are several obstacles that impede access such as depots, works areas and rail tracks.

TABLE 26 . Bus route description.

Route	Bus stop sequence
4	11 4 26 32 20 9 13
8	30 40 42 43 16 27 25 6 5 4 26 24 33 20 9 14 21 34 31
16	11 4 26 24 33 20 9 14 21 34
18	35 28 16 44 1
28	25 27 16 43 42 40 30
32	25 6 5
35	11 4 26 32 20 9 14 21
37	11 4 5 3 1
45	11 4 5 3 1 12
47	25 27 28 36 39 41
52	25 27 28 16 44 1
53	25 6 5 3 1
54	25 27 28 36 37 38
55	11 4 26 24 33 23 22 21

Source: Author's elaboration based on GGPTE routes and Table 23.

In order to standarize units weighting factors for waiting and walking times were used. These values, taken from the Bradford Bus Study (1976), were the result of a behavioural analysis made on a sample of 180 people which concluded that walking and waiting times could reasonably be weighted by 1.75 and 2.25 respectively.

Due to lack of empirical evidence two variables were given arbitrary values. They were DOPTIM whose value was assumed to be 15% and c, here called extra penalty for transferring, whose value was set to zero. However, it is important to notice that the DOPTIM value was taken relatively low in order to avoid the registration of options that could include backtracking.

Routes run by the GGPTE operate basically a stage system with a maximum fare for any journey over four stages for adults and over two stages for children. Therefore, the fare was calculated according to the distance covered by the passengers in the form of stages, being the average distance between two fare stages of approximately 0.55 miles. The fare structure in operation at the time of the study is presented in Table 27.

TABLE 27 . Bus fare structure in Glasgow. (January-June 1975)

N° of stages	Fare in pence
1-2	6
3-4	12
over 4	15

Source: GGPTE.

One important addition to the stage fare system was the Transcard which was introduced in 1974 to benefit both the regular traveller on bus/underground services and the operator. This facility allows the user of GGPTE bus and underground services freedom of travel at any time and on any service other than the premium night services. The Transcard price was based on the cost of 40 journeys per four-weekly period at the maximum fare, less a discount of a fixed amount. Two cards were in operation, one for adults and another for juveniles.

It is also worth mentioning that at the time of the Public Transport Survey there was a concessionary travel scheme for pensioners who were paying a 1 penny fare on Glasgow Corporation vehicles.

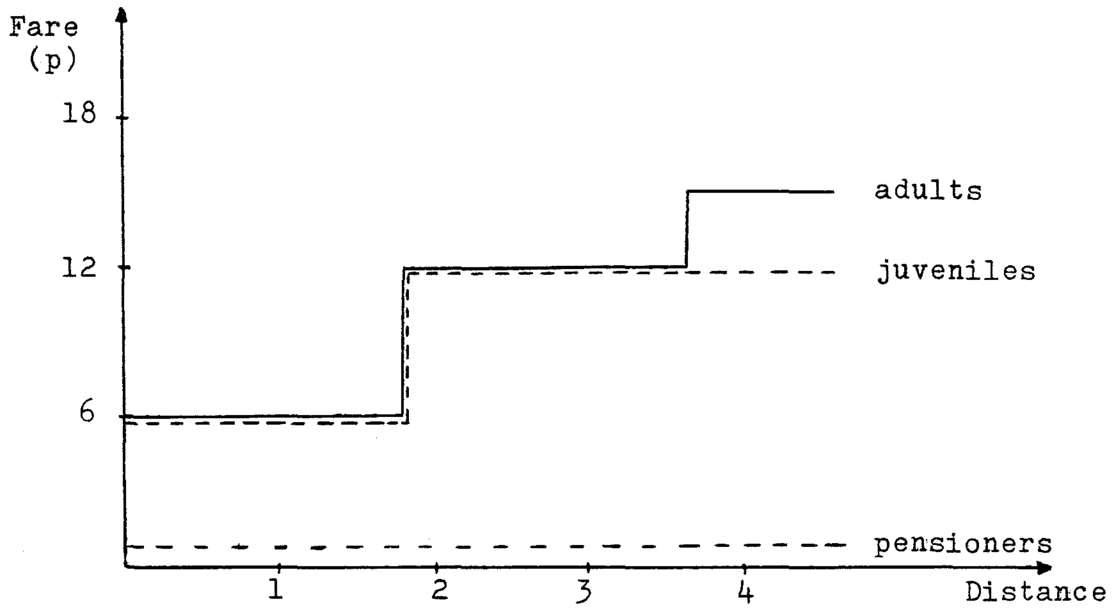
The relationship between fare and distance travelled is given in Figure 28 , taking into account three different groups:

adults, juveniles and pensioners. Then, with the view to considering these groups the fare function was estimated by regressing the fare on the distance travelled and the following expression was obtained:

$$\text{Fare} = 3.16 + 1.80\ D$$

where D is given in kilometres and fare in pence.

FIGURE 28 . Relationship between fare and distance travelled.



Source: GGPTE and author's elaboration.

As for the purpose of this study the fare should be expressed in units of time (seconds) then a behavioural value of ride time of 38 pence per hour at September 1975 prices was used (BBS, 1976). This means conversion of fares at the rate of 94.7 seconds per penny, and once again assuming an average travel speed of 21.2 Km/hour, then the fare would be given by the following function:

$$\text{Fare} = 300 + 1.0\ D$$

whereFare and D are given in seconds. Therefore, variables A and

B of submodel ROUTE-2 are given by 1.0 and 300 respectively.

Given a set of options to travel between two nodes, the probability that a person chooses option i was given by the expression:

$$p_i = \frac{(1/C_i)^m}{\sum_{j=1}^{I} (1/C_j)^m}$$

where C_i generalised costs option i;
I number of options; and
m parameter expressing traveller's sensitivity.

For this particular run it was assumed that m = 2 as there was no empirical evidence that suggested something different. For traffic assignment based on attractivity Overgaard suggests a figure between 5 and 6.[2]

Seating capacity per bus per service was taken from the transport survey and is given in Table 28 .

TABLE 28. Bus capacity per route.

Route	Capacity (passengers)
4,8,16,18,28,32,37,45,47,53,55	76
52	70
35,54	65

Source: Springburn Public Transport Survey.

In order to get the approximate number of buses per route, the bus journey time was divided by the bus route frequency. The bus journey time was considered to be made up of three components: travel time between bus stops, time at bus stops and time at termini. But since neither the time spent at bus stops nor the time at termini were known it was assumed that the bus journey time was approximately equal to 1.2 times the travel time between stops. The calculation was made by submodel CHANGE which was introduced in the previous chapter and the results given in Table 29. It is worth remembering that only route 8 runs completely inside the Study Area and the other routes are cut before they cross the Study Area boundary. Therefore, the number of buses obtained by CHANGE do not have to correspond exactly to the number of buses in operation for those particular routes.

TABLE 29 . Number of buses per route.

Route	Number of buses	Route	Number of buses
4	3	37	3
8	8	45	4
16	4	47	5
18	4	52	3
28	6	53	4
32	3	54	4
35	3	55	4

Source: Calculations made by the author.

It was assumed that when a bus arrives at a bus terminal after completing a journey then its next departure will be programmed according to a schedule requirement. In other words, if the bus arrives late then there will be no time of recovery for the driver and if it is early then he will wait in order to comply with the particular schedule.

Time spent by buses at stops depends basically on boarding and alighting times. An average boarding time of 3.28 seconds/passenger was obtained as the mean of a set of values given by Chapman (1976) for different cities in Britain. On the other hand, an average alighting time of 1.2 seconds/passenger was assumed (Chapman, 1976).

Finally, with respect to the correlation between successive journey times this model assumed the form $D = e^{-ah}$ where a is an appropriate constant and h is the bus headway. The value of a was assumed to be 8/60 = 0.133 (Gerrard and Brook, 1972) which means that in five minutes the correlation would be reduced by about half.

One important aspect is worth mentioning before the results of the model are presented. This concerns the options obtained by ROUTE-2 to travel between two nodes of the network. It is clear that this submodel selects options with the lowest total travel time and this is perfectly consistent with its purpose. However, this does not mean that all the options found will be considered reasonable from the traveller's point of view. This situation is illustrated by means of the following example: one possibility for travelling between 21 and 40 (see Map 3), given the data set used in this case study, consists in taking a bus from 21 to 14 (routes 8, 16, 35), then walking to node 42 to take either route 8 or 28 to reach the final destination. This option has a total travel time of 4794 seconds made up of 284 sec. of riding time, 571 of waiting time, 3055 sec. of walking time and 884 sec of fare equivalent. The optimum option to make that trip is to take route 8 for a direct journey having a total travel time of 4260 seconds, made up of 1609 sec. of riding time, 742 sec of waiting time and 1909 of fare equivalent.

Although both options were consistent with the theoretical framework of the model and both were registered by ROUTE-2, only the second one was considered by submodel ASSIGN. the first option was eliminated on the grounds that a traveller would be more prepared to do the whole journey on foot rather than taking two different buses and walking more than half the total distance. Therefore, submodel ASSIGN was provided with the ability to make a further selection of options taking into consideration the proportion of walking time to total travel time. Due to lack of empirical evidence the criteria to eliminate options had to be established after a careful examination of both the results obtained by ROUTE-2 and the bus route network. Then, the following type of options were eliminated by ASSIGN:

i) Options in which walking time was more than 50% of the total travel time;

ii) Options with any single walk of more than 1200 generalised seconds;

iii) Options with more than three buses and zero walks; for instance three buses and one walk.

3. Results and conclusions.

Examination of the results of the test runs can be considered as an intuitive procedure based mainly upon judgement, with the help of several measures of performance borrowed from applied mathematics and statistics. Here, three measures of performance were utilized to examine the results:

i) Graphic comparison. This is a simple and commonplace way of analysing results and consists in plotting on a graph observations of the prototype and the corresponding estimates given by the model. This procedure allows the analyst with his own judgement to visualize how well the model compares to the real world.

ii) Root mean square (rms) or standard deviation. This is a method of measuring the difference which exists between the prototype and the results of the model. The rms is given by the following expression:

$$\sigma = \text{rms} = \sqrt{\frac{\sum (X_o - X_e)^2}{n}}$$

where: X_o is the observed value

X_e is the expected value

n is the number of observations

The rms is used to determine the percentage of error or coefficient of variance V which is given by the expression:

$$V = \frac{\text{rms}}{\bar{X}_o}$$

This coefficient measures the relative difference between the observed and expected data. Although the values given by rms and V could be interpreted differently by different analysts, a key rule is that for a good model rms and V should be low; for a perfect fit evidently V should be zero (Catanese, 1972).

iii) Chi-square test which was already discussed in a introductory chapter.

As a matter of procedure results will be presented in the same chronological order as they were obtained from the different submodels, the same suggested by the General Flowchart of the Proposed Public Assigment Model: PATH ⟶ ROUTE-2 ⟶ ASSIGN ⟶ SORTASSIGN ⟶ LOGAR ⟶ ARRIVE ⟶ SORTING ⟶ TRAVAR ⟶ SIMULA.

To begin with, an illustration of how ROUTE-2 and ASSIGN produce their results is shown in Table 30 . The first example gives three different options to travel between 4 and 37. It should be noted that the second and third alternatives differ in the sense that in the latter the traveller instead of

Table 30

Examples of results obtained from ROUTE-2 and ASSIGN (in seconds)

T.T.Time	R.Time	W.Time	Wk.Time	Fare	Passengers	Option Description
From 4 to 37						
3344	712	1020	0	1612	8	4 37/45 1 18/52 28 54 37
3338	726	1286	0	1326	8	4 8 25 54 37
3488	644	755	845	1244	7	4 walk 5 8/32/53 25 54 37
From 25 to 1						
1646	403	540	0	703	30	25 53 1
1834	477	580	0	777	25	25 52 1

Abbreviations: T.T. = total travel ; R. = riding ; W. = waiting ; Wk. = walking waiting at 4 for route 8 walks to node 5 where he widens his choice and so reduces his waiting time and fare. The second example produces two direct alternatives to make a trip between 25 and 1 using two different ways or paths. In the first case there are 23 people wishing to travel from 4 to 37 and submodel ASSIGN apportions them according to the rule that the higher the total travel time the less the number of people choosing that option. In the same way, there are 55 passengers wanting to travel fom 25 to 1 and ASSIGN distributes them accordingly between the different options.

From Table 30 and in general from the results obtained by ROUTE-2 it is interesting to note how this submodel in addition to deriving the optimum option for travelling between two points can produce, where possible, meaningful and rational alternatives for making the trip. It is also worth noting how it decomposes a trip into single journeys breaking down total travel time into its components: riding, waiting and walking time, and fare.

Another input to submodel ASSIGN is the O-D matrix that refers to total movements inside the Study Area (door-to-door). This matrix gives a total number of trips of 20,690, and according to the survey data the total number of journeys inside the area is 26,491; meaning that approximately 25% of the people take more than one bus to complete their trip. Then, with the decomposed options obtained by ROUTE-2, ASSIGN transforms the O-D matrix into a new one that only registers existing service movements (stop-to-stop). The total number of journeys obtained was 24,356, meaning a difference of about 8% with respect to the total given by the survey. This point will be discussed below after presenting the results of SORTASSIGN.

According to the way ROUTE-2 decomposes a trip into single journeys, an option could be classified according to the number of buses and walks constituting it. For instance, in Table 30 the first option has three buses, the second two buses and the third two buses and one walk, etc. It is evident that if an option has one bus then the maximum number of walks will be two; similarly, for two buses there should be no more than three walks, and so forth. Taking into account these considerations submodel SORTASSIGN classifies the options according to their complexity and number of people taking them. These results could help the analyst in the understanding of the transport network and the way people use it. A summary of the results is presented in Table 31.

Having obtained the first results it is worth analysing the differences with respect to the survey data. Assuming that the expanded O-D matrix is correct then the number of journeys obtained by the model is lower than the observed figure and the discrepancies could in part be explained by the following reasons:

i) The aggregation of zones and the location of nodes and bus stops in the Study Area.

ii) The inability to measure travel times and the other elements of perceived cost to a high degree of accuracy. One factor that could be critical is the weighting of the walking mode, as this mode apart from being a feeder to the public transport system sometimes competes with it. On the other hand, the fare used could also be an important source of discrepancy since for higher fares people tend to substitute other modes of transport, such as walking, for public transport, taking into account the elasticity of demand.

iii) The parameter m that expresses the user's sensitivity in the process of distribution of passengers between options. This variable is crucial as the higher its value the less the number of people apportioned to the non-optimum options and therefore the higher the number of passengers assigned to the optimum option.

iv) The fact that the model works with average perceived costs when in fact people could assess the value of the parameters in a different way.

TABLE 31. Composition of trips obtained by the model.

Number of buses	Number of walks	Trips	Journeys
1	0	9290	9290
1	1	7056	7056
1	2	866	866
2	0	1368	2736
2	1	1614	3228
2	2	308	616
3	0	188	564
TOTAL		20690	24356

Source: Calculations made by the author.

However, another source of discrepancy which cannot be completely ruled out is the possibility of deficiencies in the expanded O-D matrix either because the information used was incorrect or the expansion itself was faulty. According to the critique to the survey methodology[3], about one-third of the questionnaire cards were completely unusable, for several reasons: insufficient time to complete questionnaire, inability of passenger to read or understand questionnaire, refusal to cooperate in survey, illegibility of writing, etc. The remainder of the cards, about 8000 in total, although not entirely complete were considered to contain enough information for the purpose of the survey. Therefore, some doubts can be held about the credibility of the survey data.

As the results obtained by submodel ROUTE-2 and ASSIGN are to be used as input to the following submodels then the previous considerations, and specifically an underestimated O-D matrix produced by ASSIGN, should be taken into account when analysing the final results.

Assuming a Poisson distribution for the arrival of passengers at bus stops, the transformed O-D matrix is the basis for the generation of arrival times. Although the O-D matrix represents movements between 0700 and 1900 hours, for the purpose of this study case ARRIVE only generated enough arrival times for one hour of system simulation plus a pre-simulation period to allow buses to take position in the network.

Generation of pseudo-random numbers plays an important role in the arrival process and for this purpose the NAG library[4] was used. Although it is assumed that a generator like this one has already been widely tested some additional tests were undertaken. The Chi-square test is often applied and here the distribution of the set of numbers generated was compared to the uniform distribution, obtaining satisfactory results at a 90% significance level.

The Chi-square test was also applied to measure the difference between the observed O-D matrix and the number of arrivals generated using pseudo-random numbers. The computed χ^2 was 123.0 for 208 degrees of freedom. This means that the probability that there is no difference between the observed data and the model results is high or in other words, there is no significant difference between the observed and generated values.

Finally, working basically with the bus frequencies, the bus route description, the cumulative distribution function for inter-stop travel time, the passengers arrivals per bus stop, the bus capacity and the number of buses per route, the simulation was run with a view to reproducing the pattern of movements within the area.

Having the number of buses per route to run a particular schedule, SIMULA assumed that half that number were in one terminal and half in the other. In order to remove this initial bias the following procedure was used:

i) Buses were allowed to run during 20 minutes without picking up passengers. This was done to locate them in different positions of the network. In this process no statistics were gathered by the model.

ii) The simulation was then started, and stopped after another 20 minutes. This time passengers were loaded and unloaded.

Having done this, both buses and passengers in the system were left where they were when the process was halted after the 20 minutes of simulation, and the simulation was restarted with statistics being gathered from the point of restart. Finally, the model simulated bus operations in the Study Area during 60 minutes.

In oder to compare results it was once again necessary to turn to the survey data which gave the number of passengers carried inside the area per route and direction during the 12 hours of survey. Then, for comparison it was assumed that one-twelfth of the total number of passengers carried in 12 hours was equal to the total number of passengers boarded in one hour.[5] Figure 29 shows the total number of persons boarded in one hour per route and direction for both the survey and the simulated results.

In order to measure the difference between observed and simulated data two tests were applied: the root mean square and the Chi-square. The results obtained were as follows:

rms = σ = standard deviation = 17.5 passengers

V = coefficient of variance = 22.2%

$\chi^2 = 104 > 40.1$ (for 27 degrees of freedom and 95% significance level)

From Figure 29 and the application of the previous tests several conclusions could be drawn:

i) In general the results given by the simulation are quite close to the observed data, with the exceptions of routes 55 and 32 direction 1 and routes 4, 8 and 45 direction 2.

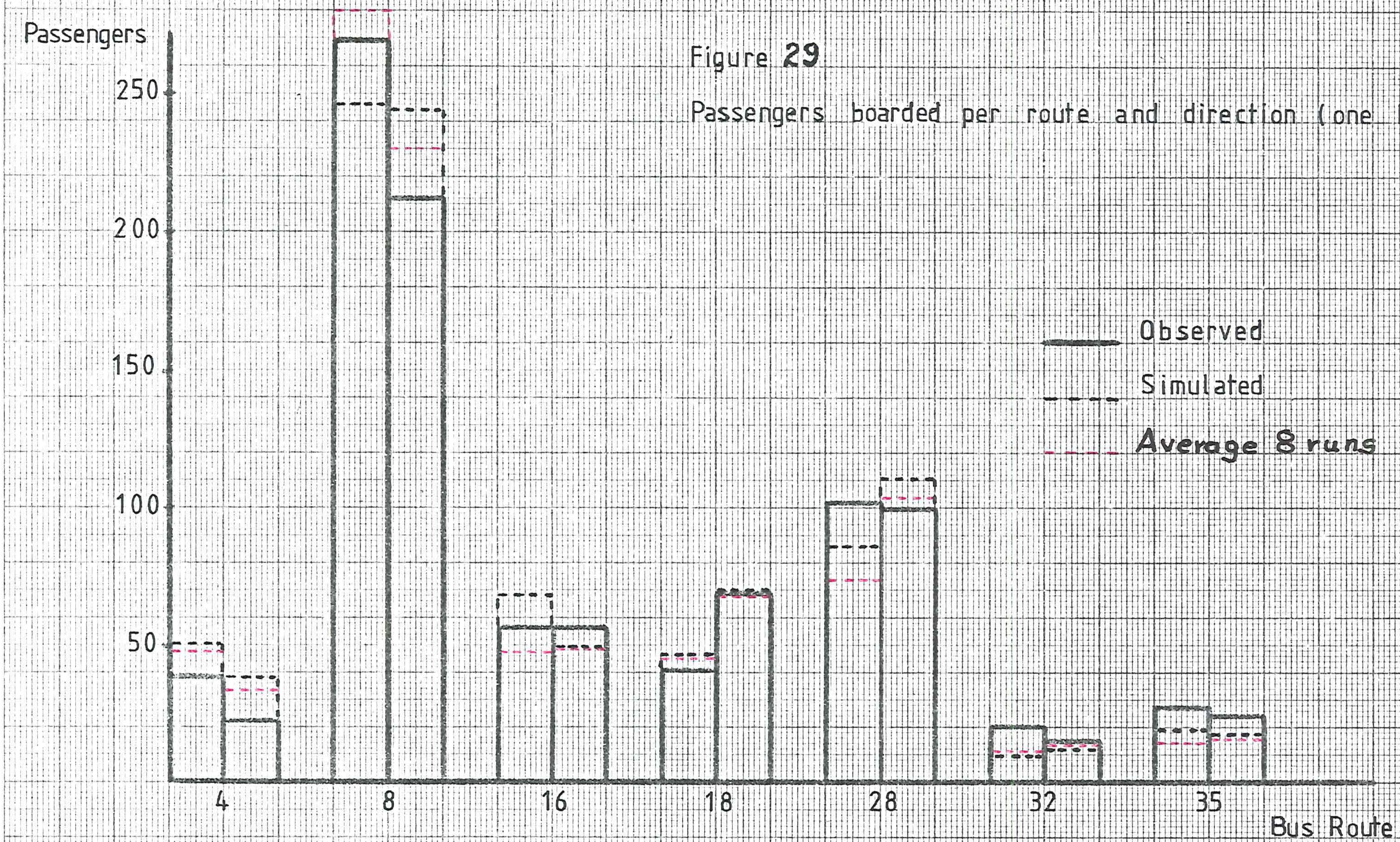
ii) Although the statistical results are not so satisfactory as might be desired, it is worth commenting that for the computation of the Chi-square test, the routes mentioned above (18% of the total) are to blame for 55% of its value.

iii) Taking into account the fact that SIMULA has to work with outputs produced by other submodels and that they did not completely correspond to the observed data mainly for reasons discussed above, it is then possible to conclude that in general the results are very encouraging.

Bearing in mind that one of the main objectives of the simulation submodel is to gather statistics and to organize them

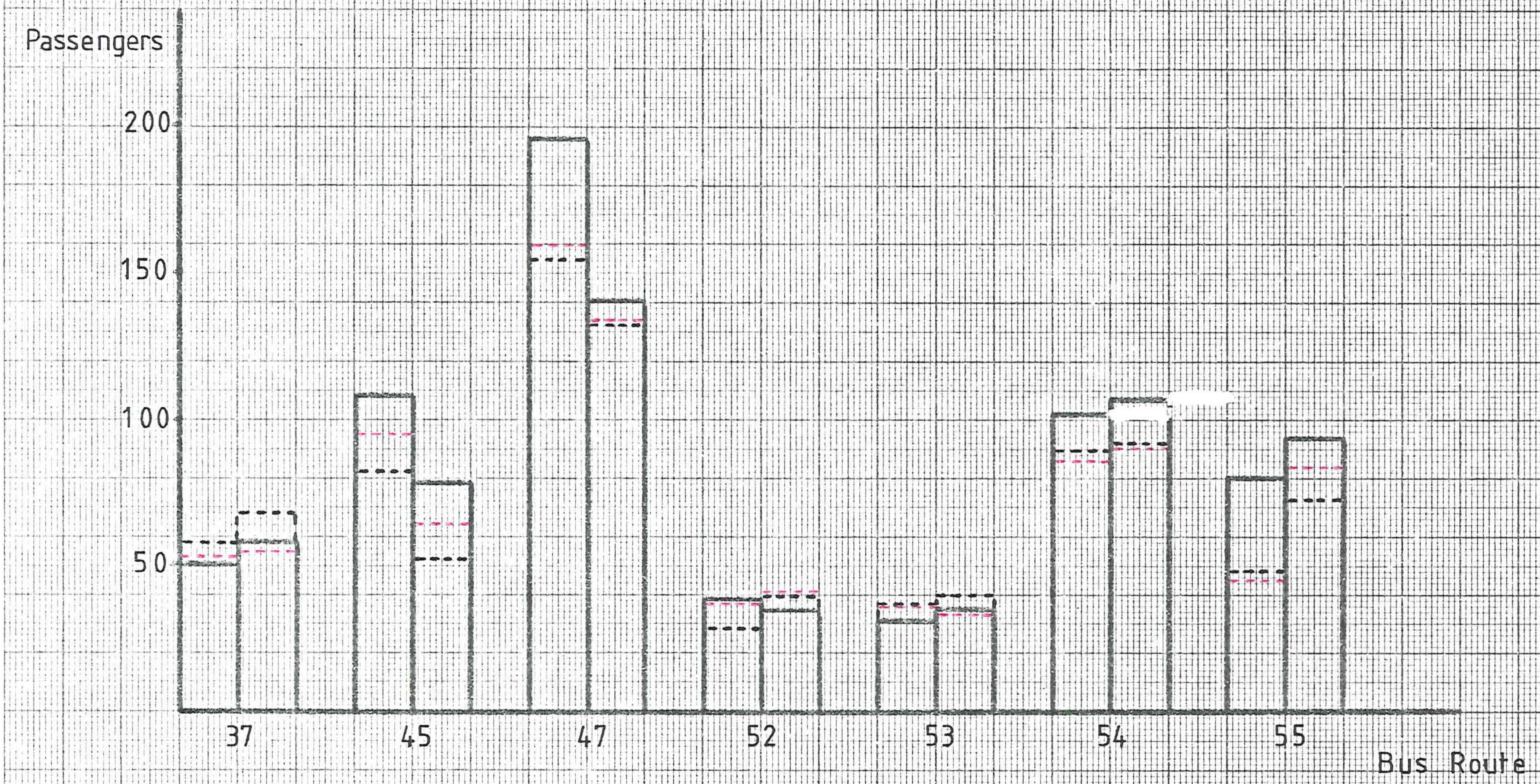
Figure 29

Passengers boarded per route and direction (one hour)+



+ For each route direction 1 on the left and direction 2 on the right

Figure 29 (continued)



in a suitable way to help planners in policy making, then SIMULA also produces additional information useful for assessing the performance of the public system, such as:

i) Average occupancy of buses between stops per route and direction, and summary per direction and for all transport system.

ii) Average waiting times per bus stop per route and direction, and summary per direction and for all transport system.

iii) Passengers not boarded due to capacity constraint per bus stop per route and direction, and summary per direction and for all transport system.

iv) Average time between bus arrival at terminal and next schedule departure per terminal, and summary for all transport system.

Figures 30 and 31 present for instance the bus occupancy between stops for route 8 directions 1 and 2 respectively compared to the observed data. As for the observed values total capacity between bus stops was unknown then the average capacity was used to estimate the percentage occupancy and therefore the results are not strictly comparable. However, it is possible to observe from these figures that specially the eastwards direction follows a very acceptable tendency.

There was no information available to compare additional results obtained by the simulation; however, it is worth showing a sample of tables produced by SIMULA. Table 32, for instance, illustrates inter-stop information for route 8 in both directions after one hour of simulation. Position, passengers on board and alighting matrix are displayed for each bus of the route when the

Figure 30

Occupancy/Capacity between stops, Route 8 - Eastwards

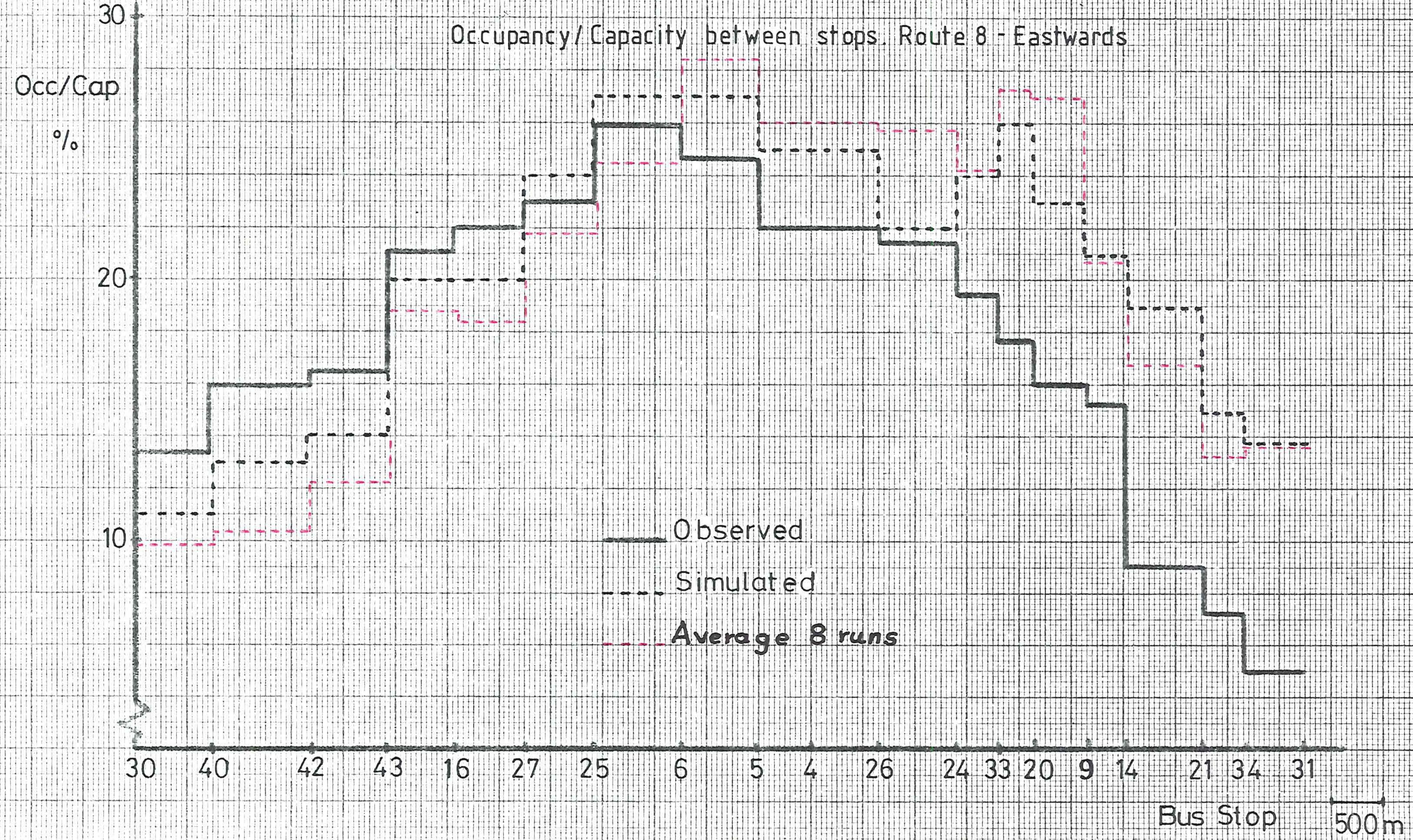


Figure 31

Occupancy/Capacity between stops

Route 8 - Westwards

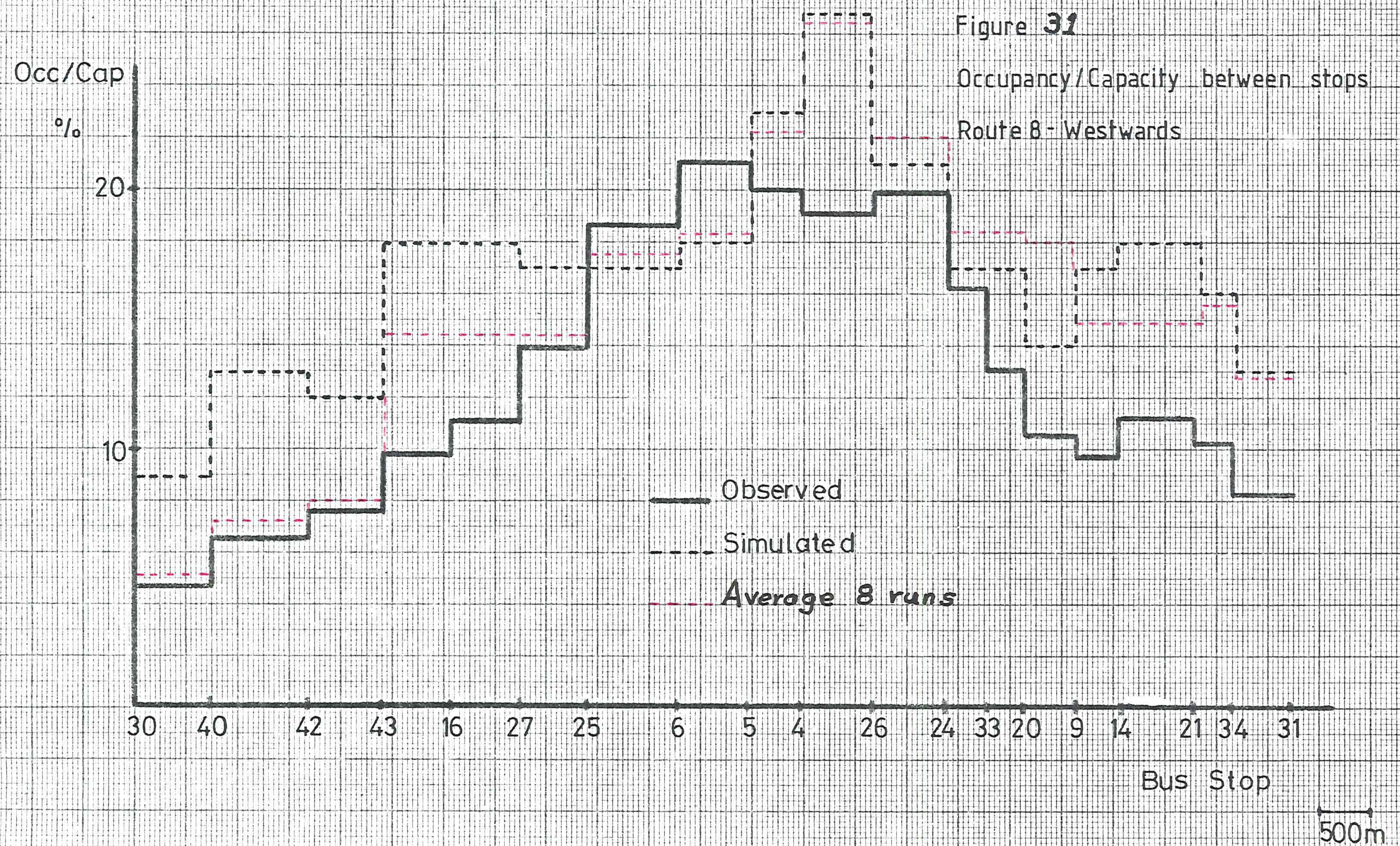


TABLE 32. <u>Inter-stop information for route 8.</u>

BUS	BC	PB	PO	BPO	D	TN	30	40	42	43	16	27	25	6	5	4	26	24	33	20	9	14	21	34	31
1	76	10	8	40	1	2	0	0	0	2	1	1	0	0	1	1	0	0	0	0	0	0	0	0	5
2	76	3	1	30	7	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
3	76	19	2	14	2	1	3	2	0	7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6
4	76	12	3	26	2	1	1	0	0	0	0	1	0	1	4	5	0	0	0	0	0	0	0	0	0
5	76	12	4	34	2	2	0	0	0	1	0	0	0	1	1	4	0	0	1	0	1	1	2	0	0
6	76	0	5	31	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	76	24	6	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	4	3	0	10
8	76	19	7	25	1	1	0	0	0	0	0	0	0	3	4	1	2	0	0	1	0	1	3	0	4

***DIRECTION : 1 *** →	30	40	42	43	16	27	25	6	5	4	26	24	33	20	9	14	21	34	31
OCCUPANCY / CAPACITY		11	13	14	26	20	24	27	27	25	25	22	24	26	23	21	19	15	14
AVE.WAITING T.(SEC)		120	264	200	252	260	303	332	365	218	229	0	170	268	354	327	0	449	0
PASSENGERS BOARDING		32	13	8	28	7	35	13	13	14	27	0	13	19	2	20	0	2	0
PASSENGERS ALIGHTING		0	5	2	6	14	0	17	25	28	5	3	8	10	30	11	22	2	57
PASSENGERS NOT BOARDED		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

***DIRECTION : 2 *** ←	30	40	42	43	16	27	25	6	5	4	26	24	33	20	9	14	21	34	31
OCCUPANCY / CAPACITY		6	13	12	12	14	17	17	16	23	27	21	17	17	14	17	18	16	13
AVE.WAITING T.(SEC)		16	227	133	122	119	122	201	262	356	300	222	236	442	160	236	292	197	337
PASSENGERS BOARDING		2	4	4	8	24	5	10	13	20	17	22	5	16	8	3	10	13	61
PASSENGERS ALIGHTING		23	14	2	32	7	16	0	17	28	41	0	3	5	6	21	7	6	0
PASSENGERS NOT BOARDED		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

<u>Abbreviations</u>: BC=bus capacity ; PB=passengers on board ; PR=previous bus (to control overtaking)
BPO=bus position ; D=direction; TN=trip number

the simulation was halted. Table 33 shows the summary per route and direction for the whole bus system.

TABLE 33. <u>Summary per route and direction and for the whole bus system.</u>

ROUTE	DIR/TER	OCC.(%)	AWT(MIN)	TTT(MIN)	S-A(MIN)	P.BOARD	N.BOARD
1	1	4.77	4.20	9.23	1.64	51	0
1	2	3.43	5.05	6.87	10.95	39	0
2	1	20.72	4.94	17.13	9.68	246	0
2	2	17.01	4.43	13.63	3.26	245	0
3	1	15.94	5.03	13.11	12.33	68	0
3	2	15.37	3.40	6.89	10.62	49	0
4	1	3.60	3.95	7.36	5.14	47	0
4	2	5.05	3.20	7.02	5.43	71	0
5	1	7.71	2.72	7.69	8.41	66	0
5	2	9.21	3.27	6.97	8.47	111	0
6	1	6.20	2.68	10.57	4.25	[illegible]	0
6	2	4.81	2.30	6.08	13.96	11	0
7	1	2.58	2.12	6.74	3.15	19	0
7	2	2.58	3.12	7.67	15.90	17	0
8	1	5.26	2.02	5.20	1.28	53	0
8	2	6.19	2.56	6.74	7.91	63	0
9	1	6.71	2.24	6.70	6.42	62	0
9	2	4.01	3.01	6.72	5.20	52	0
10	1	21.55	3.13	6.05	3.44	154	0
10	2	15.46	3.06	6.43	10.10	132	0
11	1	4.42	2.67	12.74	8.01	27	0
11	2	6.81	3.15	6.11	7.27	40	0
12	1	11.26	2.73	5.82	6.27	37	0
12	2	3.21	2.14	4.41	7.36	40	0
13	1	12.39	3.56	8.33	5.91	89	0
13	2	11.62	3.20	3.29	6.66	92	0
14	1	3.60	5.20	6.44	9.93	47	0
14	2	7.67	4.41	9.79	8.65	72	0
FOR SYSTEM		7.75	3.57	9.92	7.09	2058	0

Source: Calculations made by the author.

<u>Abbreviations</u>: DIR=direction ; TER= terminal ; OCC= occupancy ; AWT=average waiting time ; TTT=total travel time; S-A=difference between expected and actual departure time ; P.BOARD=passengers boarded; N.BOARD= passengers not boarded

In summary, taking into account

i) the complexity of the proposed public transport assignment model which implies the use of several variables interacting in the different submodels;

ii) the possibility of deficiencies in the information used;

iii) the need to borrow from a wide variety of sources because of the unavailability of data for the Springburn area and;

iv) the time limitation which prevented study of the viability of using different parameters compatible with the Study Area, or even to modify the bus stop location and the aggregation of zones.

Then it is possible to conclude that the results of the application of the model to the data set gathered are very encouraging as they are in general meaningful and consistent with the prototype which they are intended to simulate.

The computing time for the application of the model in the Springburn area was approximately 4.5 minutes of Order Code Processor (OCP). Of this time ROUTE-2 used almost 50%.

Subsequently, it was decided that some additional tests were necessary to complete this initial validation process. The details are recorded overleaf.

For a given set of parameters the sampling error in a simulation model of this type can be reduced in one of two ways (Naylor et al, 1968): either the run length can be increased or repeated runs of a given length can be made using different sets of random numbers.

It was decided to use the latter method to extend the validation tests here by running the simulation model (ARRIVE-SORTING-SIMULA) an additional seven times. Sample averages of the eight runs are included in Figures 29,30 and 31. The Chi-square was recalculated using the average of these runs and the value obtained was 69.5 which was still not completely satisfactory but much lower than the one achieved before. The coefficient of variance was 18.6% , also lower. The following Table shows the variability of the results of the eight runs.

Summary for the whole bus system.

Run	OCC(%)	AWT(min)	TTT(min)	S-A(min)	P.boarded
1	7.75	3.57	9.92	7.09	2058
2	7.77	3.63	9.56	7.27	1991
3	7.66	3.74	9.77	7.37	1964
4	7.70	3.61	9.37	7.28	2029
5	7.71	3.63	9.43	7.09	2040
6	7.72	3.64	9.88	7.07	2044
7	7.83	3.71	9.81	7.04	2034
8	7.72	3.61	9.67	7.00	2066
AVERAGE	7.73	3.64	9.68	7.15	2028

This result was considered to be particularly satisfactory. Had the results shown greater variation it would not have been possible without further work to detect which part of the whole model was faulty. These results, however,do give confidence that no serious flaws exist.

NOTES TO CHAPTER III.

1) Only services 174, 177, 172 (inward) and 182 (outward) have no restrictions.

2) Overgaard, K.R.? Traffic estimation in urban transportation planning. Acta Polytechnica Scandinavica, 1966, Copenhagen.

3) GGPTE, Survey Study, volume 1, 1975.

4) Numerical Algorithms Group (NAG) Library is an application software already mounted on the ICL-2976 of the University of Glasgow. It consists in a collection of algorithms for the solution of numerical problems on computers.

5) For a long period of time (i.e. 12 hours) the total number of people carried tends to be equal to the total number of passengers boarded.

CHAPTER IV

SUMMARY AND CONCLUSIONS

A public transport system affects different sectors of the community and each sector regards the system in a different way according to its own interests. Transport planners have considered generalised cost to be a good measure of the way people perceive the level of service provided by the system as it embraces cost in time and money of the trip to be made. Any measure which reduces generalised cost can be regarded, therefore, as a potential improvement of the system from the users' point of view.

The bus operator in control of his resources can combine them in such a way that specific objectives are met, taking into account that the system operates on a balance of costs and service and that for a fixed amount of input improvements in some aspects may affect other sectors of the operation adversely unless there is an increase in productivity or ridership.

The use which the operator makes of his resources cannot be guaranteed to produce a good service as the influence of some external factor may affect the reliability of the system, such as the pattern of arrival of passengers at bus stops. Although there may be no way of avoiding such influences the operator may try to allow for them by the use of extra inputs.

The interdependent action of the different elements involved in the operation of bus services suggests the need for analysing the system as a whole. Within this framework, the purpose of this study was to develop a model capable of evaluating the operation of the entire system and of giving useful information to planners and operators in their task of improving the productivity of the service and of spreading benefits to the users.

Given a specific level of demand, the proposed model can be used to evaluate a public transport system in such a way that service alterations can quickly and accurately be tested before decisions are taken. The approach used was to assign the transport demand to the existing transport facilities and in doing so three different stages were considered: i) determination of individual options to make a trip; ii) process of loading passengers onto selected options; and iii) process of loading passengers onto routes and buses. The model was composed of several linked submodels and a major goal was the formulation of a model which was capable of yielding an accurate description of the system while minimising computing time.

The process of determining options to make a trip relied basically on matrix network analysis as it was assumed that people chose their routes so as to minimise their total travel time. The process of loading passengers onto routes and buses was accomplished by using simulation techniques which have proved to be very useful in the analysis of complex queueing processes and interactions between them.

The basic data for the model consisted of:

i) Bus stop spacing and inter-node distances; average and maximum speeds; average walking speed;

ii) Bus frequencies in seconds;

iii) Bus route description in which each route was given as a string of consecutive bus stops with two termini;

iv) O-D matrix that constituted the transport demand within the area of study and which was assumed to be concentrated either at nodes or at bus stops;

v) Random number generator;

vi) Set of parameters such as:

- scale factors for waiting and walking times;
- slope and independent term of the linear fare function;
- extra penalty for transferring;
- number of nodes in the network;
- percentage parameter which defined the range of options to be recorded;
- power for assignment function;
- length of simulation in time and unit time of O-D matrix;
- bus capacities;
- number of buses per route;
- layover time at termini;
- correlation factor for inter-stop travel variation.

The main outputs of the model consisted of:

i) Minimum distances and path description between every pair of nodes in the network.

ii) Set of options to travel by public transport between every pair of nodes in the network and number of people taking every option.

iii) Walking links which were used with more frequency by passengers and the origin-destinations that involved more walking according to the registered number of trips.

iv) Statistics collected at three different levels: stop (or inter-stop), route and entire transport system:

- average bus occupancy in percentage;
- average waiting time;
- total travel time;
- number of passengers boarding and alighting;
- passengers left behind because of seat unavailability when bus arrived at bus stop;
- average differences between expected and actual departure time at termini.

Some characteristics and advantages of the submodels which have been developed are particularly noteworthy, and it is worth noting their differences from other models which accomplish similar tasks. Submodel PATH, for instance, constitutes an improvement upon the the traditional matrix approach for obtaining minimum paths in a network in the sense that is faster and can register the sequence of nodes in a convenient and practical way for retrieval. This submodel was the basis of ROUTE-1, a first version of a matrix algorithm dealing with minimum paths in public transport networks and which in turn was the basis of a more sophisticated model, ROUTE-2.

Submodel ROUTE-2 constitutes a simple though effective and precise way of obtaining public transport options to make a trip between any two pair of nodes in a network. This submodel, like some other non-matrix algorithms of this kind, works on the basis that total travel time is an adequate measure of travel resistance although the generalised cost function is more complex since it also involves walking time, fare and extra penalties for transferring. However, it is worth noting that the fact that people perceive costs in a different way could be regarded as a limitation of models which use average generalised costs.

By contrast with other algorithms, this submodel does not require dummy links to be coded in order to provide the respective level of service of the walking mode. ROUTE-2 works on the basis that this mode not only competes with the transport system but is also its natural feeder, so that this submodel can obtain options which include the walking mode by using the matrix of minimum walking distances.

The fact that matrices can keep information in a concise and convenient way make them very suitable for mathematical and computing manipulation. Some other important features of ROUTE-2 depend on this characteristic:

i) By running the computer program once it is possible to obtain the minimum paths between any pair of points in a network. This characteristic may be very useful in applications similar to the one made in this study; however, when paths for only a few pair of points are required other approaches could be more useful and practical.

ii) Apart from the optimum option to make a trip, this submodel is also capable of registering alternative paths, where these are feasible. Therefore, ROUTE-2 constitutes an improvement upon the other algorithms presented in this study which have to work on the basis of anall-or-nothing assignment.

iii) The process of retrieving information is simple and works on the basis that if a trip from A to B requires a transfer at C, then the minimum path from A to C constitutes the first part of the trip.

With respect to the kind of options which this submodel is capable of recording, it may be worth emphasising two important aspects:

1) ROUTE-2 selects options with the lowest total travel time but this does not mean that all recorded options to make a trip may be found reasonable from the passenger's point of view. In fact some options could even be ludicrous. However, this is not to be considered a disadvantage of the submodel though criteria to eliminate undesirable options should be established after a more detail study of this topic and the proportion of walking, riding, waiting time and fare equivalent in total travel time.

2) ROUTE-2 registers all inspected paths falling in the range defined by the optimum total travel time and a percentage figure given in the input process which denotes the acceptable range. Though the definition of this range may look rather simple its specification requires a good knowledge of the travel pattern in the area of study. The higher the value of that figure the greater the probability of registering options which include backtracking; therefore, a careful examination of this aspect may be advisable in future studies.

Having obtained the different options for making a trip between two nodes, submodel ASSIGN assumes that the higher the total travel time of the option the less the number of people taking it. The fundamental expression of this submodel includes a constant parameter $\mathbf{m}$, which reflects passenger sensitivity. In mathematical terms, the higher the value of $\mathbf{m}$ the more the concentration of passengers onto lower travel time options, in other words the less the importance of higher travel time options. Given the relevance of this parameter in the assignment process it is important to determine it in a precise way.

Two different simulation model developed for general studies of the operation of bus route networks in urban areas were presented in the first chapter of this study. Although there are many similarities between those models and the simulation presented here, there are also important differences, principally the representation of passengers and the way with which queues at bus stops are dealt.

The simulation model developed by Jackson (1972) obtains the number of arrivals at a bus stop for a small interval of time dt sampling from a Poisson distribution, and these numbers are summed in order to get a sample over a particular period. This model considers only direct trips and the O-D information for pair of stops with direct service is given as p_{IJ}, the probability that a passenger boarding at stop I will alight at stop J. Then, the number of arrivals is distributed among the potential destinations using the probabilities p_{IJ}. Passengers are classified into types according to their destinations and when a bus cannot accommodate all the people wishing to board at a stop, then those to be carried for each destination are selected according to a probability proportional to the time since the last arrival of a passenger who was able to board a bus for that destination.This means that this model takes no strict account of the order of passengers in the queue.

In the model developed by Gerrard and Brook (1972), passenger inter-arrival times between arrival times of two consecutive buses are generated according to a Poisson distribution. The number of passengers generated is equal to a new queue segment. Passengers are classified by passenger type according to the route choice available at the stop in question. As a queue may be formed by one or more queue segments, when a bus arrives at a stop and there are not enough places for people waiting to board, then each segment is processed in sequence. Passenger arrival times are not recorded within each segment. The number of passengers alighting at each stop is generated using a binomial distribution.

The proposed model here deals with the aspects mentioned above by means of submodels ARRIVE and SIMULA. The first one works with the exponential distribution in order to generate the exact times of arrival of passengers at bus stops. To do so this submodel assumes that if there are P people wishing to travel from I to J during the period of time T, then the mean rate of arrival per unit of time λ is equal to P/T. Then, by using the stop-to-stop O-D matrix and the respective bus routes to make direct trips a file which stores information about every passenger is created where each register contains his origin, destination, time of arrival and bus services for the journey. This file is classified by stop, origin and arrival time in order to be input to submodel SIMULA.

When a bus arrives at a bus stop and can accommodate more people, then the created file is read in order to board the passengers that have arrived before the bus departure and that were waiting for that particular route service. People are boarded on the basis of first come first served but taking into consideration bus seat availability. Given the destination of the people boarded, the array which keeps information on the number of passengers alighting at each bus stop is updated accordingly. In this way, these two submodels, in contrast with the two simulation models developed elsewhere, consider passengers individually until they get on the bus which makes easy to work out waiting and total travel times.

Apart from the advantages of using simulation in the analysis of complex systems, it is worth emphasising that SIMULA works with outputs created by other submodels which means that in some sense this submodel also serves for the evaluation of the previous processes. This could be particularly useful when there is no data to test or validate the results of the other submodels.

At this point it is worth reconsidering two important exercises on public transport planning which were presented in the first chapter in order to highlight the general approach of these models with the one developed in this study. Other important models also reviewed fall by and large within the scope of these two exercises.

Lampkin and Saalmans (1967) developed a heuristic algorithm to design a bus network system which minimised total travel time to the passengers, taking into account restrictions of financial order. TRANSEPT on the other hand, was developed by the Local Government Operational Research Unit for evaluation of bus route networks. This complex model can take into account changes in public transport demand and in the long term can evaluate changes in the road network and in the land use. It includes a modal split process between car, walking mode and public transport. This model assumes that people travel between two points using the path with minimum generalised cost and passengers are assigned to routes and buses by means of a dynamic relationship which tries to find a stable solution.

With respect to the first model, the proposed model differs fundamentally in the sense that it is not an optimisation process. The proposed model was designed for the evaluation of alternatives to the existing system in order to recommend minor changes or even the implementation of a new design. Therefore, the model presented in this study falls much more within the scope of TRANSEPT although they differ basically in three main points. First there is, of course, the wider aspects covered by TRANSEPT; secondly there is the algorithm developed here for obtaining minimum paths in public transport networks; and thirdly there is the use of simulation techniques in the process of loading passengers onto routes and buses, whereas TRANSEPT uses analytical techniques.

Thus, it may be concluded that the proposed model constitutes a different approach to the analysis and evaluation of public transport systems in urban areas. Its basic data are input to the model in a simple way and the results obtained, which tend to cover the main aspects of the operation, are presented in a suitable manner to the analyst. It is clear, however, that some of the information needed may prove to be difficult to collect as sometimes may happen with such models. This is the case, for instance, of the door-to-door O-D matrix and some parameters which were mentioned earlier in this chapter.

This study has been concerned mainly with bus systems in urban areas; however, it should be emphasised that the proposed model would also allow consideration of other modes of public transport such as tramcars, underground and the like. It is important to note, however, that one of the main assumptions here concerns the independence of vehicles and passengers arrivals at boarding points and this may be a reason not to consider some kind of urban trains within the analysis.

The set of submodels designed in the present study was tested in a real situation using data gathered in a suburb of Glasgow. The results obtained may be considered to be rather satisfactory taking into account some possible limitations in the data used. However, those results suggest the need for further work in order to reach a more extensive validation of the model. This might well be complemented with further analysis of some aspects of the planning process and collection of further data on specific items. Some examples of these topics were suggested earlier in this chapter; in particular the way people use the transport system in order to make a trip and how they behave when they have several options to choose from require further study.

Finally, it is evident that every person tends to conceptualise a given problem in a different way which explains much of the difference between this and previous exercises, and hopefully this study may be regarded as a new contribution to knowledge in public transport planning. However, since model building forces the analyst to organise, evaluate and examine the validity of his thoughts, perhaps the most important result of this research is its effect on the author who now realises that he may hardly have begun to explore all the possibilities.

APPENDIX A. <u>Network example: PATH</u>.

INPUT DATA

MATRIX OF DISTANCES D (IN MINUTES)

	1	2	3	4	5	6	7	8	9
1	9999	6	9999	7	9999	9999	9999	9999	9999
2	4	9999	8	9999	9	9999	9999	9999	9999
3	9999	5	9999	9999	9999	10	9999	9999	9999
4	9	9999	9999	9999	5	9999	8	9999	9999
5	9999	10	9999	6	9999	11	9999	4	9999
6	9999	9999	7	9999	9	9999	9999	9999	7
7	9999	9999	9999	6	9999	9999	9999	4	9999
8	9999	9999	9999	9999	5	9999	3	9999	5
9	9999	9999	9999	9999	9999	9	9999	2	9999

ITERATION NO. 1

APPENDIX A - 2

MATRIX OF DISTANCES D (IN MINUTES)

	1	2	3	4	5	6	7	8	9
1	0	6	14	7	12	23	15	16	21
2	4	0	8	11	9	18	19	13	18
3	9	5	0	16	14	10	24	18	17
4	9	15	23	0	5	16	8	9	14
5	14	10	18	6	0	11	7	4	9
6	16	12	7	15	9	0	16	9	7
7	15	21	29	6	9	20	0	4	9
8	18	15	23	9	5	14	3	0	5
9	20	17	16	11	7	9	5	2	0

MATRIX OF NODE SEQUENCE REGISTRATION MT

	1	2	3	4	5	6	7	8	9
1	0	0	2	0	4	5	4	5	8
2	0	0	0	1	0	3	4	5	8
3	2	0	0	1	2	0	4	5	6
4	0	1	2	0	0	5	0	5	8
5	2	0	2	0	0	0	8	0	8
6	2	3	0	5	0	0	8	9	0
7	4	1	2	0	8	5	0	0	8
8	4	5	2	7	0	9	0	0	0
9	4	5	6	7	8	0	8	0	0

TOTAL DISTANCE : 888 MINUTES

ITERATION NO. 2

APPENDIX A - 3

MATRIX OF DISTANCES D (IN MINUTES)

	1	2	3	4	5	6	7	8	9
1	0	6	14	7	12	23	15	16	21
2	4	0	8	11	9	18	16	13	18
3	9	5	0	16	14	10	21	18	17
4	9	15	23	0	5	16	8	9	14
5	14	10	18	6	0	11	7	4	9
6	16	12	7	15	9	0	12	9	7
7	15	19	25	6	9	18	0	4	9
8	18	15	21	9	5	14	3	0	5
9	20	17	16	11	7	9	5	2	0

MATRIX OF NODE SEQUENCE REGISTRATION MT

	1	2	3	4	5	6	7	8	9
1	0	0	2	0	4	5	4	5	8
2	0	0	0	1	0	3	8	5	8
3	2	0	0	1	2	0	8	5	6
4	0	1	2	0	0	5	0	5	8
5	2	0	2	0	0	0	8	0	8
6	2	3	0	5	0	0	8	9	0
7	4	5	6	0	8	9	0	0	8
8	4	5	6	7	0	9	0	0	0
9	4	5	6	7	8	0	8	0	0

TOTAL DISTANCE : 868 MINUTES

MINIMUM PATHS-FINAL RESULTS APPENDIX A - 4

DIST	PATH DESCRIPTION				
6	1	2			
14	1	2	3		
7	1	4			
12	1	4	5		
23	1	4	5	6	
15	1	4	7		
16	1	4	5	8	
21	1	4	5	8	9
4	2	1			
8	2	3			
11	2	1	4		
9	2	5			
18	2	3	6		
16	2	5	8	7	
13	2	5	8		
18	2	5	8	9	
9	3	2	1		
5	3	2			
16	3	2	1	4	
14	3	2	5		
10	3	6			
21	3	2	5	8	7
18	3	2	5	8	
17	3	6	9		
9	4	1			
15	4	1	2		
23	4	1	2	3	
5	4	5			
16	4	5	6		
8	4	7			
9	4	5	8		
14	4	5	8	9	
14	5	2	1		
10	5	2			
18	5	2	3		
6	5	4			
11	5	6			
7	5	8	7		
4	5	8			
9	5	8	9		
16	6	3	2	1	
12	6	3	2		
7	6	3			

MINIMUM PATHS-FINAL RESULTS APPENDIX A - 5

DIST	PATH DESCRIPTION				
15	6	5	4		
9	6	5			
12	6	9	8	7	
9	6	9	8		
7	6	9			
15	7	4	1		
19	7	8	5	2	
25	7	8	9	6	3
6	7	4			
9	7	8	5		
18	7	8	9	6	
4	7	8			
9	7	8	9		
18	8	7	4	1	
15	8	5	2		
21	8	9	6	3	
9	8	7	4		
5	8	5			
14	8	9	6		
3	8	7			
5	8	9			
20	9	8	7	4	1
17	9	8	5	2	
16	9	6	3		
11	9	8	7	4	
7	9	8	5		
9	9	6			
5	9	8	7		
2	9	8			

APPENDIX B. Network example: ROUTE-1.

INPUT DATA

APPENDIX B - 1

MATRIX OF DISTANCES D (IN MINUTES)

0.0	6.0	999.0	7.0	999.0	999.0	999.0	999.0	999.0
4.0	0.0	8.0	999.0	9.0	999.0	999.0	999.0	999.0
999.0	5.0	0.0	999.0	999.0	10.0	999.0	999.0	999.0
9.0	999.0	999.0	0.0	5.0	999.0	8.0	999.0	999.0
999.0	10.0	999.0	6.0	0.0	11.0	999.0	4.0	999.0
999.0	999.0	7.0	999.0	9.0	0.0	999.0	999.0	7.0
999.0	999.0	999.0	6.0	999.0	999.0	0.0	4.0	999.0
999.0	999.0	999.0	999.0	5.0	999.0	3.0	0.0	5.0
999.0	999.0	999.0	999.0	999.0	9.0	999.0	2.0	0.0

BUS FREQUENCIES FR (IN MINUTES)

10.0 12.0 8.0 15.0 10.0 8.0

MATRIX OF ROUTES MR

0	0	0	1	2	0	7	0	0	-2	0	0	7	0	0	7	0	0	7	0	0	7	0	0	7	0	0
-2	0	0	0	0	0	-6	1	2	7	0	0	-4	6	0	7	0	0	7	0	0	7	0	0	7	0	0
7	0	0	-2	6	0	0	0	0	7	0	0	7	0	0	1	2	0	7	0	0	7	0	0	7	0	0
1	2	0	7	0	0	7	0	0	0	0	0	3	4	0	7	0	0	-4	0	0	7	0	0	7	0	0
7	0	0	-6	4	0	7	0	0	-4	-3	0	0	0	0	-5	3	0	7	0	0	5	6	0	7	0	0
7	0	0	7	0	0	-2	0	0	7	0	0	-3	5	0	0	0	0	7	0	0	7	0	0	-5	1	3
7	0	0	7	0	0	7	0	0	1	4	0	7	0	0	7	0	0	0	0	0	-6	0	0	7	0	0
7	0	0	7	0	0	7	0	0	7	0	0	-6	-5	0	7	0	0	1	6	0	0	0	0	-3	0	0
7	0	0	7	0	0	7	0	0	7	0	0	7	0	0	-3	5	0	7	0	0	1	3	0	0	0	0

*NUSTO : 9 *NUROU : 6 *NUROC : 3

ITERATION NO. 1 — APPENDIX B - 2

MATRIX OF DISTANCES D (IN MINUTES)

0.0	8.7	16.7	13.0	20.3	26.7	28.5	25.7	34.7
10.0	0.0	9.6	17.0	11.6	20.7	32.5	17.0	26.0
15.0	7.4	0.0	22.0	18.0	12.7	37.5	22.0	22.0
11.7	17.7	25.7	0.0	7.6	20.0	15.5	13.8	22.8
20.3	12.6	22.0	8.6	0.0	13.2	11.0	6.2	15.2
22.0	18.0	13.0	19.0	11.2	0.0	22.2	11.2	8.5
20.0	26.0	34.0	9.0	13.0	26.2	0.0	8.0	17.0
23.0	19.0	27.0	14.0	7.2	18.0	5.2	0.0	9.0
25.0	23.2	24.2	16.0	11.4	11.2	10.0	4.2	0.0

MATRIX OF ROUTES MR

0	0	0	1	2	0	1	2	0	-2	0	0	-4	6	0	1	2	0	-4	0	0	6	0	0	-3	0	0
-2	0	0	0	0	0	-6	1	2	-2	0	0	-4	6	0	1	2	0	-4	0	0	6	0	0	-3	0	0
-2	0	0	-2	6	0	0	0	0	-2	0	0	6	0	0	1	2	0	-4	0	0	6	0	0	1	0	0
1	2	0	1	2	0	1	2	0	0	0	0	3	4	0	3	0	0	-4	0	0	5	6	0	-3	0	0
1	2	0	-6	4	0	-6	0	0	-4	-3	0	0	0	0	-5	3	0	6	0	0	5	6	0	-3	0	0
-2	0	0	-2	0	0	-2	0	0	-3	0	0	-3	5	0	0	0	0	6	0	0	1	3	0	-5	1	3
1	0	0	1	0	0	1	0	0	1	4	0	-6	0	0	-5	3	0	0	0	0	-6	0	0	-3	0	0
1	0	0	-6	0	0	-6	0	0	1	0	0	-6	-5	0	-3	0	0	1	6	0	0	0	0	-3	0	0
1	0	0	-6	0	0	-2	0	0	1	0	0	-6	-5	0	-3	5	0	1	0	0	1	3	0	0	0	0

MATRIX OF TRANSFERS MT

0	0	0	0	2	0	4	2	8
0	0	0	0	0	0	4	0	8
0	0	0	0	0	0	4	0	0
0	0	0	0	0	0	0	5	8
4	0	0	0	0	0	0	0	8
0	0	0	0	0	0	5	0	0
0	0	0	0	0	5	0	0	8
0	0	0	0	0	0	0	0	0
0	8	6	0	8	0	0	0	0

TOTAL DISTANCE : 1245.93 MINUTES

MATRIX OF DISTANCES D (IN MINUTES)

0.0	8.7	16.7	13.0	20.3	26.7	28.5	25.7	34.7
10.0	0.0	9.6	17.0	11.6	20.7	20.0	17.0	26.0
15.0	7.4	0.0	22.0	18.0	12.7	25.0	22.0	22.0
11.7	17.7	25.7	0.0	7.6	20.0	15.5	13.8	22.8
20.3	12.6	22.0	8.6	0.0	13.2	11.0	6.2	15.2
22.0	18.0	13.0	19.0	11.2	0.0	17.0	11.2	8.5
20.0	23.0	31.0	9.0	13.0	26.0	0.0	8.0	17.0
23.0	19.0	27.0	14.0	7.2	18.0	5.2	0.0	9.0
25.0	23.2	24.2	16.0	11.4	11.2	10.0	4.2	0.0

MATRIX OF ROUTES MR

0	0	0	1	2	0	1	2	0	-2	0	0	-4	6	0	1	2	0	-4	0	0	6	0	0	-3	0	0
-2	0	0	0	0	0	-6	1	2	-2	0	0	-4	6	0	1	2	0	6	0	0	6	0	0	-3	0	0
-2	0	0	-2	6	0	0	0	0	-2	0	0	6	0	0	1	2	0	6	0	0	6	0	0	1	0	0
1	2	0	1	2	0	1	2	0	0	0	0	3	4	0	3	0	0	-4	0	0	5	6	0	-3	0	0
1	2	0	-6	4	0	-6	0	0	-4	-3	0	0	0	0	-5	3	0	6	0	0	5	6	0	-3	0	0
-2	0	0	-2	0	0	-2	0	0	-3	0	0	-3	5	0	0	0	0	1	0	0	1	3	0	-5	1	3
1	0	0	-6	0	0	-6	0	0	1	4	0	-6	0	0	-3	0	0	0	0	0	-6	0	0	-3	0	0
1	0	0	-6	0	0	-6	0	0	1	0	0	-6	-5	0	-3	0	0	1	6	0	0	0	0	-3	0	0
1	0	0	-6	0	0	-2	0	0	1	0	0	-6	-5	0	-3	5	0	1	0	0	1	3	0	0	0	0

MATRIX OF TRANSFERS MT

0	0	0	0	2	0	4	2	8
0	0	0	0	0	0	0	0	8
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	5	8
4	0	0	0	0	0	0	0	8
0	0	0	0	0	0	0	0	0
0	0	0	0	0	8	0	0	8
0	0	0	0	0	0	0	0	0
0	8	6	0	8	0	0	0	0

TOTAL DISTANCE : 1209.49 MINUTES

ROUTE1 : FINAL RESULTS APPENDIX B - 4

DIST(MIN)	PATH DESCRIPTION												
8.7	1	1	2	0	2								
16.7	1	1	2	0	3								
13.0	1	-2	0	0	4								
20.3	1	1	2	0	2	-4	6	0	5				
26.7	1	1	2	0	6								
28.5	1	-2	0	0	4	-4	0	0	7				
25.7	1	1	2	0	2	6	0	0	8				
34.7	1	1	2	0	2	6	0	0	8	-3	0	0	9
10.0	2	-2	0	0	1								
9.6	2	-6	1	2	3								
17.0	2	-2	0	0	4								
11.6	2	-4	6	0	5								
20.7	2	1	2	0	6								
20.0	2	6	0	0	7								
17.0	2	6	0	0	8								
26.0	2	6	0	0	8	-3	0	0	9				
15.0	3	-2	0	0	1								
7.4	3	-2	6	0	2								
22.0	3	-2	0	0	4								
18.0	3	6	0	0	5								
12.7	3	1	2	0	6								
25.0	3	6	0	0	7								
22.0	3	6	0	0	8								
22.0	3	1	0	0	9								
11.7	4	1	2	0	1								

ROUTE1 : FINAL RESULTS APPENDIX B - 5

DIST(MIN)	PATH DESCRIPTION											
17.7	4	1	2	0	2							
25.7	4	1	2	0	3							
7.6	4	3	4	0	5							
20.0	4	3	0	0	6							
15.5	4	-4	0	0	7							
13.8	4	3	4	0	5	5	6	0	8			
22.8	4	3	4	0	5	5	6	0	8	-3	0	0	9
20.3	5	-4	-3	0	4	1	2	0	1			
12.6	5	-6	4	0	2							
22.0	5	-6	0	0	3							
8.6	5	-4	-3	0	4							
13.2	5	-5	3	0	6							
11.0	5	6	0	0	7							
6.2	5	5	6	0	8							
15.2	5	5	6	0	8	-3	0	0	9			
22.0	6	-2	0	0	1							
18.0	6	-2	0	0	2							
13.0	6	-2	0	0	3							
19.0	6	-3	0	0	4							
11.2	6	-3	5	0	5							
17.0	6	1	0	0	7							
11.2	6	1	3	0	8							
8.5	6	-5	1	3	9							
20.0	7	1	0	0	1							
23.0	7	-6	0	0	2							

ROUTE1 : FINAL RESULTS APPENDIX B - 6

DIST(MIN)	PATH DESCRIPTION							
31.0	7	-6	0	0	3			
9.0	7	1	4	0	4			
13.0	7	-6	0	0	5			
26.0	7	-6	0	0	8	-3	0	0	6
8.0	7	-6	0	0	8			
17.0	7	-6	0	0	8	-3	0	0	9
23.0	8	1	0	0	1			
19.0	8	-6	0	0	2			
27.0	8	-6	0	0	3			
14.0	8	1	0	0	4			
7.2	8	-6	-5	0	5			
18.0	8	-3	0	0	6			
5.2	8	1	6	0	7			
9.0	8	-3	0	0	9			
25.0	9	1	0	0	1			
23.2	9	1	3	0	8	-6	0	0	2
24.2	9	-3	5	0	6	-2	0	0	3
16.0	9	1	0	0	4			
11.4	9	1	3	0	8	-6	-5	0	5
11.2	9	-3	5	0	6			
10.0	9	1	0	0	7			
4.2	9	1	3	0	8			

APPENDIX C. Network example: ROUTE-2 & ASSIGN.

INPUT DATA APPENDIX C - 1

*WWK: 1.40 *A: 0.54 *B: 644 *C: 600 *DOPTIM: 0.20 *WWT: 2.25 *NUSTO: 9

MATRIX OF DISTANCES (IN SECONDS) WALKING

	1	2	3	4	5	6	7	8	9
1	0	2972	6936	3467	5945	11397	7432	7928	10406
2	1980	0	3963	5450	4458	8919	7928	6441	8919
3	4458	2476	0	7928	6936	4954	10406	8919	8423
4	4458	7432	11397	0	2476	7928	3963	4458	6936
5	6936	4954	8919	2972	0	5450	3467	1980	4458
6	7928	5945	3467	7432	4458	0	5945	4458	3467
7	7432	9414	12388	2972	4458	8919	0	1980	4458
8	8919	7432	10406	4458	2476	6936	1485	0	2476
9	9910	8423	7928	5450	3467	4458	2476	989	0

MATRIX OF DISTANCES (IN SECONDS) RIDING

	1	2	3	4	5	6	7	8	9
1	0	360	9999	420	9999	9999	9999	9999	9999
2	240	0	480	9999	540	9999	9999	9999	9999
3	9999	300	0	9999	9999	600	9999	9999	9999
4	540	9999	9999	0	300	9999	480	9999	9999
5	9999	600	9999	360	0	660	9999	240	9999
6	9999	9999	420	9999	540	0	9999	9999	420
7	9999	9999	9999	360	9999	9999	0	240	9999
8	9999	9999	9999	9999	300	9999	180	0	300
9	9999	9999	9999	9999	9999	540	9999	120	0

BUS ROUTES

R	FREQ(SEC)	1	2	3	4	5	6	7	8	9
1	600	1	2	3	6	9	8	7	4	1
2	720	4	1	2	3	6				
3	480	4	5	6	9	8				
4	900	7	4	5	2					
5	600	9	6	5	8					
6	480	3	2	5	8	7				

OPTION ASSIGNMENT PROCESS-FINAL RESULTS APPENDIX C - 2

Route-set classes

Route-set number	Routes	Route-set number	Routes
1	-6	15	-5 1 3
2	-6 -5	16	1 3
3	-5	17	1 4
4	-4	18	3 4
5	-3	19	-6 4
6	-4 -3	20	-3 5
7	-2	21	5
8	1	22	-2 6
9	-6 1	23	-4 6
10	-5 1	24	5 6
11	1 2	25	1 6
12	-6 1 2	26	7
13	3	27	6
14	-5 3	28	4

★ O-D MATRIX

	1	2	3	4	5	6	7	8	9	
1	0	0	45	0	33	2[illegible]	66	93	120	385
2	0	0	0	36	0	87	24	41	59	247
3	31	0	0	14	65	0	93	110	13	341
4	0	12	26	0	0	45	0	55	66	203
5	33	0	46	0	0	0	27	0	12	118
6	87	76	0	65	0	0	43	32	0	303
7	40	26	37	0	65	32	0	0	18	218
8	36	63	59	19	0	2[illegible]	0	0	0	206
9	69	93	50	26	33	0	12	0	0	303
	316	270	263	160	214	220	262	331	258	2324

POWER FOR ASSIGNMENT : 2 WALKING MODE : 26

***OPTION ASSIGNMENT PROCESS-FINAL RESULTS

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 1> TO < 5> ***							
3990	900	717	0	2373	12	4	* 1* 11* 2* 23* 5*
4156	719	1161	0	2276	11	4	* 1* 7* 4* 18* 5*
4576	420	810	2476	870	10	2	* 1* 7* 4* 26* 5*
*** FROM < 1> TO < 7> ***							
4825	1319	966	0	2600	22	4	* 1* 11* 2* 27* 7*
5095	900	1822	0	2373	20	4	* 1* 7* 4* 4* 7*
4642	2158	675	0	1809	24	1	* 1* 8* 7*
*** FROM < 1> TO < 8> ***							
4548	1139	906	0	2503	45	4	* 1* 11* 2* 27* 8*
4365	1976	675	0	1712	48	1	* 1* 8* 8*
*** FROM < 3> TO < 7> ***							
3122	1259	540	0	1323	48	1	* 3* 27* 7*
* 3349	1319	675	0	1355	42	1	* 3* 8* 7*
*** FROM < 3> TO < 8> ***							
2845	1079	540	0	1226	59	1	* 3* 27* 8*
* 3072	1139	675	0	1258	51	1	* 3* 8* 8*
*** FROM < 4> TO < 8> ***							
3368	539	650	0	2179	15	4	* 4* 18* 5* 24* 8*
3788	240	299	2476	773	12	2	* 4* 26* 5* 24* 8*

	T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 4> TO < 8> ***								
	3436	300	351	1980	805	14	2	* 4* 18* 5* 26* 8*
	3491	1499	540	0	1452	14	1	* 4* 13* 8*
*** FROM < 5> TO < 1> ***								
	4341	839	1161	0	2341	10	4	* 5* 19* 2* 7* 1*
	3898	600	351	1980	967	12	2	* 5* 19* 2* 26* 1*
	3990	900	717	0	2373	11	4	* 5* 6* 4* 11* 1*
*** FROM < 6> TO < 4> ***								
	2569	900	540	0	1129	37	1	* 6* 5* 4*
*	2980	1079	675	0	1226	28	1	* 6* 8* 4*
*** FROM < 7> TO < 2> ***								
	3258	1260	675	0	1323	12	1	* 7* 8* 2*
*	2937	1139	540	0	1258	14	1	* 7* 1* 2*
*** FROM < 7> TO < 3> ***								
	3997	1739	675	0	1583	17	1	* 7* 8* 3*
*	3676	1619	540	0	1517	29	1	* 7* 1* 3*
*** FROM < 7> TO < 6> ***								
	4920	2339	675	0	1906	6	1	* 7* 8* 6*
	4864	1319	945	0	2600	6	4	* 7* 17* 4* 13* 6*
	4573	1199	839	0	2535	7	4	* 7* 1* 5* 14* 6*

***OPTION ASSIGNMENT PROCESS-FINAL RESULTS

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 7> TO < 6> ***							
4629	1079	1080	0	2470	7	4	* 7* 1* 8* 5* 6*
4436	839	540	1980	1097	6	2	* 7* 26* 8* 5* 6*
*** FROM < 7> TO < 9> ***							
3798	539	1080	0	2179	6	4	* 7* 1* 8* 5* 9*
3625	300	540	1980	805	7	2	* 7* 26* 8* 5* 9*
4029	240	540	2476	773	5	2	* 7* 1* 8* 26* 9*
*** FROM < 8> TO < 6> ***							
* 2796	960	675	0	1161	13	1	* 8* 3* 6*
2476	839	540	0	1097	16	1	* 8* 5* 6*
*** FROM < 9> TO < 2> ***							
3719	1559	675	0	1485	44	1	* 9* 8* 2*
3557	899	540	989	1129	49	2	* 9* 26* 8* 1* 2*
*** FROM < 9> TO < 3> ***							
4458	2039	675	0	1744	13	1	* 9* 8* 3*
4296	1379	540	989	1388	14	2	* 9* 26* 8* 1* 3*
4474	960	1109	0	2405	13	4	* 9* 20* 6* 7* 3*
5034	1499	839	0	2696	10	4	* 9* 16* 8* 1* 3*
*** FROM < 9> TO < 5> ***							
2605	1080	299	0	1226	15	1	* 9* 20* 5*

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 9> TO < 5> ***							
2393	300	299	989	805	13	2	***** ***** ***** * 9* 26* 8* 2* 5* ***** ***** *****
*** FROM < 1> TO < 3> ***							
2303	840	366	0	1097	45	1	***** ***** * 1* 11* 3* ***** *****
*** FROM < 1> TO < 6> ***							
3226	1439	366	0	1421	28	1	***** ***** * 1* 11* 6* ***** *****
*** FROM < 1> TO < 9> ***							
4181	1859	675	0	1647	120	1	***** ***** * 1* 8* 9* ***** *****
*** FROM < 2> TO < 4> ***							
2469	660	810	0	999	36	1	***** ***** * 2* 7* 4* ***** *****
*** FROM < 2> TO < 6> ***							
2672	1080	366	0	1226	87	1	***** ***** * 2* 11* 6* ***** *****
*** FROM < 2> TO < 7> ***							
2661	960	540	0	1161	24	1	***** ***** * 2* 27* 7* ***** *****
*** FROM < 2> TO < 8> ***							
2384	780	540	0	1064	41	1	***** ***** * 2* 27* 8* ***** *****
*** FROM < 2> TO < 9> ***							
3627	1499	675	0	1453	59	1	***** ***** * 2* 8* 9* ***** *****
*** FROM < 3> TO < 1> ***							
2284	539	810	0	935	31	1	***** ***** * 3* 7* 1* ***** *****

***OPTION ASSIGNMENT PROCESS-FINAL RESULTS

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 3>	TO < 4> ***						
2930	959	810	0	1161	14	1	***** ***** * 3* 7* 4* ***** *****
*** FROM < 3>	TO < 5> ***						
2476	839	540	0	1097	83	1	***** ***** * 3* 27* 5* ***** *****
*** FROM < 3>	TO < 9> ***						
2888	1019	675	0	1194	13	1	***** ***** * 3* 8* 9* ***** *****
*** FROM < 4>	TO < 2> ***						
2395	900	356	0	1129	12	1	***** ***** * 4* 11* 2* ***** *****
*** FROM < 4>	TO < 3> ***						
3134	1380	366	0	1388	26	1	***** ***** * 4* 11* 3* ***** *****
*** FROM < 4>	TO < 6> ***						
2661	960	540	0	1161	44	1	***** ***** * 4* 13* 6* ***** *****
*** FROM < 4>	TO < 9> ***						
3307	1379	540	0	1388	66	1	***** ***** * 4* 13* 9* ***** *****
*** FROM < 5>	TO < 3> ***						
2846	1080	540	0	1226	46	1	***** ***** * 5* 1* 3* ***** *****
*** FROM < 5>	TO < 7> ***						
1830	420	540	0	870	27	1	***** ***** * 5* 27* 7* ***** *****
*** FROM < 5>	TO < 9> ***						
2605	1080	299	0	1226	12	1	***** ***** * 5* 14* 9* ***** *****

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 6> TO < 1> ***							
2930	959	810	0	1161	87	1	***** ***** * 6* 7* 1* ***** *****
*** FROM < 6> TO < 2> ***							
2561	719	810	0	1032	76	1	***** ***** * 6* 7* 2* ***** *****
*** FROM < 6> TO < 7> ***							
2426	719	675	0	1032	43	1	***** ***** * 6* 8* 7* ***** *****
*** FROM < 6> TO < 8> ***							
1773	539	299	0	935	32	1	***** ***** * 6* 16* 8* ***** *****
*** FROM < 7> TO < 1> ***							
2704	900	675	0	1129	40	1	***** ***** * 7* 8* 1* ***** *****
*** FROM < 7> TO < 5> ***							
2014	539	540	0	935	65	1	***** ***** * 7* 1* 5* ***** *****
*** FROM < 8> TO < 1> ***							
2981	1080	675	0	1226	36	1	***** ***** * 8* 8* 1* ***** *****
*** FROM < 8> TO < 2> ***							
2568	899	540	0	1129	63	1	***** ***** * 8* 1* 2* ***** *****
*** FROM < 8> TO < 3> ***							
3307	1379	540	0	1388	59	1	***** ***** * 8* 1* 3* ***** *****
*** FROM < 8> TO < 4> ***							
2150	540	675	0	935	19	1	***** ***** * 8* 8* 4* ***** *****

***OPTION ASSIGNMENT PROCESS-FINAL RESULTS

APPENDIX C - 9

T.T.TIME(SEC)	R.TIME(SEC)	W.TIME(SEC)	WK.TIME(SEC)	FARE(SEC)	PASSENGERS	T	OPTION DESCRIPTION
*** FROM < 9> TO < 1> ***							
3165	1199	675	0	1291	89	1	***** ***** * 9* 8* 1* ***** *****
*** FROM < 9> TO < 4> ***							
2334	660	675	0	999	26	1	***** ***** * 9* 8* 4* ***** *****
*** FROM < 9> TO < 7> ***							
1780	300	675	0	805	12	1	***** ***** * 9* 8* 7* ***** *****
3131	1079	626	96	1330	2324		GENERALISED COST : AVERAGE PASSENGER JOURNEY
100	34	19	3	44	2324		GENERALISED COST : AVERAGE PASSENGER JOURNEY

OPTION ASSIGNMENT PROCESS-FINAL RESULTS APPENDIX C - 10

* MODIFIED O-D MATRIX

	1	2	3	4	5	6	7	8	9	
1	0	79	45	41	0	26	24	48	120	385
2	10	0	0	36	12	87	46	86	59	336
3	31	0	0	14	85	0	90	110	10	341
4	11	12	26	0	40	50	20	14	66	239
5	0	22	46	11	0	7	27	27	12	152
6	87	76	13	65	0	0	43	32	0	316
7	40	26	37	6	72	6	0	18	0	205
8	36	112	83	19	18	42	0	0	13	323
9	89	44	15	26	15	10	12	10	0	222
	304	371	263	218	240	253	262	345	283	2519

* WALKING-LINKS

I	J	WK(M)	PER	ACCPER
9	6	1335	35.74	35.74
4	5	907	24.28	60.03
5	6	462	12.37	72.40
7	8	429	11.49	83.88
2	1	396	10.60	94.48
8	9	206	5.52	100.00

* WALKING-OD

I	J	WK(M)	PER	ACCPER
4	8	957	25.64	25.64
9	2	807	21.62	47.25
7	9	437	11.71	58.96
1	5	412	11.04	70.00
5	1	396	10.61	80.61
9	5	296	7.93	88.53
9	3	230	6.16	94.70
7	6	198	5.30	100.00

REFERENCES

ACKOFF, R.L. and M.W. SASIENI (1968) Fundamentals of operations research, John Wiley & Sons Inc.

ADLER, H.A. (1971) Economic Appraisal of transport projects: a manual with case studies, Indiana University Press, Bloomington/London.

ALLEN, W.G. Jr. and F. DICESARE (1976) Transit service evaluation: preliminary identification of variables characterizing level of service. Transportation Research Record, 606, 41-47.

ALTER, C.H. (1976) Evaluation of public transit services: the level-of-service concept. Transportation Research Record, 606, 37-40.

ANDERSON, R.B. and L.A. HOEL (1974) Estimating latent demand and cost for statewide transit service. Transportation Research Record, 519, 26-35.

BAKKER, J.J. (1976) Transit operating strategies and levels of service. Transportation Research Record, 606, 1-5.

BARTON, R.F. (1970) A primer on simulation and gaming, Prentice-Hall Inc.

BJELKÄKER, S. and S. SCHÉELE (1973) Optimisation methods for bus systems in urban areas: a literature review. Proceedings of a seminar on theoretical models for fixed scheduled and demand actuated bus routing in urban areas. Linköping University.

BLUNDEN, W.R. (1971) The land-use / Transport system: analysis and synthesis, Pergamon Press.

BLY, P.H. (1973) Optimisation of a fixed route bus network. Transport and Road Research Laboratory, SR 7 UC, Crowthorne.

BLY, P.H. and R.H. OLDFIELD (1974) Optimisation of a simple model bus network. Transport and Road Research Laboratory, SR 27 UC, Crowthorne.

BLY, P.H. and R.L. JACKSON (1974) Evaluation of bus control strategies by simulation. Transport and Road Research Laboratory, LR 637, Crowthorne.

BLY, P.H. (1976) The effect of fares on bus patronage. Transport and Road Research Laboratory, LR 733, Crowthorne.

BROWNE, J.J. (1967) Simulation of public transportation operations. Proceedings of the 18 th Annual Institute Conference and Convention, American Institute of Industrial Engineers, 301-306.

BRUTON, M.J. (1975) Introduction to transportation planning, Hutchinson of London.

BUCHANAN, C. et al (1968) Traffic in towns, in Transport edited by D. Munby, Penguin Books.

BUCHANAN, C.M. (1976) The Telford short-term bus study. Proceedings of Seminar in Public Transport, P.T.R.C P.141, 1-13

BULLOCK, D.A. (1970) Scientific methods in bus control assessments. Public Road Transport Analysis, P.T.R.C., London.

C.I.D. (1969) Alternativas para el desarrollo urbano de Bogotá D.E. Universidad Nacional de Colombia.

CATANESE, A.J. (1972) Scientific methods of urban analysis, Leonard Hill Books.

CHANDAN, A. (1968) A scientific approach to bus service planning. Proceedings of a P.T.R.C. Seminar: Public Transport Analysis II. Birkbeck College London.

CHAPMAN, R.A. (1975) Bus boarding times- a review of studies and suggestions for interpretation. Transport Operations Research Group, Working Paper 8. University of Newcastle upon Tyne.

CHAPMAN, R.A. (1976) Bus reliability- definition and measurement. Working Paper 18. Transport Operations Research Group, University of Newcastle upon Tyne.

CHAPMAN, R.A., H.E. GAULT and I.A. JENKINS (1976) Factors affecting the operation of urban routes. Transport Operations Research Group, Working Paper 23. University of Newcastle upon Tyne.

CHAPMAN, R.A., H.E. GAULT and I.A. JENKINS (1977) The operation of urban bus routes. 1. Background and preliminary analysis. Traffic Engineering and Control, 18, 6, 294-298.

CHAPMAN, R.A., H.E. GAULT and I.A. JENKINS (1977) Factors affecting the operation of urban bus routes. 2. Sources of irregularity. Traffic Engineering and Control, 18, 7/8, 364-367.

CHAPMAN, R.A., H.E. GAULT and I.A. JENKINS (1977) Factors affecting the operation of urban bus routes. 3. Wider aspects of operation. Traffic Engineering and Control, 18, 9, 416-419.

CHORAFAS, D.N. (1965) Systems and simulation. Mathematics in Science and Engineering, 14. Academic Press Inc.

CHRIQUI, C. and P. ROBILLARD (1975) Common bus lines. Transportation Science, 9, 115-121.

COE, G.A. and R.L. JACKSON (1977) Measurement of waiting times in assessments of operational changes in bus services. Department of the Environment, T.R.R.L. S.R. 298.

COSTINETT, P.J. (1970) Transportation study models in public transport planning. Public Road Transport Analysis, P.T.R.C. London.

DAJANI, J.S. and G. GILBERT (1978) Measuring the performance of transit systems. Transportation Planning and Technology, 4, 2, 97-103.

DALY, A.J. (1973) TRANSEPT- A multipath public transport assignment model. Seminar Proceedings Urban Traffic Annual Meeting, University of Sussex. P.T.R.C. P.81.

DEWEES, D.N. (1976) Urban bus and railroad performance. Journal of Transport Economics and Policy, 10, 1, 16-25.

DIAL, R.B. (1967) Transit pathfinder algorithm. Highway Research Record, 205, 67-85.

EMSHOFF, J.R. and R.L. SISSON (1970) Design and the use of computer simulation models. The Macmillan Company.

ERIKSEN, A.R. and E.A. BEIMBORN (1973) Effectiveness of alternative downtown area bus distribution systems. Highway Research Record, 459, 29-39.

FORTRAN: Language. Technical Publication 6800. ICL, 1978.

FREY, H. (1976) Electronic system to assist the management of bus operations. Control in Transportation Systems. International Federation of Automical Control.

FRIEDMAN, F.L. and E.B. ELLIOT (1977) Problem solving and structured programming in FORTRAN, Addison-Wesley Publishing Company.

GAZIS, D.C. (1974) Traffic science. John Wiley & Sons.

GEORGE, J.A. and H.G. DAELLENBACH (1978) Introduction to operations research techniques, Allyn and Bacon Inc.

GERRARD, J.P.D. (1970) Applications of computers in bus operations: a state-of-the-art survey, IBM U.K. Report 0003.

GERRARD, J.P.D. and D. BROOK (1972) A computer simulation model for bus operation planning. IBM U.K. Report 0040.

GOODE, A.P. (1973) The bus demonstration programme. Moving people in cities. A conference held at the TRRL, Crowthorne. Transport and Road Research Laboratory, SR 6 UC. Crowthorne.

GORDON, G. (1978) System simulation, Prentice-Hall,Inc.

HANEY, D.G. and G.E. KLEIN (1972) Community values in transport network evaluation. Proceedings of the A.S.C.E. Transportation Engineering Journal, 98, TE4 , 823-835.

HAY, A. (1973) Transport for the space economy: a geographical study. The Macmillan Press Ltd.

HEDGES, C.A. (1974) Let's attack the real urban transportation problem.Transportation Research Record, 519, 10-25.

HOLLINGDALE, S.H. (1978) Methods of operational analysis, in Newer uses of mathematics edited by J. Lighthill, 176-280, Penguin Books.

HOLM, J.(1973) A model for allocation of buses in a route network. Proceedings of a Seminar on Theoretical models for fixed scheduled and demand-actuated bus routing in urban areas. Linköping University, 259-275.

HOLROYD, E.M. (1967) The optimum bus service: a theoretical model for a large uniform urban area, in Vehicular traffic science edited by L.C. Edie. American Elsevier Publishing Company, Inc, New York.

HOLROYD, E.M. and D.A. SCRAGGS (1966) Waiting times for buses in Central London. Traffic Engineering and Control, 8, 3, 158-160.

HSU, J. and V.H. SURTI (1975) Framework of route selection in bus network design. Transportation Research Record, 546, 44-57.

HSU, J. and V.H.SURTI (1977) Decomposition approach to bus network design. Transportation Engineering Journal, 103, TE 4 , Proceedings of the A.S.C.E. 447-459.

IRWIN, N.A. (1966) Criteria for evaluating alternative transportation systems. Highway Research Record,148, 9-12.

JACKSON, H.S. (1972) Computer simulation modelling of public transport networks. (Unpublished PhD thesis) University of Leeds.

JACKSON, H.S. (1973) Computer simulation models of a bus system. Seminar Proceedings Operational Research in the Bus Industry. University of Leeds,23-25.

JACKSON, H.S. and A. WREN (1972) Simulation of urban bus operations. University of Leeds, Operational Research Unit, ULORU-10.

JACKSON, R.L. (1976) An evaluation of bus control systems by simulation. Transport and Road Research Laboratory, LR 712, Crowthorne.

JACOBS, G.D. and P.R. FOURACRE (1974) Urban transport in developing countries. Traffic Engineering and Control, 15, 20, 926-928.

JAMIESON, G.B., W.K. MACKAY and J.C. LATCHFORD (1967) Transportation and land use structures. Urban Studies, 4, 3.

JENKINS, I.A. (1976) A comparison of several techniques for simulating bus routes. Transport Operations Research Group, Working Paper 14, University of Newcastle upon Tyne.

JOLLIFFE, J.K. and T.P. HUTCHINSON (1975) A behavioural explanation of the association between bus and passenger arrivals at a bus stop. Transportation Science, 9, 3, 248-282.

JONES, I.S. (1977) Urban transport appraisal. The Macmillan Press Ltd.

KAUFMANN, A. (1963) Methods and models of operations research, Prentice-Hall, Inc.

KLEIJNEN, J.P.C. (1974) Statistical techniques in simulation. Part 1, Marcel Dekker, Inc.

KOBAYASHI, K.,Y. AOKI and A. TANI (1975) A method for evaluating urban transportation planning in terms of user benefit. Transportation Research, 9, 1, 67-79.

KRAFT, W.H. and H. DEUTSCHMAN (1977) Bus passenger service-time distributions. Transportation Research Record, 625, 37-43.

LAMPKIN, W. and P.D. SAALMANS (1967) The design of routes, service frequencies and schedules for a municipal bus undertaking: a case atudy. Operational Research Quarterly, 18, 4, 375-397.

LANE, R. , T.J. POWELL and P.P. SMITH (1971) Analytical transport planning, Duckworth, London.

LAST, A. and A.J. DALY (1975) Modelling the walk mode. Proceedings of the Seminar on Urban Traffic Models. P.T.R.C., P122, 333-340.

LAST, A. and S.E. LEAK (1976) Transept: a bus model. Traffic Engineering and Control, 17, 1, 14-17.

LEHMAN, R.S. (1977) Computer simulation and modelling: an introduction, LEA Publishers.

LESLEY, L.J.S. (1975) Bus priority and modal split. Traffic Engineering and Control, 16, 3, 127-129.

LESLEY, L.J.S. (1976) Optimum bus-stop spacing. Traffic Engineering and Control, 17, 10, 399-401.

LESLEY, L.J.S. (1976) Optimum bus-stop spacing. Traffic Engineering and Control, 17, 11, 472-475.

LILLIENBERG, S. (1973) A computer system for public transport analysis. Proceedings of a Seminar on Theoretical Models for fixed scheduled and demand-actuated bus routing in urban areas. Linköping University, 288-293.

LOCAL GOVERNMENT OPERATIONAL RESEARCH UNIT (1973) Planning urban bus routes, L.G.O.R.U Report C149.

LODATO, V.A. and P.A. GORDON (1975) A mass transit simulator in APL/370. Modelling and simulation vol 6. Proceedings of the Sixth Annual Pittsburgh Conference, 721-729.

MACDOUGALL, E.B. (1976) Computer Programming for spatial problems, Edward Arnold Ltd.

McCAMERON, F.A. (1970) FORTRAN IV, Irwin-Dorsey Ltd.

MAISEL, H. and G. GNUGNOLI (1972) Simulation of discrete stochastic systems. Science Research Associates, Inc.

MANHEIM, M.L. (1967) Principles of transport systems analysis. Highway Research Record, 180, 11-20.

MARTIN, F.F. (1968) Computer modelling and simulation, John Wiley & Sons, Inc.

MYERS, S. (1967) Transit technology: the soft revolution. Future directions for research in urban transportation. O.E.C.D.

NAYLOR, T.H., J.L. BALINTFY, D.S. BURDICK and K. CHU (1968) Computer simulation techniques. John Wiley & Sons, Inc.

NEWELL, G.F. and R.B. POTTS (1964) Maintaining a bus schedule. Proceedings of the Second Australian Road Research Board Conference, 388-393.

O.E.C.D. (1972) Optimisation of bus operation in urban areas: a report prepared by an O.E.C.D. Road Research Group.

O'FLAHERTY, C.A. and D.O. MANGAN (1970) Bus passenger waiting times in central areas. Traffic Engineering and Control, 11,419-421.

OJEDA, S. (1977) Mass transit in Bogotá. Transportation Engineering, 47, 11.

PARKER, G.B. (1977) The Bradford Bus Study. 2. The results: role of buses, rerouting and costing. Traffic Engineering and Control, 18, 1, 22-25.

PASLEY, T.J. and J. SUGDEN (1972) A network rationalisation of public transport services in Teesside. Traffic Engineering and Control, 13, 11/12, 517-520.

PIKE, D.H., P.I. FULLER and M.T. WHITE (1976) Telford transportation study : 1 An aid to the preliminary evaluation of long-term public transport options. Traffic Engineering and Control, 17, 2, 52-55.

Plan de Estructura para Bogotá: Informe técnico sobre el estudio de desarrollo urbano de Bogotá Fase II, 1974.

POOLE, T.G. and J.Z. SZYMANKIEWICZ (1977) Using simulation to solve problems, Mc Graw-Hill.

POTTS, R.B. (1978) Networks, in Newer uses of mathematics, edited by James Lighthill, 281-325. Penguin Books.

POTTS, R.B. and E.A. TAMLIN (1964) Pairing of buses. Australian Road Research, 2, 2, 3-9.

R. TRAVERS MORGAN & PARTNERS (1976) Bradford bus study. Final report.

RAPP, M.H. and A. ROBERT-GRANDPIERRE (1976) Interactive graphics system for transit route optimisation. Transportation Research Record, 559, 73-88.

REEVES, L.H.D. (1970) Omnibus- a new computer program for public transport planning. Public Road Transport Analysis, PTRC, London.

SCHNEIDER, J.B. and J.W.CLARK (1976) High-performance bus rapid transit systems: a design-process experiment. Transportation Research Record, 606, 30-36.

SEDDON, P.A. and M.P. DAY (1974) Bus passenger waiting times in Greater Manchester. Traffic Engineering and Control, 15, 9, 442-445.

SESHAGIRI, N. and others (1969) Computer generated time-tables and bus schedules for a large bus transport network. Transportation Science, 9, 69-85.

SHANNON, R.E. (1975) Systems simulation, the art and science. Prentice-Hall Inc.

SHARMA, S. (1976) Intra-urban transport optimisation through operations research techniques. Traffic Quarterly, 30, 3.

SMEED, R.J. (1968) Traffic studies and urban congestion. Journal of Transport Economics and Policy, 2, 33-70.

SMITH, E. (1973) An economic comparison of urban railways and express bus services. Journal of Transport Economics and Policy, 7, 20-31.

SPRINGBURN Public Transport Survey; Greater Glasgow Passenger Transport Executive, 4 volumes,(1975).

STAIRS, S. (1968) Selecting an optimal traffic network. Journal of Transport Economics and Policy, 2, 218-231.

STÖCKER, K. (1974) Simulation of traffic flow through large networks-a tool of transportation system analysis. Urban Traffic Models, II, PTRC.

STOPHER, P.R. and A.H. MEYBURG (1975) Urban transportation modeling and planning. Lexington Books.

Transport and Road Research Laboratory (1974) Symposium on public transport fare structure: papers and discussion, SR 37 UC. Crowthorne.

TAAFFE, E.J. and H.L. GAUTHIER (1973) Geography of transportation. Prentice-Hall Inc.

THOMPSON, G.L. (1977) Planning considerations for alternative transit route structures. <u>Journal of the American Institute of Planners</u>, <u>43</u>, 2, 158-168.

TINKER, J. (1975) How Delhi made the buses run on time. New Scientist, 64-66.

The Bradford Bus Study: see R. Travers Morgan & Partners.

TOCHER, K.D. (1963) The art of simulation. The English Universities Press Ltd. London.

TOMAZINIS, A.R. (1974) Planning urban transportation systems for productivity, efficiency, and quality of services. Transportation Research Record, 491.

VUCHIC, V.R. and R.M. STANGER (1973) Lindelwold rail line and Shirley busway: a comparison. Highway Research Record, 459, 13-28.

VUCHIC, V.R. and E.L. TENNYSON (1974) Application of guidelines for improving transit service and operating efficiency. Transportation Research Record, 519, 66-72.

WEBSTER, F.V. (1970) The optimum size of buses for urban areas. Symposium Proceedings Public Road Transport Analysis, PTRC. London.

WELDING, P.I. (1957) The instability of a close-interval-service. <u>Operational Research Quarterly</u>, <u>8</u>, 3, 133-148.

WELLS, G.R. (1975) Comprehensive transport planning. Charles Griffin & Company Ltd.

WHITE, P.R. (1976) Planning for public transport. Hutchinson of London.

WOHL, M. and B.V. MARTIN (1967) Traffic system analysis. For engineers and planners; McGraw-Hill.

WREN, A. (1971) Computers in transport planning and operation. Allan. London.

YOUNG, R.N., C.W. OCKERT and C.E.PIXTON (1971) Evaluating transportation systems. Transportation Engineering Journal, 97, 2, 347-352. Proceedings of the A.S.C.E.