

A Transportation Model For Demand Responsive Fleet Operation In A Manufacturing Firm

E.O. Oyatoye, University of Lagos, Nigeria
J.A.O. Magbagbeola, University of Lagos, Nigeria

ABSTRACT

We study the problem of determining the number of vehicles needed to provide a demand responsive transit service with a predetermined quality for the user in terms of waiting time at the stops and maximum allowed detour. We propose a probabilistic model that requires only the knowledge of the distribution of the demand over the service area and the quality of the service in terms of time windows associated with pickup and delivery nodes. This methodology can be much more effective and straight forward compared to a simulation approach whenever detailed data on demand patterns are not available. Computational results under a fairly broad range of test problems show that the model can provide an estimation of the required size of the fleet in several different scenarios.

Keywords: Transportation model, demand responsive, fleet operation, predetermined quality.

INTRODUCTION

Every decision-making process produces a final choice or opinion. It begins when we need to do something but we do not know what. Therefore, decision-making is a reasoning process, which can be rational or irrational, and can be based on explicit assumptions or tacit assumptions.

Decision-making is said to be a psychological construct. This means that although we can never “see” a decision, we can infer from observable behavior that a decision has been made. Therefore, we conclude that a psychological event that we call “decision making” has occurred. It is a construction that imputes commitment to effect the action. The manager is forever beset by the necessity to choose among alternatives the outcome of which is definitely unknown. Strategic decisions take place at three different levels in an organization. These levels are functional, businesses and corporate. The corporate strategic manager seeks to maximize the attainment of long-term organization priorities such as maximizing shareholders wealth and developing managerial ability.

The political and business terrain in many developing countries in recent time has been that of biting economic hardship, great uncertainties and instability. Most organizations in these nations are having serious problems in coping with their corporate functions and responsibilities. Of particular interest is the fleet operation problem manufacturing organizations have been encountering. The decision on the choice of route of transportation which tells more on the profitability and success of such organizations, has posed a great challenge to the Fleet Operation Departments of many of these organizations. Hence, this study considers a many-to-many demand responsive transit service with predefined itineraries and schedules.

The Problem

The economic downturn in many African countries, particularly Nigeria, has been having serious effects on some organizations in the manufacturing sector to the extent that a number of such organizations have had to close shop due to escalating costs of production and distribution. However, the distribution cost outweighs the production

cost. We investigated to what extent a transportation model could help in solving the problem and which type of model will be most suitable for use in order to:

- Achieve the attainment of corporate objectives
- Optimize the profitability level

Related Studies

A good deal of research has been devoted to the investigation of issues similar to ours, and it is interesting to compare different approaches in order to draw lessons from our problem. As will be shown in the following sections, the proposed methodology relies on a variety of research fields. In what immediately follows, we take a brief look at the past work of various authors in the area being investigated.

Freight Distribution Systems

One of the related fields that can provide useful insights for our problem is the abundance literature on the modeling of freight distribution systems. Bearwood, Halton and Hammersley (1959) and Eilon, Watson-Grandy and Christofides (1971) estimate the length of a TSP tour for a fleet of vehicles through simulation. In distribution problems it is usually possible to divide the service area into several zones each of which is served with one vehicle and each path is estimated using the formulas for the TSP tour. This very popular technique, called “cluster-first, route second”, has been successively used in many papers modeling distribution problems. Larson and Odoni (1981) provide useful insights for the multi-route problem, while a generalization of the TSP formula for zones of different shapes is provided in Daganzo (1984a).

Some authors [for example Adebisi and Hurdle (1982); Aldaihani, Quadrifoglio, Dessouky and Hall (2004); Quadrifoglio, Hall and Dessouky (2005)] adapt a model for fixed lines bus systems to flexible services (that is, services in which the buses can deviate from their pre-defined path to serve requests off the route). In those cases, the decision variable usually considered is the headway between two successive vehicles or the slack in the schedule. This kind of service is different from our DRT system, since in our case there are no predefined paths and so headways cannot be defined. For this reason, a model for conventional transit system cannot be used in our case.

In the 1970's due to the diffusion of Para transit services, some researchers proposed different methodologies to model simplified variants of a DRT system in order to compare them with conventional bus line networks (Ward, 1975). The issue of the design of an integrated urban public transportation system was investigated by Batchelder and Kullman (1977). However in this case the model for the dial-a-ride system was based on computer simulations calibrated on real datasets. Wilson and Hendrickson (1980) focus on performance models, where the decision variable is related to the quality of the service, and provide an excellent comparative analysis on the different methodologies that have been proposed. They also report from previous unpublished research empirical models for the determination of the number of vehicles that were calibrated on real data. It is well known that empirical models are difficult to use in a context that is different from the one upon which they have been calibrated.

Another modeled system is “many-to-one”, where there are many origins and a single destination. In this case the service area is usually divided into several zones and in each of these zones, only one vehicle can operate. A typical example of this is the waste disposal problem in Lagos and Oyo states in Nigeria. Each vehicle collects the request in its zone and delivers them to a central location. It is straightforward to see that in this case it is possible to decompose the problem into several smaller TSP, and to successfully apply the previously mentioned “cluster-first, route-second” methodologies for the estimation of both the number of vehicles and the distance traveled. In this case, the decision variable can be the size of the zone or the capacity of the vehicle.

Some research has focused on “many-to-many” systems, where there are many origins and destinations. Arrillaga and Medville (1974) and Flushberg and Wilson (1976) make use of regression models, whereas Lerman and Wilson (1974) and Daganzo (1978) propose stochastic models in which the customer's arrival at a stop is a Poisson distributed queuing process. The general drawback of the former approach is the limitation of the validity range depending on the datasets used for the calibration, whereas the latter may present problems in the case of un-

congested systems. Daganzo (1984b) performs a comparison of fixed and flexible transit systems by modeling their costs. In this work, door to door service is a limiting case of a jitney service, that is the considered routing strategy consists of dispatching vehicles with constant headways and the stops, without waiting passengers are skipped.

An interesting theoretical discussion is provided in Stein (1978a, 1978b) where on the basis of a probabilistic analysis; a class of scheduling rules is suggested. The outcome is that a decomposition algorithm in which buses serve a small zone and passengers across different zones have to transfer seems to asymptotically outperform systems in which the vehicles can travel in the whole area and a customer is inserted on the basis of the cost minimization.

AGENT-BASED MODEL SIMULATION (ABMS)

An attempt at aiding the understanding of this study will necessitate a little excursion into the explanation of Modeling Technique, as explained through the understanding of Agent-Based Simulation. Agent-based simulation is establishing itself as a serious, useful area of study.

The essential idea of agent-based modeling and simulation (ABMS) is that many phenomena, even very complex ones, can best be understood as systems of autonomous agents that are relatively simple and follow relatively simple rules for interaction. Applications range from modeling agent behaviour in the stock market (Arthur et al. 1997) and supply chains (Macal 2003, 2004) and modeling bacterial cell behavior (Emonet et al. 2005). Agent-based modeling and simulation is also an experimental technique, a framework for developing electronic laboratories in which the most detailed assumptions about individual agents, their behaviors and interactions can be varied and explored in silicone.

ABMS and Traditional M&S Techniques

Agent-based modeling can provide an overarching framework for model based on other modeling techniques. For example, models may be composed of agents whose decision-making behaviors are represented by formal optimization problems or by informal decision heuristics. Another example is agent behaviors represented as statistical models deriving agent behaviors from the agents' input information. Agent-based modeling can also be used as a complement to other modeling techniques: for example, an agent model that builds system behavior from the behaviors of the individual agents can be "docked" (use in conjunction) with a more aggregate systems dynamics model of the system, to see whether the two approaches yield similar results over a range of test cases.

The goal of this study is to model a many-to-many demand responsive transit service without predefined itineraries and schedules. In this case, the fleet has to be dispatched exclusively on the basis of the list of requests, like in taxicab systems, the difference being the possibility of serving customers with some detours in order to share the ride. We believe that this kind of service is of particular interest for the possibility of offering a high quality service with an efficient allocation of the resources. To achieve this, we have modeled a service in which time windows are associated with each pickup and delivery point.

The definition of time window is different from the notion of "time deadline" that can be found in previous works, for example concerning hauling services (Hall, 1996). Although Daganzo (1987) modeled a distribution problem considering time windows associated with each delivery point, the suggested methodology is not suitable when temporal constraints are tight as in the case we are considering. Thus, we need a procedure that is not easily derivable from existing methodologies. For example, comparing our problem to the previously discussed ones, it can be observed that in our case, it is impossible to model it as a fixed-line service since we cannot define a "path" or "headway" between the vehicles. On the other hand, the joint need of avoiding transfers for any pair of pickup and delivery points and of limiting the maximum ride time for every customer prevents us from dividing the area into several service zones served by a single vehicle, hence, a "cluster-first, route-second" model is not appropriate.

METHODOLOGY

Specification of the Studied System

Assumption 1

In the following, we will start by partially adopting the operating scenario described by Jaw et al. (1986).

- Our demand responsive transit (DRT) system consists of a fleet of vehicles with no predefined schedules.
- The vehicles travel at a constant speed and cannot idle. We later show where relaxing the no idling assumption by considering a more idealized scenario.
- The service time at the locations is zero and we do not consider capacity constraints since in most practical cases they are dominated by time window constraints.
- When making a reservation, the customer has to specify the origin and the destination of the trip, as well as the pickup time.
- The coordinates of the pickup and the delivery points are random variables drawn from the same distribution. Hence, given this distribution, it is possible to compute the distribution of the Euclidean travel distances between any pair of points.

Parameters

Let a and b be two parameters that are specified by the scheduler, with $a \geq 0$ and $b \geq 1$.

Variables

Since the vehicles travel at a constant speed v , the direct ride time is simply $L(D_k, P_k)/v$ assuming the Euclidean metric, where P_k and D_k are the pickup and the delivery points of request k , respectively.

Model 1 – Computation of Time Window

Let $L(A,B)$ be a random variable from the latter distribution, representing the distance between point A and point B . In order to ensure an acceptable quality of the service, the vehicle has to pick up the customer no earlier than the pickup time and no later than a specified time interval from the pickup time. The vehicles cannot pickup a customer earlier than the pickup time because customers may not be there at that time. Also the maximum length of the trip must be somewhat limited. To do this, we fix a maximum wait state WS , which is the same for all the customers, and we compute a maximum ride time MRT_k for each request k .

The maximum ride time is defined as an increasing function of the direct ride time that is the time needed to serve the request without deviations. The maximum ride time of each customer k is computed as: $MRT_k = a + b \cdot L(D_k, P_k)/v$

The above scheduling constraints related to the maximum wait state and maximum ride time for each request k define the quality of the service. The most practical way to take them into account in the scheduling process is to define time windows for all the pickup and delivery locations.

Let EPT_k and LPT_k be the earliest pickup time and latest pickup time of customer k , respectively; while EDT_k and LDT_k are the corresponding earliest delivery time and latest time respectively. Define (EPT_k, LPT_k) and (EDT_k, LDT_k) as the time windows associated with the pickup and delivery time for customer k , respectively. It is possible to define these time windows on the basis of $[L(D_k, P_k)]/v$, WS and MRT_k in several different ways, each method having its benefits and drawbacks as discussed in Diana and Dessouky (2004). In this paper, we use the following method to compute the time windows (see Figure 1).

$$\begin{aligned}
 LPT_k &= EPT_k + WS \\
 EDT_k &= EPT_k + [L(P_k, D_k)]/v \\
 LDT_k &= EPT_k + MRT_k = EPT_k + a + b [L(P_k, D_k)]/v
 \end{aligned}$$

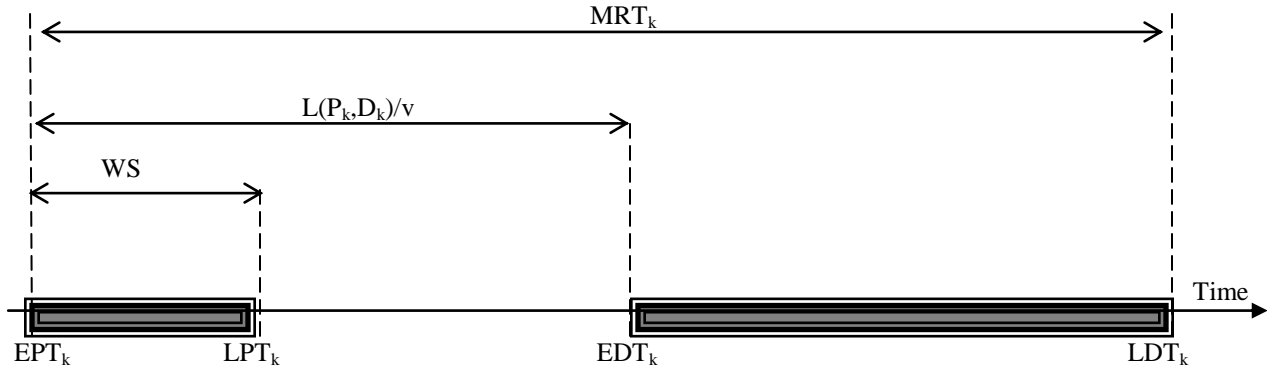


Figure 1: Definition of Time Windows

A Model for Estimating the Required Number of Vehicles

Expected Number of Vehicles

We have a list of n requests scattered in a service area. Our objective is to estimate the number of vehicles needed to serve these request using the DRT system introduced earlier.

Let r_m be the probability of serving a set of m requests out of the n total requests with the same vehicle. By the above definition of the time windows, $r_1=1$; that is, each request can be satisfied if assigned to a vehicle. If for example we state that each vehicle cannot serve more than two requests, then there will be on average $(n/2) r_2$ vehicles that serve two requests and $n(1-r_2)$ that serve the remainder. The expected total number of vehicles $E(z)$ needed to serve n requests is thus

$$E(z) = (n/2) r_2 + n(1-r_2)$$

Now, if we suppose that each vehicle can serve three requests, there will be on average $(nr_3)/3$ vehicles that serve three requests, $[nr_2(1-r_3)]/2$ that serve two requests (where $r_2(1-r_3)$ is the joint probability of serving two requests with a vehicle that could not serve the three requests) and finally, $n(1-r_2)(1-r_3)$ that serve only one request. Thus, the expected number of vehicles is

$$E(z) = (nr_3)/3 + [nr_2(1-r_3)]/2 + n(1-r_2)(1-r_3)$$

The expected number of vehicles needed to serve n requests can be computed generalizing the above equation:

$$E(z) = n \sum_{i=1}^n r_i \prod_{j=i+1}^m (1-r_j) \tag{1}$$

It should be noted that the succession of the probabilities r_1, r_2, \dots, r_n rapidly converges to zero so that we need to determine only the first m values, with $m \ll n$.

The Probability of Serving m Requests with One Vehicle

The General Case

From the definition of our problem, if one vehicle has to serve m requests it will have to visit $2m$ nodes (m pickups and m deliveries). Theoretically, there are $(2m)!$ possible visiting sequences and we compute the probability associated with each one. If we assume that the fleet dispatching process seeks for cost minimization, then the scheduler would choose the visiting sequence that maximizes the possibility of serving all the m requests. It follows that r_m would simply be the maximum of all the probabilities of success that are associated to the $(2m)!$ possible visiting sequences. However, the presence of the pairing constraints (each pickup point must be visited before the corresponding delivery point) limits the number of feasible sequences (that is, of the sequences that have probability greater than zero) to $(2m)!/2^m$.

Focusing our attention on the easiest case, that is for $m = 2$; we want to compute the probability of success in serving any pair of requests (say, 1 and 2) among the n requests waiting to be served with one vehicle. The vehicle must then visit four nodes: the pickup and delivery point of the first and of the second request each one having the above defined time window. If we denote these points with P_1, D_1, P_2 and D_2 respectively, considering the pairing constraint, the feasible sequences are:

$$P_1D_1P_2D_2, P_1P_2D_1D_2, P_1P_2D_2D_1, P_2D_2P_1D_1, P_2P_1D_2D_1, P_2P_1D_1D_2.$$

Now, assuming that r_2 is equal to the probability of realizing the most likely sequence among the above six, each sequence is determined by three different events; for example, the first one is feasible if and only if we can serve first P_1 and then D_1, D_1 and then P_2 , and P_2 and then D_2 .

Assumption 2

We assumed that the location of any point is not related to the location of all the others, the travel times of these three events are independent. However, the arrival time at P_2 is dependent on the travel time of the first two legs. In order to simplify the computation of the joint probability of the realization of the above sequence (P_1 to D_1 to P_2 to D_2), we assume that it is the product of the probabilities of the single events. This assumption of independence of the events related to a sequence overlooks the links between the arrival time at a node and the departure time from the same node. It may be a more severe limitation as the time window width decreases and the vehicle is running late.

We refer to the probabilities of the single events in a sequence as “elementary probabilities” and pd_{ij} denotes the probability of success in visiting the pickup point p of request i and then the delivery point d of request j . This same indication goes for dp_{ij}, pp_{ij} and dd_{ij} . By so doing, r_2 can be expressed as:

$$r_2 = \max \{pd_{11} dp_{12} pd_{22}, pp_{12} pd_{21} dd_{12}, pp_{12} pd_{22} dd_{21}, pd_{22} dp_{21} pd_{11}, pp_{21} pd_{12} dd_{21}, pp_{21} pd_{11} dd_{12}\} \tag{2}$$

Considering the definition of the time windows, we set $pd_{ii} = 1$ for every i . To determine all the other elementary probabilities, we proceed as follows. We will extensively show the procedure of computing dp_{ij} for a case and give only the results for the other three cases since the steps are very similar.

Since the nodes D_i and P_j have a time window, the vehicle can serve both only if the travel time $L(D_i, P_j)/v$ between them is within a certain range. The upper limit of this range is reached if the vehicle visits D_i at the earliest time and P_j at the latest. If the vehicle is not allowed to idle, there is also a lower limit represented by the trip duration when the vehicle visits D_i at the latest time and P_j at the earliest time.

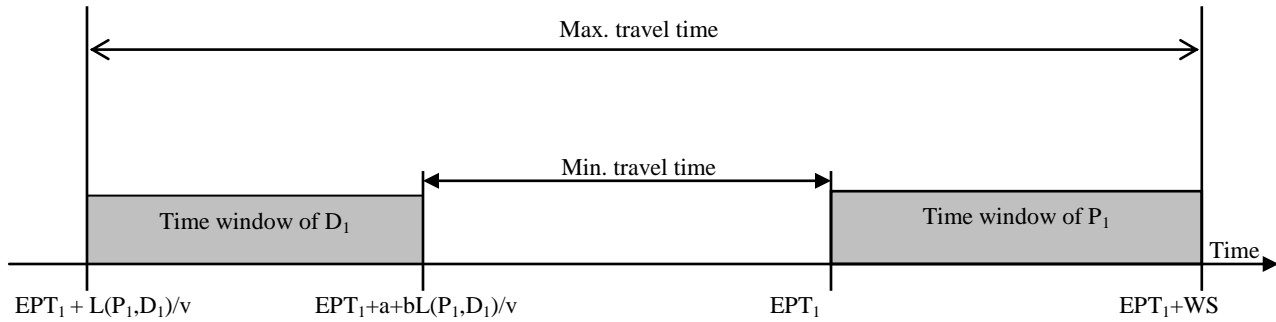


Figure 2: Computation of the Elementary Probability pd_{ij}

Figure 1 shows the relationship between these two limits and the time windows. The following inequalities translate this graphical relationship into a mathematical expression:

$$EPT_j - [EPT_i + a + b \cdot L(D_i, P_j)/v] \leq L(D_i, P_j)/v \leq EPT_j + WS - [EPT_i + L(D_i, P_j)/v]$$

The above can be expressed as:

$$\begin{aligned} v \cdot [EPT_j - EPT_i - a] &\leq L(D_i, P_j) + b \cdot L(P_j, D_i) \\ L(D_i, P_j) + L(P_j, D_i) &\leq v \cdot [EPT_j - EPT_i + WS] \end{aligned} \tag{3}$$

For the other three elementary probabilities, the procedure is the same and only the time windows need be changed. The probability intervals associated with each of these elementary probabilities are given below.

$$pp_{ij} : v \cdot [EPT_j - EPT_i - WS] \leq L(P_i, P_j) \leq v \cdot [EPT_j - EPT_i + WS] \tag{4}$$

$$\begin{aligned} pd_{ij} : v \cdot [EPT_j - EPT_i - WS] &\leq L(P_i, D_j) - L(P_i, D_i) \\ L(P_i, D_j) - bL(P_i, D_i) &\leq v \cdot [EPT_j - EPT_i + a] \end{aligned} \tag{5}$$

$$\begin{aligned} dd_{ij} : v \cdot [EPT_j - EPT_i - a] &\leq L(D_i, D_j) + bL(P_i, D_i) - L(P_j, D_j) \\ L(D_i, P_j) + L(P_i, D_i) - bL(P_j, D_j) &\leq v \cdot [EPT_j - EPT_i + a] \end{aligned} \tag{6}$$

In the above expressions, EPT_i , EPT_j and $L(\bullet, \bullet)$ are random variables, while v , a , b and WS are constants.

Problem Reduction

On the basis of equations [3 to [6], given the distributions of the random variables, it is theoretically possible to define the associated probability intervals, and thus, the values of the elementary probabilities. However, in order to have a computationally tractable problem, it is necessary to make some simplifying assumptions concerning these distributions.

Assumption 3

In the following, we will assume that EPT_i , EPT_j and $L(\bullet, \bullet)$ are drawn from the same distribution for every value of the argument. Another issue concerns the number of times we need to apply this procedure to compute all the elementary probabilities and the number of sequences to be considered when m is increasing.

From expression [2] we observe that we need to repeat the above procedure for computing these eight elementary probabilities eight times, viz: dp_{12} , pp_{12} , pd_{21} , dd_{12} , dp_{21} , pp_{21} , pd_{12} and dd_{21} . Furthermore, as m increases the number of elementary probabilities that need to be computed explodes.

Computational Experiments

Distributions of the Time Intervals between Pickup Times

In order to implement our model, we need to specify the probability density function of the time intervals between two successive pickup times, $f(g)$. We assume Poisson arrivals, hence, it follows that $f(g)$ is an exponential distribution with parameter $\lambda = 1/E(g)$.

Distribution of the Leg Lengths in A Vehicle Route

We also need to specify the probability density function $f(L(\bullet, \bullet))$ of the distance between two successive points in a route served by one vehicle. We assume a complete randomness for the spatial point pattern. This implies that the number of service points in any planar region with area A follows a Poisson distribution with parameter $\lambda' = N/A$, where N is the number of service points, and that the point coordinates are an independent random sample from a uniform distribution. Even under this assumption, it is not straightforward to represent the probability density function of the distance between two successive points of a vehicle route.

We shall start by illustrating the procedure for computing $E(L(\bullet, \bullet))$. Consider the case limit in which there are no time windows and ignoring the precedence constraints, the problem is reduced to a standard Traveling Salesman Problem (TSP). Previous research showed that when the number of points, p , is large, the length L_T of a TSP tour, assuming Euclidean metric and a square area A , is

$$L_T \approx 0.75 \sqrt{A} \sqrt{p}.$$

Assuming the vehicles are routed like in a TSP in our problem, if each vehicle serves m requests then it has to visit m pickup and m delivery locations starting and coming back to the depot. So, the expected length of the tour would be $L_T \approx 0.75 \sqrt{A} \sqrt{(2m+1)}$

Experimental Design and Results

Consider a square area of 10 x 10 miles and a planning period of 2 hours (a short planning period is used in this illustration since we are focusing on determining the fleet size during the peak period). The complete spatial randomness assumption implies that the pickup and delivery points are independently and uniformly distributed over the square area. In both cases, we used the above specified distributions of g and $L(\bullet, \bullet)$ and we varied the number of requests from 12 to 120 (corresponding to a mean value of g ranging from 10 to 1 minute).

Also, we considered different time windows since in our DRT system the time window width directly affects the quality of the service; thus, this sensitivity analysis is the key to assess the trade off between a higher quality of service and the corresponding increase of the costs, in terms of a greater number of vehicles needed.

In order to simplify the presentation of our results and their subsequent analysis, we have allowed the vehicle to idle in the following but we will keep $b = 1$. Appendix 1 shows that we could as well consider a system in which either $b > 1$ or the vehicles are not allowed to idle. Furthermore, we set $a = WS$, implying that the pickup and the delivery time windows are the same for all the requests.

The lognormal model we introduced in section “Distribution of the Leg Lengths in a Vehicle Route” seems to satisfactorily approximate $f(L(\bullet, \bullet))$ in those cases. Figure 3 shows the plots of the sampled distribution of the leg lengths when there are 120 requests and time windows of 10 and of 30 minutes, together with their respective approximation when $m = 7$.

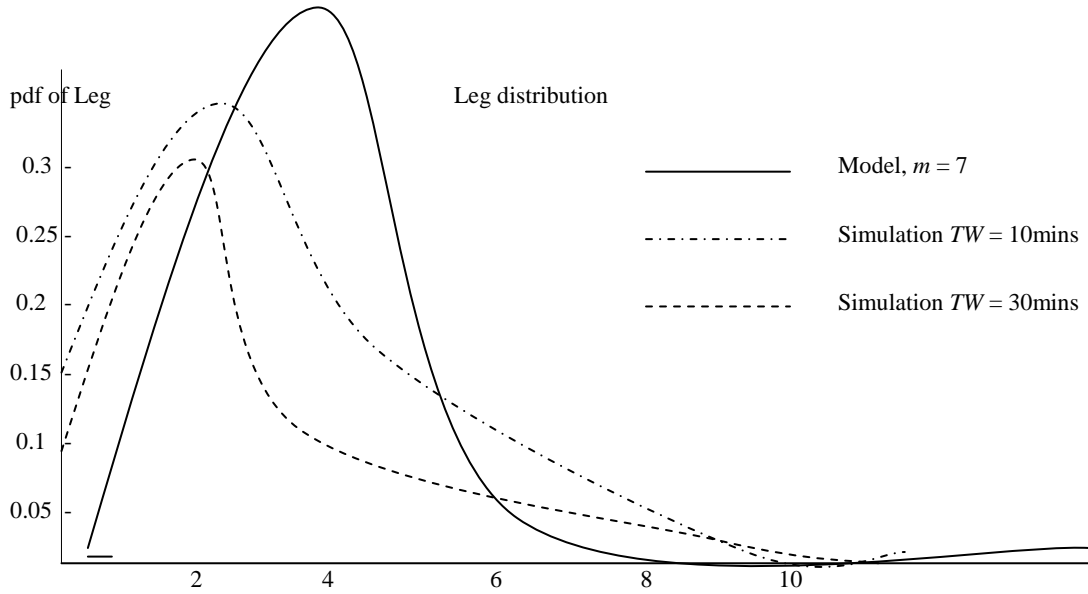


Figure 3: Distribution of the Leg Lengths from the Simulation and the Model when 120 Requests are Serviced and the Time Window Width TW is 10 and 30 Minutes

As mentioned in subsection “The Probability of Serving m Requests with One Vehicle”, when we compute the probabilities of serving 2, 3, ..., m requests with one vehicle, we do not need to fix $m = n$. In fact, the values for r_i are decreasing when i increases, and the corresponding counter j in equation [1] become less and less influential on the value of $E(z)$. Thus, we compute the series of probabilities r_2, r_3, \dots, r_m until the expected number of vehicles cannot be changed by the counter related to the probabilities r_{m+1}, \dots, r_n . However, we also tested a more stringent stopping criterion. Since the scheduling horizon is of two hours, we impose that the time needed to serve m requests with a vehicle must be less than the deadline of delivering the latest request; hence, we compute the probabilities only until m satisfies the inequality:

$$2m/v E(L(\bullet, \bullet)) < m \cdot E(hg) + TW + 1/v E(L(\bullet, \bullet)) \tag{7}$$

Where $E(hg)$ is the expectation of the $f(hg)$ distribution, v is the speed of the vehicles and TW is the time window length.

In order to benchmark our planning model, we compare it to a simulation approach that requires determining the complete daily schedule. In the simulation model, requests were generated that followed the above mentioned distributions. The requests were scheduled using a parallel regret insertion algorithm (Diana and Dessouky). The regret insertion method allowed us to find the minimum number of vehicles required to service all the requests. In order to do this, we performed the first run of the algorithm with a very high number of vehicles, and later progressively lowered this number in the successive runs until some requests could not be scheduled. The result of the computational experiments and the simulation are shown in Table 1.

N.B: When computing r_2, \dots, r_m , we noticed that the most likely sequence is always the following one: $P_1 D_1 P_2 D_2 \dots P_m D_m$. This consistently occurred throughout the entire experimental plan. However, it depends on the value of the elementary probabilities and cannot be shown to be a general rule. Nevertheless, when solving these problems we could always take the above sequence as the most likely.

PRESENTATION AND ANALYSIS OF RESULTS

Analysis of Model Results on the Required Number of Vehicles

Table 1: Number of Required Vehicles from the Model and from the Simulation (In Brackets) when the Demand Density and the Time Windows Width is Changing

Demand Density	Time Window			
	10 minutes	15 minutes	20 minutes	30 minutes
120 requests	22.9 (21.2)	18.7 (18.4)	17.6 (16.0)	17.2 (13.4)
60 requests	11.8 (13.0)	9.5 (11.0)	8.9 (9.6)	8.6 (8.0)
24 requests	5.2 (7.2)	4.1 (6.2)	3.7 (5.8)	3.4 (4.4)
12 requests	2.9 (4.4)	2.2 (3.2)	1.8 (3.2)	1.7 (2.6)

It can be seen that the difference between the model results and those from the simulation (shown in brackets) is almost always less than two vehicles. Only when we have to serve 120 requests and the time windows are of 30 minutes do we overestimate the number of needed vehicles by a little more than 2.1. This could be due to the approximation of the leg lengths that we used. We believe that the model presented in this paper would outperform the simulation model whenever the demand density is greater than 120 requests for any time window width.

Analysis of Expected Leg Length

Table 2: Expected Leg Lengths from the Model and from the Simulation (in Brackets) when the Demand Density and the Time Windows Width are Changing

Demand	Time Window			
	10 minutes	15 minutes	20 minutes	30 minutes
120 requests	3.97 (3.51)	3.89 (3.27)	3.81 (3.03)	3.72 (2.70)
60 requests	4.08 (3.96)	3.99 (3.62)	3.90 (3.32)	3.81 (2.98)
24 requests	4.30 (4.14)	4.20 (4.12)	4.16 (3.93)	3.99 (3.80)
12 requests	4.45 (4.58)	4.22 (4.33)	4.22 (4.33)	4.09 (3.78)

Table 2 presents the values of $E(L(\bullet, \bullet))$ computed from equation [7] and those derived from the schedules of the simulation. It can be seen that there was overestimation in the expected value of the leg lengths with larger time windows when we have to serve a higher number of requests. In particular, when we have 120 requests and the time windows are of 30 minutes, an estimate of a mean leg length of 3.72 miles was made, whereas the one from the simulation is 2.70 miles. The reason for the overestimation in this case is because there is a significant amount of ridesharing when there are a large number of requests and a wide time window. With a significant amount of ridesharing our approach in equation [7] under-weights the trips that follow the TSP tour versus those that have a longer expected length of 0.52 = 5.2 miles. In this scenario, only a small number of trips go directly from pickup to delivery. Thus, only a small fraction of trips have a mean length of 5.2 miles. This suggests deriving a new weighting scheme in equation [7] when there is a significant amount of ridesharing.

Considering that a better approximation of $f(L(\bullet, \bullet))$ would even improve the model results, it is believe that the proposed methodology is an effective way to quickly estimate the number of vehicles needed to provide a DRT under a fairly broad range of cases (systems with different levels of demand and different quality requirements).

Viability of Two Possible Approximations for the Distribution of the Leg Lengths in a Demand Responsive Service Vehicle Route

The first approach is to consider the distribution of the distance between any two random points in the service area $f(d)$ that can be obtained from the distributions of the coordinates of the points through convolution. Another possible strategy is to consider the s^{th} nearest-neighbor distance density $f(d_s)$ from a given point in the service area. In other words, either the distance d between any two points in the service area or between a point and its s^{th} nearest neighbor, d_s , could be used as an approximation of the vehicle route $L(\bullet, \bullet)$.

Under the complete spatial randomness hypothesis, some results related to the distributions $f(d)$ and $f(d_s)$ are available in the published literature. Christofides and Eilon (1969), and Eilon, Watson-Gandy and Christofides (1971) derived the expected distance between two random points for different shapes of the service area. The distribution of the distances $f(d)$, their mean $E(d)$ and variance $VAR(d)$ for the cases of points uniformly scattered over a unit service area are reported in Table 3. Spatial analysis textbooks such as Mathai (1999) reports the probability density function $f(d_s)$ of the nearest, second-nearest, ..., s^{th} nearest point, assuming Poisson arrivals in a plane. However, if we consider a finite area, there are boundary effects that alter these latter distributions. In fact, we would expect that the value of d_s is greater and is increasing when we consider points that are nearer the edge. Considering the case of the nearest neighbor ($s=1$) and the related distribution of the distance d_1 , Donnelly (1978) determined correction terms through simulation for $E(d_1)$ and $VAR(d_1)$, that are sufficiently accurate when there are more than seven points and the shape of the region is sufficiently smooth. We report the expressions for $f(d_1)$, $E(d_1)$ and $VAR(d_1)$ both considering and not considering edge effects in Table 4.

Table 3: Distribution of the Distances between any Two Points that are Uniformly Scattered in a Unit Square

$f(d)$	$\begin{aligned} & 2\pi d - 8d^2 + 2d^3 & : 0 \leq d \leq 1 \\ & 4d \arcsin((2-d^2)/d^2) + 8d\sqrt{(d^2-4d-1)-2d^2} & : 1 < d \leq \sqrt{2} \\ & 0 & : \text{otherwise} \end{aligned}$
$E(d)$	0.5214
$VAR(d)$	0.0615

Table 4: Distribution of the Nearest-Neighbor Distance from a Given Point in a Unit Square (Assuming Poisson Arrivals of N Points)

	Options	
	Not considering edge effects	Including Donnelly (1978) correction terms
$f(d_i)$	$2\pi N d_i e^{-\pi N d_i^2} \quad d_i \geq 0$	-
$E(d_i)$	$\frac{1}{2}\sqrt{N}$	$\frac{1}{2}\sqrt{N} + \frac{1}{N}(0.0514 + 0.041/\sqrt{N})$
$VAR(d_i)$	$(4 - \pi) / 4\pi N$	$(4 - \pi) / 4\pi N + 0.037/N\sqrt{N}$

In order to check the possibility of approximating $f(L(\bullet, \bullet))$ through either $f(d)$ or $f(d_i)$, we ran some simulations on standard problems (pickup and delivery points uniformly scattered in a unit square area). To schedule the vehicles in the simulation, we used a parallel regret insertion heuristic (Diana and Dessouky). The results showed that $E(L(\bullet, \bullet))$ is about 20% to 40% less than the value of $E(d)$ indicated in Table 3, and the gap increases when the time windows are larger and the density of the requests is higher. This is rather an intuitive result since relaxing the scheduling constraints leads to more efficient routing of the vehicles. On the other hand, $E(d_i)$ seriously underestimates $E(L(\bullet, \bullet))$ when there are more than 3-4 requests to serve and the time window width is not too loose even if we include the correction terms proposed by Donnelly. Also, the shape of the sampled distribution is quite different from those reported in Tables 3 and 4.

To sum up, the considered approximations have proven to be rather poor for our purposes. When scheduling the service, the vehicles are normally dispatched to the “best” point that satisfies the time, precedence, and coupling constraints in order to increase the efficiency of the system. The definition of “best” point obviously depends on the heuristic used to schedule the service, but in any case this is rather unlikely to be either a random or the nearest-neighbor point. In this case, the regret insertion algorithm tries to anticipate the insertion of requests that

could be difficult to insert in a later stage of the process, as explained in detail in Diana and Dessouky. One could argue that using a nearest-neighbor-based heuristic to schedule the service could allow for a better approximation of $f(L(\bullet, \bullet))$ through $f(d_1)$. However the inferiority of the nearest-neighbor rule over an insertion-based algorithm when we consider a routing problem with time windows is an established result (Solomon, 1987).

SUMMARY OF THE STUDY

This paper has presented a continuous approximation model to forecast the number of vehicles needed to operate a demand responsive transit service. In contrast with current mathematical programming techniques, our approach simply requires the knowledge of the demand density over the service area since it may be hard or even impossible to have more detailed data in the planning phase. Computational results showed that the proposed methodology can provide reliable result under different circumstances. A critical point is the approximation of the distribution of the leg lengths that can alter the results when the time windows are wide, and further research is needed at this point in order to improve the performances of the proposed model.

SUMMARY OF THE FINDINGS

The interest in using an approximation model lies in the possibility of the planner to perform sensitivity analysis through the construction of several different scenarios. In this way, the choice of the best compromise between quality of service and financial resources is much more effective. The problem presented in this paper, as described in section 3, is sufficiently general to envision the application of our methodology in different contexts, particularly the problems of distribution of goods in which there are severe time constraints. Another useful generalization of the present work might be the inclusion of the proposed methodology in a demand-supply equilibrium model for a general DRT system, similar to what was proposed by Chang and Schonfeld (1991) and Chang and Lee (1993) for the specific case of a deviation service.

CONCLUSION AND RECOMMENDATION

We believe that an application of particular interest of this model is the study of the tradeoff between the number of vehicles needed and the time windows associated with the locations, in analogy with what is shown in Table 1. This is a research field that deserves more attention and that may be a key issue in developing DRT services that are more cost-effective but still satisfying for the customers, hence, it is highly recommended for further study to serve as antidote to reducing, to the barest minimum (or non-existing), fleet operations problems when noticed or envisaged to occur (a proactive-approach) especially in organizations that have big supply-chain of customers to satisfy with their products or services as applicable.

AUTHOR INFORMATION

Emmanuel Olateju Oyatoye is a Senior Lecturer in the Department of Business Administration of the University of Lagos, Akoka, Lagos, Nigeria. He holds a B.Sc. (Hons.) degree in Mathematics and Statistics; an M.Sc. degree in Applied Mathematics (Computational Methods in Optimization); and a Ph.D. degree in Management Sciences (Operations Research). His current research focus is using the Analytic Hierarchy Process and Analytic Network Process in addressing effective health care delivery problem in Nigeria; poverty alleviation among Nigerian women; and the problem of the development of the Niger Delta region of Nigeria. He is married with five children.

Magbagbeola Joshua Adekunle O, Lecturer at Joseph Ayo Babalola University (JABU), Lecturer at DLI (UNILAG), Research Fellow with Dept. Of Bus Admin (UNILAG). He bagged BSc (Maths)Ed, PGD (Computer Science), MBA & MSc (Operations Research) all from University of Lagos, Akoka. Currently MTech/PhD student at Federal University of Technology (FUTA), Akure. He is an Associate Fellow (AFIOR) and member of Technical Committee of INFORN (Institute for Operations Research of Nigeria). Member German Operation Research Society (GOR), AMNIM. Conference Speaker, Publisher and Author. Area of interest is Mathematical Modelling with emphasis on Transportation Modelling. Also Pastor with CAC (Worldwide) and happily married.

REFERENCES

1. Adebisi, O. and Hurdle, V.F., "Comparing fixed-route and flexible-route strategies for intra-urban bus transit"; *Transportation Research Record* 854, Transportation Research Board, National Research Council, Washington D.C., pp 37-43, 1982.
2. Aldaihani, M.M., Quadrioglio, L., Dessouky, M. and Hall, R.W., "Network design for a grid hybrid transit service"; *Transportation Research*, 38A, 511-530, 2004.
3. Arrillaga, B. and Medville, D.M., "Demand, supply and cost modeling framework for demand-responsive transportation systems"; *Transportation Research Board Special Report* 147, pp.32-47, 1974.
4. Arthur, W.B., J.H. Holland, B. LeBaron, R. Palmer, and P. Tayler, "*The Economy as an Evolving Complex System II*"; Chapter Asset pricing under endogenous expectations in an artificial stock market, pages 15-44, Addison-Wesley, Reading, MA; 1997.
5. Bagni, R., R. Berchi and P. Cariello, "A comparison of simulation models applied to epidemics," *Journal of Artificial Societies and Social Simulation*, Vol. 5, No. 3, 2002. <http://jasss.soc.surrey.ac.uk/5/3/5.html>.
6. Batchelder, J.H. and Kullman, B.C., "Analysis of integrated urban public transportation systems"; *Transportation Research Record* 639, Transportation Research Board, National Research Council, Washington D.C., 25-29, 1977.
7. Beardwood, J., Halton, J.H. and Hammersley, J.M., "The shortest path through many points"; *Proceedings of the Cambridge Philosophical Society*, 55, pp 299-327, 1959..
8. Christofides, N. and Eilon, S., "Expected distances in distribution problems"; *Operational Research Quarterly*, 20, pp 437-443, 1969.
9. Daganzo, C.F., "An approximate analytic model of many-to-many demand responsive transportation systems"; *Transportation Research*, 12, 325-333, 1978.
10. Daganzo, C.F., "The length of tours in zones of different shapes"; *Transportation Research B*, 18B, 135-145, 1984a.
11. Daganzo, C.F., "Checkpoint dial-a-ride systems"; *Transportation Research B*, 18B, 318-327, 1984b.
12. Daganzo, C.F., "Modeling distribution problems with time windows: Part I"; *Transportation Science*, 21, pp 171-179, 1987.
13. Diana, M. and Dessouky, M.M., "A new regret insertion heuristic for solving large-scale dial-a-ride problems with time windows"; *Transportation Research B*, 38B, pp 539-557, 2004.
14. Donnelly, K.P., "*Simulations to determine the variance and edge effect of total nearest neighbor distance*"; *Simulation Studies in Archaeology* ed I. Hodder, pp. 91-95, Cambridge University Press, Cambridge, 1978.
15. Eilon, S., Watson-Gandy, C. and Christofides, N., "*Expected distances in distribution problems*", *Distribution Management: Mathematical Modeling and Practical Analysis*, pp. 151-179, Griffin, London, 1971.
16. Emonet, T., C. M. Macal, M. J. North, C. E. Wickersham and P. Cluzel, "AgentCell: A Digital Single-Cell Assay for Bacterial Chemo taxis"; *Bioinformatics*, Vol. 21, No. 11, pp. 2714-2721, 2005.
17. Flusberg, M. and Wilson, N.H.M., "A descriptive supply model for demand responsive transportation system planning"; *Proceedings of the 17th Annual Meeting of the Transportation Research Forum*, pp 425-431, 1976.
18. Hall, R.W., "Pickup and delivery systems for overnight carriers"; *Transportation Research A*, 30A, pp 173-187, 1996.
19. Jaw, J.J., Odoni, A.R., Psaraftis, H.N. and Wilson, N.H.M., "A heuristic algorithm for the multi-vehicle many-to-many advance request dial-a-ride problem"; *Transportation Research B*, 20B, pp 243-257, 1986.
20. Larson, R.C. and Odoni, A.R., "*Urban Operations Research*"; Prentice-Hall, Englewood Cliffs, 1981..
21. Lerman, S.R. and Wilson, N.H.M., "Analytic model for predicting dial-a-ride system performance"; *Transportation Research Board Special Report* 147, pp 48-53, 1974.
22. Mathai, A.M., "*The nearest neighbor problem on a plane from Poisson arrivals of random points*"; *An Introduction to Geometrical Probability*, Gordon and Breach, Amsterdam, pp. 223-228, 1999.
23. Macal, C., "Effects of global information availability in networks of supply chain agents"; in *Proc. Agent 2003: Conf. on Challenges in Social Simulation*, Eds., D. Sallach, C. Macal and M. North, Chicago, Oct. 2-4, pp. 235-252, Argonne National Laboratory, 2003.

24. Macal, C., "Emergent structures from trust relationships in supply chains"; in Proc. Agent 2004: Conf. on Social Dynamics: Interaction, Reflexivity and Emergence, Eds., C. Macal, D. Sallach and M. North, Chicago, Oct. 7-9, pp. 743-760, Argonne National Laboratory, MASON,2005
<http://cs.gmu.edu/~eclab/projects/mason/>.
25. Quadrifoglio, L, Hall, R. W., and Dessouky, M. M., "Design of Mobility Allowance Shuttle Transit (MAST) Services: bounds on the maximum longitudinal velocity"; under revision *Transportation Science*, 2005.
26. Solomon, M. M., "Algorithms for the Vehicle Routing and Scheduling Problems with Time Windows Constraints; *Operations Research*, 35, pp. 254 – 265, 1987.
27. Stein, D.M., "Scheduling dial-a-ride transportation problems"; *Transportation Science*, 12, pp 232-249, 1978a.
28. Stein, D.M., "An asymptotic probabilistic analysis of a routing problem"; *Mathematics of Operations Research*, 3, pp 89-101, 1978b.
29. Ward, D.E., "Theoretical comparison of fixed route bus and flexible route subscription bus service in low density areas", Transportation Systems Center, USDOT, 1975.
30. Wilson, N.H.M. and Hendrickson, C., "Performance models of flexibly routed transportation services"; *Transportation Research B*, 14B, pp 67-78, 1980.

APPENDIX 1: COMPUTATION OF THE ELEMENTARY PROBABILITIES

Distribution of the Intervals between Adjacent Pickup Times

Let us define a ranking order for the list of requests by ascending EPT. Define a random variable g that represents the temporal gap between the two earliest pickup times of the requests k and $k + 1$, if k is the index of this ranking order:

$$g = EPT_{k+1} - EPT_k$$

In the general case, if two requests are at the k^{th} and at the $(k + h)^{th}$ place according to this order, we can say that the temporal gap between their respective earliest pickup times can be represented by the distribution of the random variable

$$\sum_1^n g$$

For simplicity, we will refer to this random variable as hg . Since h is not a constant (that is, the number of demand points is a random variable that follows a discrete distribution), we point out that the associated probability density function $f(hg)$ is not equal to $h.f(g)$. It is now possible to redefine the elementary probabilities introduced in section “A Model for Estimating the Required Number of Vehicles” as a function of these interval gaps. Considering again the case $m = 2$, we have for example that for the first of the above listed six feasible sequences, the vehicle should start visiting the pickup and the delivery node of the same request. The difference of order between the requests related to the two nodes is obviously 0 since both are from the same request. So in this case $h = 0$ and we can denote the corresponding probability as pd_0 . When leaving D_1 , the vehicle must arrive to P_2 . We assume that the time gap between the two requests is $(n/2).g$. We denote the corresponding elementary probability with $dp_{n/2}$, which is the probability of picking up the $(k + n/2)^{th}$ request after delivering the k^{th} request. Defining all remaining elementary probabilities in the same way, equation [2] can be rewritten as:

$$r_2 = \max. \{ pd_0 dp_{n/2} \quad pd_0 pp_{n/2} \quad pd_{-n/2} dd_{n/2} \quad pp_{n/2} pd_0 \quad dd_{-n/2} pd_0 \quad dp_{-n/2} pd_0 \quad pp_{-n/2} pd_{n/2} \quad dd_{-n/2} pp_{-n/2} \quad pd_0 dd_{n/2} \} \tag{1}$$

Nine different elementary probabilities appear in this formula. In the general case (that is, for m greater than 2), the interval gap is $c.(n/m).g = hg$, where c is an integer constant comprised between $-(m - 1)$ and $+(m - 1)$. The number of elementary probabilities that must be determined is linearly increasing with m , being equal to $(8m - 7)$. They are the following:

$$pp_{-(m-1).n/m}, \dots, pp_{-n/m}, pp_{n/m}, \dots, pp_{(m-1).n/m}, pd_{-(m-1).n/m}, \dots, pd_{-n/m}, pd_0, pd_{n/m}, \dots, pd_{(m-1).n/m}, dp_{-(m-1).n/m}, \dots, dp_{-n/m}, dp_{n/m}, \dots, dp_{(m-1).n/m}, dd_{-(m-1).n/m}, \dots, dd_{-n/m}, dd_{n/m}, \dots, dd_{(m-1).n/m}.$$

We can see that h assumes a value of zero only when there is a request that is served without deviations; in this case the corresponding elementary probability pd_0 is 1 by definition. In all the other cases (that is, when $h \neq 0$), the definition of the elementary probabilities as a function of h allows us to compute more easily their values. In fact, assuming $b = 1$, equations [3] to [6] can be rewritten in this manner:

$$\begin{aligned} dp_n &: v.(hg - a) \leq L(D_i, P_j) + L(P_i, D_i) \leq v.(hg + WS) \\ pp_n &: v.(hg - WS) \leq L(D_i, P_j) \leq v.(hg + WS) \\ pd_n &: v.(hg - WS) \leq L(D_i, P_j) - L(P_i, D_i) \leq v.(hg + a) \\ dd_n &: v.(hg - a) \leq L(D_i, P_j) + L(P_i, D_i) - L(P_j, D_j) \leq v.(hg + a) \end{aligned}$$

Assuming $b = 1$ is useful to transform for each elementary probability the two inequalities in a probability interval; in order to have only one interval in the above equations, we could alternatively allow the vehicle to idle at every node. As discussed in section “A Model for Estimating the Required Number of Vehicles”, the lower bound would disappear in this case (if the travel time is too short, then the vehicle could wait at the second point until the

time window is met) and the assumption of $b = 1$ could be relaxed.

It is reasonable to assume that the random variables $L(D_i, P_j)$ and (hg) are independent, as well as $[L(D_i, P_j) + L(P_i, D_i)]$ and (hg) , $[L(D_i, P_j) - L(P_i, D_i)]$ and (hg) and $[L(D_i, P_j) + L(P_i, D_i) - L(P_j, D_j)]$ and (hg) . In the preceding equation with $f(\cdot)$ we indicated the probability density functions of the random variables $L(D_i, P_j)$, $[L(D_i, P_j) + L(P_i, D_i)]$, $[L(D_i, P_j) - L(P_i, D_i)]$, $[L(D_i, P_j) + L(P_i, D_i) - L(P_j, D_j)]$ and (hg) , that can be computed using convolution. Theoretically speaking, we could now compute the probabilities r_m . However, if m is increasing we still have the problem of the number of elements to consider in equation [1], that as aforesaid is equal to $(2m)!/2^m$ (in the general case there is one element for each feasible visiting sequence).