LONG HAUL TRUCKS DYNAMIC ASSIGNMENT WITH PENALTIES AND TIME WINDOWS

Antonio Martins Lima Filho (martins.lima@poli.usp.br)

Nicolau D. Fares Gualda (ngualda@usp.br)

Department of Transportation Engineering

Escola Politécnica, University of São Paulo

ABSTRACT

Long haul transportation system operational planning implies to solve a capacitated dynamic network optimization problem, aiming to perform the freight movements in an efficient and effective way, utilizing the available transportation capacity. This work employs an approximate adaptive dynamic programming to solve this kind of problem, introducing a network modeling to manage demands not attended, time windows and heterogeneous fleets. The adopted dynamic programming procedure involves solving each stage of the problem considering a concave estimate of the future stages value for each configuration of vehicles / locations. This approach greatly reduces the quantity of involved variables, allowing the utilization of more realistic mathematical models on a longer planning horizon. The method replaces advantageously real world empirical strategies and optimization techniques applied only to the next planning period. Results from a successful application of the model are presented.

Keywords: Long haul truck transportation, Network flow optimization, Adaptive Dynamic Programming

1. INTRODUCTION - THE PROBLEM

This work approaches a long haul trucks optimization problem, treating the assignment of vehicles to loads across time.

The system under consideration is composed of full-truck-loads (a single load on the vehicle, which is carried from a single origin to a single destination). The demand for transportation occurs in a random way, in several points in space, throughout time, and must be attended in all of the periods according to departure time limits.

A basic feature of this type of transport, defined by the expression Long Haul is the fact that vehicles belonging to the fleet managed by the firm (“own fleet”) do not return immediately to the operational base of origin after fulfilling the delivery of its loads. Thus, for each vehicle that becomes available at each point in space-time, a decision of operational character must be taken: to attend a load at that point, to wait for a future load at the location, or to displace the vehicle, unloaded, to another point.

Occurrence of this kind of long haul transportation is very frequent in large territory countries as USA and Brazil, and furthermore in economic blocs like Mercosul, where important cities are separated by thousands of kilometers and consequently travel times are in order of days and not hours. So, the importance of studies aiming at optimization of the involved resources.

Heterogeneous fleets should be considered, what leads to choosing the most economic available vehicle to take care of each specific task.

Tasks, furthermore, may be also attended by third party vehicles. In this case it is not necessary to consider the sequence of vehicle utilization, but it is important to take into account the availability of such third party vehicles in the location and period where they will be needed.

Some loads must departure on a single time period, while others may follow a “time window”. A time limit also exists and, in many cases, a penalty must be paid by the transportation firm if that limit is over passed. This penalty may be in pecuniary form or be related to lesser future incomes due to loss of customers.

This work proposes to approach these needs of the system on a new and comprehensive way, employing an approximate adaptive dynamic programming to solve the related optimization problem. In Section 2 a literature review is presented. In section 3, the methodology proposed for this particular problem solution is shown. In section 4 a typical problem is solved and section 5 presents conclusions of the research.

2. LITERATURE REVIEW

The problem of allocating vehicles to loads in long distance road transportation may be divided in two types of planning approaches: regular services, where the tactical planning is the most important one, and non-scheduled services, where the operational planning is prominent.

For regular services there is a previous definition of the way operations will be performed. The definition may take the form of intercity routes (Crainic and Roy, 1992), network optimization (Powell and Sheffi, 1989) and other strategies of routing, based on well known models like the shortest path problem, the travelling salesman problem, and so on.

The non-scheduled services, in turn, are the scope of this paper. Works on this subject include the treatment of empty cars on rail transportation (Haghani, 1989) and the allocation of vehicles to loads (Frantzeskakis and Powell, 1990, Powell et al., 1995, Hane et al., 1995, Powell and Carvalho, 1998). Godfrey and Powell (2002a, 2002b) and Topaloglu and Powell (2006) treat stochastic models for dynamic allocation problems of the kind.

2.1 Static x Dynamic Models

Operations planning models in transportation may be static or dynamic. Static models are proper for tactical planning, where the time factor may be left in the second plan; these models are useful for regular transportation services.

In long distance road transportation operational planning, dynamic allocation is a real need imposed by the characteristic of the problem: the resource (transportation equipment), after serving a task, is available at a different location from the original, in which new tasks waiting to be served may or may not exist. A dynamic model may capture the temporal stage of physical activities, but not the temporal stage of information; in this case the model is called deterministic (Powell, 2003).

Resources allocation may be considered as "dynamic" in three different senses (Powell, Jaillet and Odoni, 1995): the problem is dynamic, since information (position of the resources, availability of loads) changes with time; the model used to represent the problem is also dynamic, since it incorporates explicitly the interaction of activities over time; and, finally, the method of solution is dynamic in nature (Dynamic Programming), solving the problem repeatedly as new information is received.

2.2 Logistics Queuing Network - Dynamic Programming Approach

Powell (1986) proposes representing the problem of resources allocation in long distance transportation in the form of a space-time network (Fig. 1). Arcs represent vehicle movements between space-time nodes; dotted lines correspond to empty movements. This form of modeling produces a network of acyclic structure, to say, for each network node it is not possible to have an arc sequence turning back to this node.

This acyclic structure facilitates the operation of optimization algorithms. The difficulty with this type of network approach is the great extent of the problems involved, in terms of quantity of variables. Furthermore, any details of the operation, as time windows, involve new generations of a great quantity of variables. Therefore, models based on pure Linear Programming approaches, as presented by Haghani (1989) and Hane et al. (1995), are not suitable to solve real world dynamic problems.

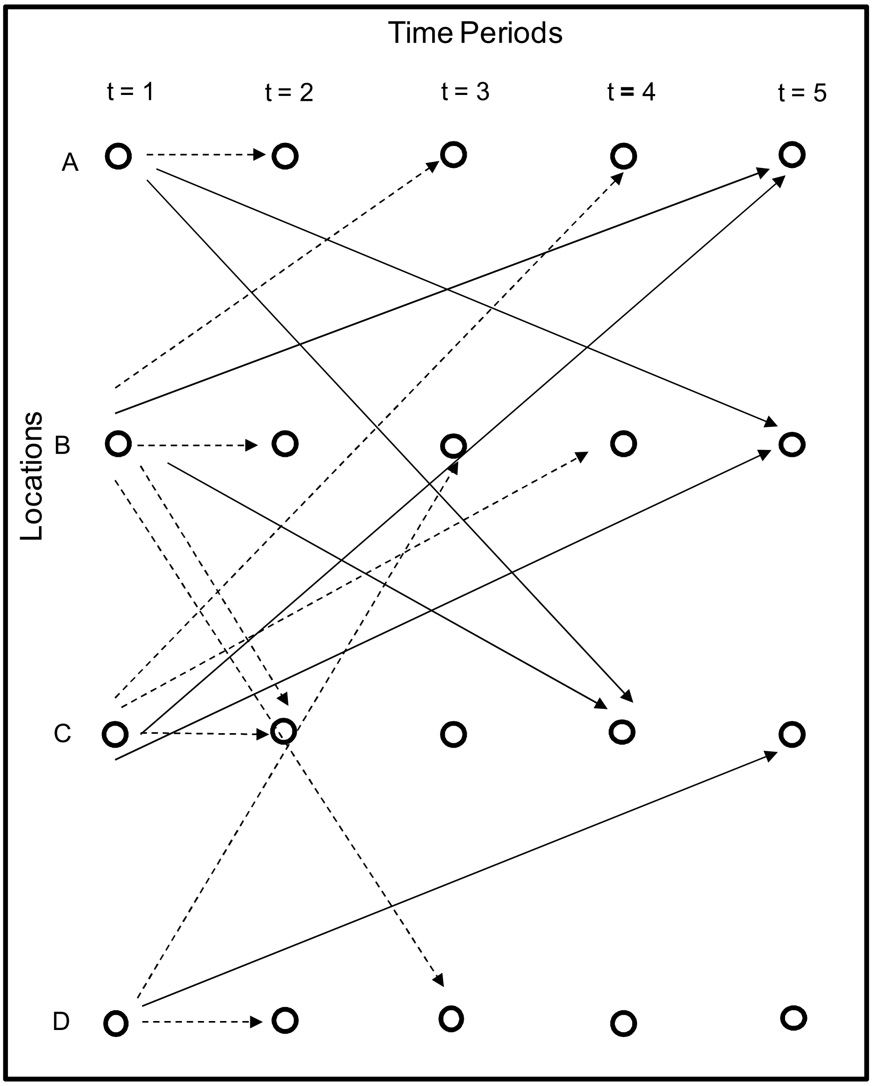

Fig. 1 - Space-Time Network

A way to bypass the above difficulties is presented by Powell et al. (1995), based on a discrete event dynamic system (DEDS), where demands are queued at terminals while waiting for available capacity. The authors call it a logistics queuing network (LQN). When a vehicle arrives at its destination, it becomes empty and available for its next movement; the flow of demands into a location may not equal the flow out of it, leading to the necessity to reposition empty vehicles from one location to another.

The purpose of this method is to replace the optimization approach based on Linear Programming with a solution by Recursive Dynamic Programming, where “recursive” is the successive application of an algorithm until a specified condition is met. The method performs the decomposition of the general solution of the network problem into a series of subproblems solutions at each location level. The general approach consists of a series of forward and backward steps across time: at each iteration, the forward step assigns vehicles to loads and, in a limited way, re-positions the vehicles; the backward step calculates gradients and an adjustment phase modifies the potential values and the limits of empty movements. The procedure continues until it ensures a desired level of convergence (Crainic, 2003).

Powell et al. (1995) and Powell and Carvalho (1998) have introduced modifications on that model and applied it to deterministic problems, utilizing a linear estimation of the value function.

2.3 Adaptive Dynamic Programming by Concave Adaptive Value Estimation

The value function depends on the vehicles quantity at a node and it is concave once at each new available vehicle the increase in value is at most the same of the precedent vehicle. The development of CAVE (Concave Adaptive Value Estimation) algorithm by Godfrey and Powell (2001) permits an accurate estimate of concave functions values based on the results of a succession of experiments without depending on the specific function of probabilities associated with the events.

The algorithm is based on generation of gradients related to the amount of available resources, taking into account the result obtained when the demand actually occurres. It allows to treat probabilistic demands. Thereafter, the algorithm executes a smoothed change in the form of the curve for the next iteration, to say, the succession of iterations is adaptive relatively to the probabilistic distribution.

Godfrey and Powell (2002a, 2002b) propose a solution focused basically in two reductions of complexity: the problem is separated by location / period (concept of logistics queuing network, as described above) and then it is solved using an estimate of the value associated with the future state with the CAVE algorithm.

Topaloglu and Powell (2006) extend the model to treat heterogeneous fleets utilizing a hybrid linear/concave estimation of the value function.

The works by Godfrey and Powell (2002a, 2002b) and Topaloglu and Powell (2006) consider that demands must be attended in a fixed time window and are lost if not attended or are attended by unlimited available chartered vehicles. Nevertheless, these works treat only homogeneous fleets or, in case of heterogeneous vehicles, the problem is simplified, not considering the decreasing value of each additional resource available at the same location (linear value function approximation).

Simão et al. (2008) addressed the problem of developing a model to simulate at a high level of detail the movements of over 6,000 drivers for Schneider National, the largest truckload motor carrier in the United States, utilizing an approximate dynamic programming technique. The goal of the model was to closely match a number of operational statistics. In addition to the need to capture a wide range of operational issues, the model had to match the performance of a highly skilled group of dispatchers while also returning the marginal value of drivers domiciled at different locations. These requirements dictated that it was not enough to optimize at each point in time (something that could be easily handled by a simulation model) but also over time.

Powell (2009) stresses the importance of utilizing Approximate Dynamic Programming (ADP) to treat resource allocation problems, including allocation of vehicles to tasks.

3. METHODOLOGY

This work applies to the Operational Planning of a Long Haul Load Transportation System, aiming to maximize the economic result of the System in function of the allocation of transportation equipments to loads. The adopted methodology is based on the Stochastic, Approximate and Adaptive Dynamic Programming Model described by Godfrey and Powell (2002a, 2002b). This methodology employs a dynamic programming model applied to each stage of the problem instead of a network global optimization procedure. A heuristic is used to estimate the value of future state functions transferred from a current stage to future ones.

As a extension of the research by Godfrey and Powell (2002a, 2002b), the present work treats not attended demands through the payment of penalties, or the utilization of third parties transportation firms to avoid the payment of such penalties. Limitation of third party vehicles availability is explicitly treated in the model. Some demands, however, may be attended in the period of appearance or in later periods, until a pre-established limit (time windows concept). As for the vehicles characteristics its heterogeneity is considered for both conformity with demand requirements and best economic allocation at each location in function of demand availability.

Given the complexity of the model, it is presented, at first, the development of a Deterministic Linear Programming model, to better illustrate the required variables and constraints. Section 3.1 is dedicated to this feature, while section 3.2 shows the solution strategy using the Adaptive Dynamic Programming model.

3.1 Deterministic Linear Programming Model

The overall optimization problem model, that is, the model to attain a solution for the entire planning horizon with no recursive feature, may be built as follows:

3.1.1 Decision Variables

Let:

T = number of planning periods in the planning horizon

$\mathcal{T}$ = {0, 1, ..., T – 1} = periods at which decisions are taken

$\mathcal{J}$ = set of physical locations in the network, indexed by i and j

$\mathcal{K}$ = set of classes of vehicles, indexed by k

$\mathcal{K}_l$ = set of classes of vehicles which can perform loads of class l, indexed by k_l (vehicle of class k may perform load of class l)

$\mathcal{L}_{ijt}$ = set of loads to be serviced at time t, with origin in i and destination in j, indexed by $l(a,b)$, with $a, b \in \mathcal{T}$, and:

a = first time period allowed for load attendance

b = last time period allowed for load attendance

$D^{l(a,b)}$ = demand (deterministic) for loads of class $l(a,b)$

$L_t^{l(a,b)}$ = quantity of loads of class $l(a,b)$ that may be performed in time t, $a \leq t \leq b$

S_{it}^{k} = total of third part vehicles, of class k, at location i, available in time t

Decision variables, for each $t \in \mathcal{T}$, $i, j \in \mathcal{J}$, may be described as:

$x_{ijt}^{l(a,b)\,k}$ = quantity of loads of class $l(a,b)$ serviced in time t ($a \leq t \leq b$) with origin i and destination j, with vehicles class k from own fleet, with reward r^{lk}

y_{ijt}^{k} = quantity of own fleet vehicles of class k relocated from i to j in time t, with cost c_{ij}^{k}

$z_t^{l(a,b)\,k}$ = quantity of loads of class $l(a,b)$, serviced in time t ($a \leq t \leq b$) with vehicles class k of third party companies, with reward s^{lk}

$m_t^{l(a,b)}$ = quantity of not attended loads of class $l(a,b)$ in time t ($t = b$), with penalty p^l

3.1.2 Objective Function

For the entire planning horizon the following expression must be maximized, encompassing the rewards for attending the loads with own and third party fleets, the cost of empty movements and the penalties incurred by not attendance of demands:

$$\sum_{t=0}^{T\text{-}1}\Bigg(\sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}}\sum_{l(a,b)\in\mathcal{L}_{ijt}}\sum_{k\in\mathcal{K}_l} x_{ijt}^{l(a,b)\,k}\, r^{l\,k} - \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}}\sum_{k\in\mathcal{K}} y_{ijt}^{k}\, c_{ij}^{k} + \sum_{i\in\mathcal{J}}\sum_{j\in\mathcal{J}}\sum_{l\in\mathcal{L}_{ijt}}\sum_{k\in\mathcal{K}_l} z_{t}^{l(a,b)\,k}\, s^{l\,k} - \sum_{l(a,b)\in\{\mathcal{L}_{ijt}|\,t=b\}} m_{t}^{l(a,b)}\, p^{l}\Bigg)$$

3.1.3 Constraints

A - Constraints related to the total attendance of each demand, considering own and third party fleets, classes of vehicles and penalties

For a given load demand:

$$\sum_{t=a}^{b}\left(\sum_{k\in\mathcal{K}_l} x_{ijt}^{l(a,b)\,k} + \sum_{k\in\mathcal{K}_l} z_{t}^{l(a,b)\,k} + m_{t}^{l(a,b)}\right) = D^{l(a,b)}$$

B - Constraints related to flow conservation for own fleet vehicles

For a class of vehicles k, a given location i and a time period t:

$$\sum_{j}\sum_{l(a,b)\in L_{ijt}} x_{ijt}^{l(a,b)\,k} + \sum_{j} y_{ijt}^{k} - \sum_{t'\in\{T|t'<t\}}\left(\sum_{j\in J:\tau_{ji}=t'\text{-}t}\ \sum_{l(a,b)\in L_{jit'}} x_{jit'}^{l(a,b)\,k} + \sum_{j\in J:\tau_{ji}=t'\text{-}t} y_{jit}^{k}\right) = R_{it}^{k}$$

where:

τ_{ij} = transit time from location i to location j, and

R^k_{it} is the quantity of vehicles of class k, existent on location i, available for act in time t, due to decisions that were taken before the beginning of the planning horizon. In this sense, the value of R^k_{it} is zero for $t > \tau_{max}$ (the greatest transit time between two locations).

C - Constraints related to third party vehicles availability

For a class of vehicles k, a given location i and a time period t:

$$\sum_{j} \sum_{l(a,b) \in L_{ijt}} z_{ijt}^{l(a,b)\,k} \leq S^k_{it}$$

D - Integer and non negativity constraints for all decision variables

3.2 Adaptive Dynamic Programming Strategy of Solution

For this class of problems, the most natural solution strategy is to use a rolling horizon procedure, solving the problem at time *t* using what is known at time *t* and a forecast of future events over some time horizon (Godfrey and Powell, 2002a, 2002b).

In case of using Linear Programming, for example, the problem defined in the previous section must be solved repeatedly for each new time period. Two issues arise: the size of the problem grows exponentially with the size of the planning horizon, and the rolling horizon procedure does not take into account the dynamic of the demand evolution over time, rendering it inappropriate for stochastic real world problems.

An adaptive dynamic programming strategy, in turn, reduces the size of the problem (solving it for one period at each time) and also allows to take in account a stochastic demand. For this, it is necessary to define a value function (a function of the value associated to the quantity of available vehicles) for each period *t*, based on an estimate of value for period *t*+1:

$$\widetilde{V}_t(R_t) = \max_{x_t, y_t} \mathrm{g}_t(x_t, y_t) + \hat{V}_{t+1}(R_{t+1})$$

Godfrey and Powell (2002a) consider the objective function “g” as depending on only two decision variables once they do not consider third party vehicles and penalties. In this work a new function, adapted from that described in the previous section as objective function, must be considered.

Godfrey and Powell (2002a) use an estimate of Value Function decomposed by location:

$$\underset{t}{\hat{V}}(R_t) = \sum \underset{it}{\hat{V}}(R_{it})$$

where functions $\hat{V}_{it}(R_{it})$ are estimated utilizing the CAVE algorithm (Godfrey and Powell, 2001). For this, at each stage it is necessary to calculate a gradient for one more and one less vehicle at each location. Furthermore, in the present work this calculation must be performed for each class of vehicle.

The problem is solved performing a “forward step” in time, establishing a set of decisions for variables x and y for a particular demand configuration, and using a particular set of value function approximations. At the end of the "forward step", the dual variables obtained for each subproblem are used to update the function approximations using the CAVE algorithm (Godfrey and Powell, 2001).

The solution is found after successive and adaptive approximations of the decision variables ("forward steps"), followed by the respective updates of the value functions. That is, known the value function estimates and an estimate for future demands, the following sequence of network subproblems is solved, starting with t = 0 and continuing until t = T – 1 :

$$\max_{x_t, y_t} g_t(x_t, y_t) + \sum_{\tau+1}^{\tau_{max}} \underset{t+1,t+\tau}{\hat{V}}\left(R_{t+1,t+\tau}\right)$$

3.2.1 CAVE Algorithm

Solution for the above subproblem may be obtained by linear programming applied to a two stage network formulation (Fig. 2) in which first stage nodes correspond to the objective function on period t and second stage nodes are movements of own fleet vehicles to a super sink that add the value of future stages to the actual subproblem.

Constraints regarding function “g” must be updated from one time period to the next ones, in order to capture new availability of resources. For the present research, otherwise, demand figures must also be updated according to time windows rules.

In regard to the movements to the super sink, each arc has a cost calculated as function of the corresponding future stage value and an upper bound equal to 1. That is, the vehicles are rewarded at the end of the first stage movements according to its future value at the new location.

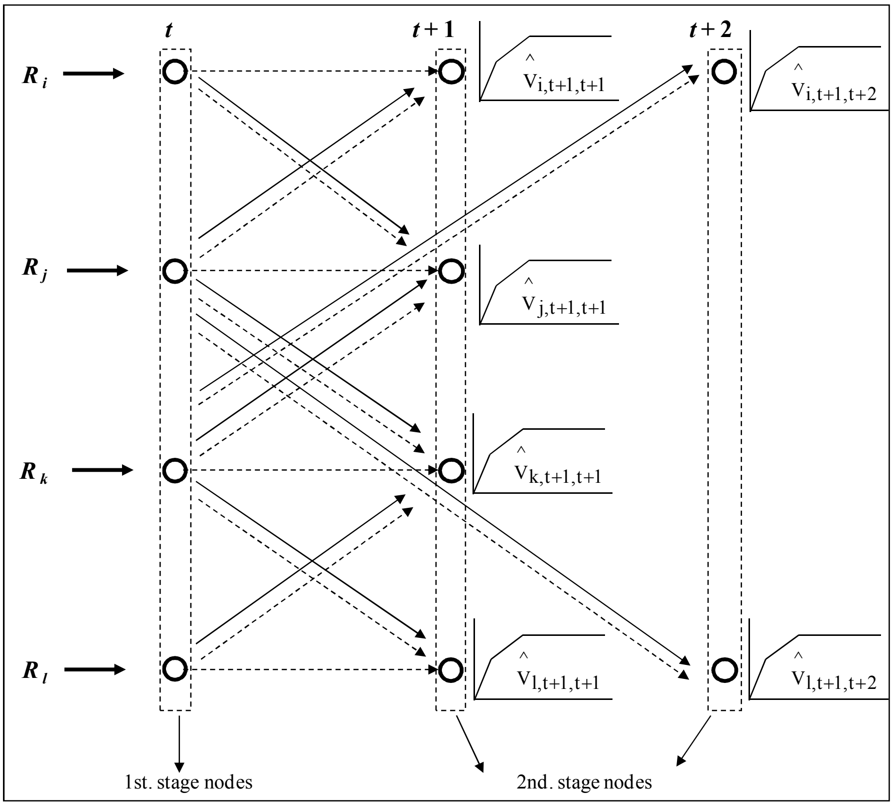

Fig. 2 - 2-stage network for time *t* subproblem

The value of the vehicles at each location is based on duals of the subproblem at that location and time period, defined by Godfrey and Powell (2002b) as:

$\pi^{+}_{i,t,t+\tau}$ = marginal value of one more resource at location i, at time $t+\tau$ in the time t sub problem (τ = 0,1,2,...,τ_{max})

$\pi^{-}_{i,t,t+\tau}$ = marginal value of one fewer resource at location i, at time $t+\tau$ in the time t sub problem (τ = 0,1,2,...,τ_{max})

New values from one iteration to the next are calculated through a smoothing process using a factor α, leading to $V_{new} = \alpha\,\pi + (1 - \alpha)\,V_{old}$, where the utilized π value is π^- for the actual quantity of vehicles and π^+ for one more vehicle at the considered location.

3.2.2 Multiperiod problems

Dual values obtained as proposed in previous section are difficult to implement due to the sequence in time of the algorithm, once the vehicles flowing to one far location in an early time period may imply a higher cost than the allocation of other vehicles from nearer locations in a later but still suitable period.

To overcome this difficulty, Godfrey and Powell (2002b) propose the utilization of the "next" duals, that is, the duals calculated for time period $t + 1$. In this manner:

$$\pi^{+}_{i,t-\tau,t} = \pi^{+}_{i,t-\tau+1,t}$$

and

$$\pi^{-}_{i,t-\tau,t} = \pi^{-}_{i,t-\tau+1,t}$$

for $\tau = 1,2,...,\tau_{max}$

This technique allows capturing, in the time t sub problem, the effect of one more decision period on future problems, mitigating the above difficulty.

This generic scheme was extended in the present research to consider actual conditions, that is, to incorporate new forms of resources (third carriers and heterogeneous vehicles), as well as conditions of compulsory tasks, time windows and cost for not attended loads. Item 4, below, details the proposed model and experimental results.

4. TYPICAL PROBLEM SOLUTION

In order to ensure the applicability of the model and the effectiveness of its results, an experiment was conducted using a problem dimensioned in such a way that it was also possible to solve it with an overall optimization model. This problem consists of 3 locations and planning horizon of 7 time periods. A total of up to 34 own vehicles were considered on the performed experiments.

4.1 Problem Characteristics

Aiming to simplify the model, but without prejudicing the evaluation of its effectiveness, the availability of loads were taken as deterministic ones. It is possible to say that the extension of the experiment to problems with probabilistic demand is guaranteed, since, for this, it is sufficient to alter the characteristics of the network and the way to allocate quantities of demand for the next iteration.

Table 1 shows, for time period 1, locations and demand characteristics, as well as the correspondent rewards and penalties. Some loads have two periods time windows and other ones must be attended in the period that they arrive.

Vehicles configuration is shown in Table 2. Own fleet vehicles appear up to period 2 as consequence of decisions taken before the planning horizon. Third party vehicles availability is considered for each location and period.

Table 1 - Demand on time period 1:

Location		Load Id.	Time Windows		Vehicles	Demand		Own fleet reward		Third party reward		Penalty
From	To		a	b	$\mathcal{K}_l$	$D^{l(a,b)}$	$L_1^{l(a,b)}$	r^{lk}		s^{lk}		p^l
								k = S	k = T	k = S	k = T	
A	B	AB1	1	1	S,T(*)	7	7	22	16	7	1	20
A	C	AC1	0	1	S,T	4	(**)	42	34	12	5	20
			1	2		0	0					
B	A	BA1	0	1	S	2	(**)	10		1		20
			1	2		0	0					
B	C	BC1	1	1	S,T	3	3	32	23	5	-1	20
B	C	BC2	0	1	T	0	(**)		21		-4	20
			1	2		0	0					
C	A	CA1	1	1	S	1	1	34		6		20
C	A	CA2	1	1	T	1	1		26		-1	20
C	B	CB1	1	1	S,T	1	1	15	27	-6	3	20

(*) - S = One axle tractor vehicles; T = Two axle tractor vehicles

(**) - Total Demand (-) not attended demand in previous periods

Two classes of vehicles were considered in the research: one axle tractor vehicles and two axle tractor vehicles. As may be seen in Table 1, some loads need one or another type of vehicle and other loads accept to be attended by both classes.

4.2 Problem Solution

A Decision Support System in Excel and Visual Basic was developed and the free software "lp_solve" (Buttrey, 2005) was utilized to solve the inherent integer linear problems.

Data extracted from a typical real world situation where organized in an Excel worksheet (see Table 1 and Table 2, above) and treated by the proposed solution approach as well as by the global integer problem formulation. Results from both applications were compared.

Figure 3 shows the computer programming process related to the proposed approach.

Ten configurations of value for the same basic structure of the network were simulated, varying the upper limit of some arcs (quantity of loads), the value of the recompense for attending loads (the arc costs) and the distribution of vehicles among the 3 locations in the initial period.

Table 2 - Vehicles availability

Time period	Location	Own fleet vehicles R_{it}^{k}		Third party vehicles S_{it}^{k}	
		k = S	k = T	k = S	k = T
0	A	2	2	4	1
	B	1	1	0	5
	C	4	2	3	2
1	A	2	2	3	0
	B	3	1	0	3
	C	6	1	5	0
2	A	1	0	7	0
	B	1	3	2	4
	C	0	2	0	1
3	A	0	0	10	0
	B	0	0	0	5
	C	0	0	3	0
4	A	0	0	6	0
	B	0	0	0	5
	C	0	0	1	0
5	A	0	0	8	0
	B	0	0	0	4
	C	0	0	1	2
6	A	0	0	4	0
	B	0	0	0	3
	C	0	0	1	2

The problem was decomposed into 7 subproblems, one for each period. Each subproblem was configured with 193 variables and 176 constraints. Iterations were performed until perceiving no improvement in the solution for the first stage. Then, at least 10 iterations were conducted to obtain the overall result (first and second stages), looking for an overall solution improvement of at least 1% as compared to the previous solution. The quantity of iterations varied, in accordance with the adopted configuration.

The adopted smoothing factor varied, among the various experiments, from 0.5 to 0.05. So the value function has been updated, from one iteration to the next, using 50 to 95% of the value accumulated on the previous iteration, and, respectively, 50 to 5% of the value for the new iteration. This factor must be calibrated for each particular case, taking into account that higher values will lead to a quicker solution, but of lower quality.

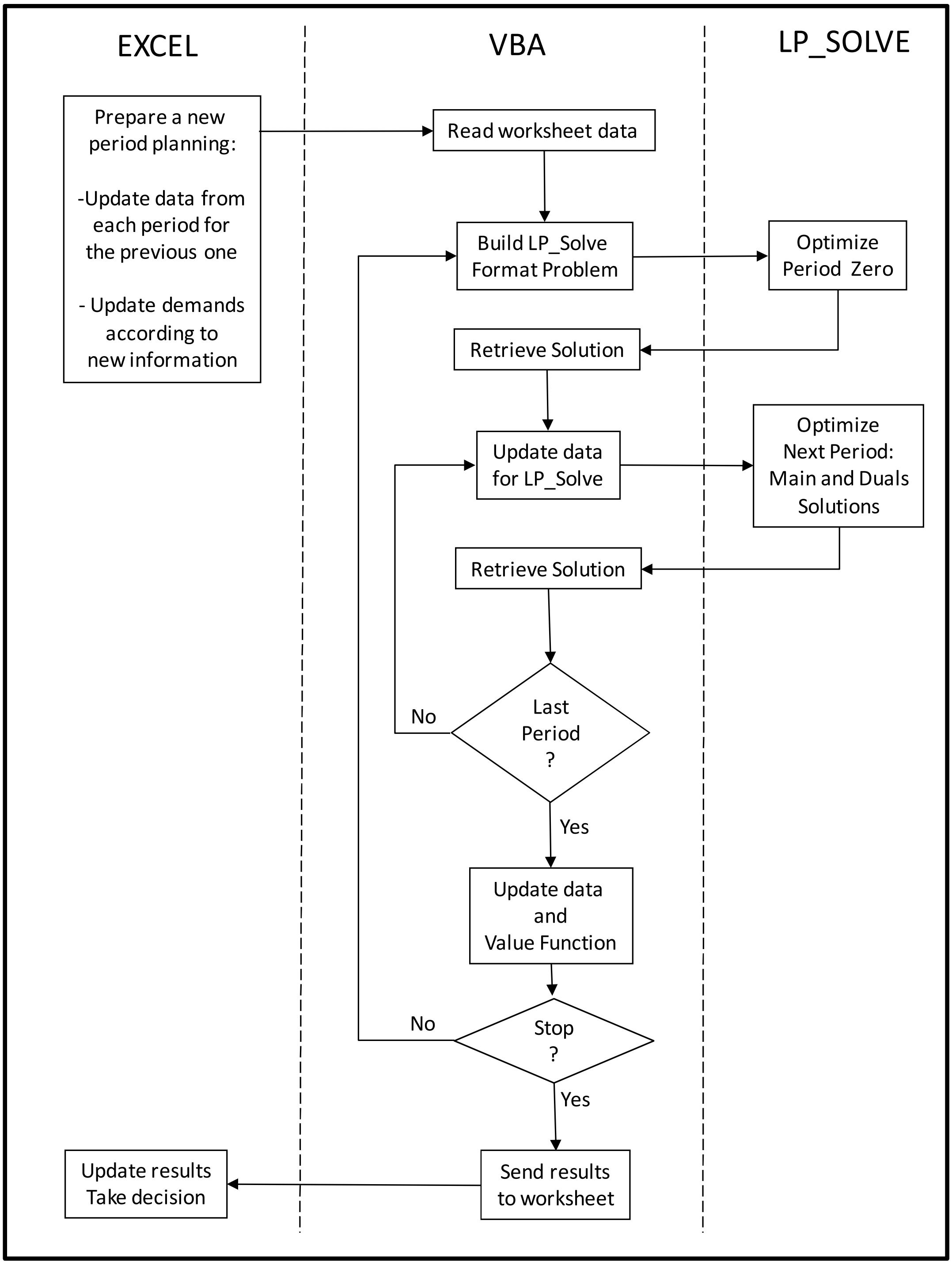

Fig. 3 - Computer Programming Process

Table 3 shows a summary of the experiments. The total reward values, in all experiments, were identical to the ones supplied by the process of global optimization of the network, showing the effectiveness of the subproblem modelling approach.

Achieved results, in terms of rewards, reproduces actual situations. For instance, in experiment 4, existence of 1 more own vehicle provides a gain of $ 42,90 when compared to experiment 1. The same occurs for experiment 10, but with a comparative smaller gain, because location C has a low demand, imposing a costly repositioning move.

Table 3. Experiments Summary

Experiment #	Changed value			Reward $	Reward change $
		From	To		
1	Original data			1,834.60	-
2	Penalty value for all loads	20	10	1,960.60	126.00
3	Penalty value for all loads	20	30	1,736.40	-98.20
4	Quantity of own fleet vehicles at A, time 0	2	3	1,877.50	42.90
5	Quantity of own fleet vehicles at A, time 0	2	1	1,776.70	-57.90
6	Quantity of third party vehicles at A, time 0	4	5	1,862.50	27.90
7	Quantity of third party vehicles at A, time 0	4	3	1,806.70	-27.90
8	Reward for load AC1 with T vehicle, time 1	34	41	1,855.60	21.00
9	Reward for load AC1 with T vehicle, time 1	34	27	1,820.60	-14.00
10	Quantity of own fleet vehicles at C, time 0	1	5	1,858.60	24.00

Although the tested problem is relatively small in size, the computational processing time was of about 1 second, which allows to consider its applicability to cases of substantially higher dimensions. It is possible to suggest that the solution in stages, typical of Dynamic Programming, does not increase the size of the problem to be solved at each stage if the time horizon is extended. There is, quite simply, a greater quantity of subproblems to compute. Thus, if the solved problem had a horizon of 30 periods, the solution of the global optimization problem would spend a computational time greater than the solution for 7 periods, but the solution time for each subproblem would remain the same.

5. CONCLUSIONS

An Adaptive Dynamic Programming Method was presented to solve the problem of resource allocation in long distance road transportation. A network modeling to manage demands not attended was introduced, considering the payment of penalties and/or the utilization of third party transportation firms, heterogeneous vehicles and time windows. The method consists of solving subproblems corresponding to each period of time, using an estimate of value for the optimization of the network throughout the planning horizon.

Tests show that the low processing time of the subproblems does not depend on the magnitude of the planning horizon. The results also indicate that the method can be applied to real world problems, replacing advantageously empirical strategies and optimization techniques applied only to the next planning period.

REFERENCES

BUTTREY, S. E. (2005) Calling the lp_solve Linear Program Software from R, S-PLUS and Excel. *Journal of Statistical Software , v. 14 (4).*

CRAINIC, T. G. (2003) *Long-haul freight transportation***.** In: Handbook of transportation science, Randolph W. Hall (ed.). 2nd. edition. Boston: Kluwer Academic Publishers.

CRAINIC, T., ROY J. (1992) Design of regular intercity driver routes for the LTL motor carrier industry. *Transportation Science*, Baltimore, v. 26, pp. 280-295.

FRANTZESKAKIS, L. F., POWELL, W. B. (1990) A successive linear approximation procedure for stochastic, dynamic vehicle allocation problems. *Transportation Science*, Baltimore, v. 24 (1), pp. 40-57.

GODFREY, G. A., POWELL, W. B. (2001) An adaptive, distribution free algorithm for the Newsvendor Problem with censored demands, with applications to inventory and distribution problems. *Management Science*, v. 47 (8).

GODFREY, G. A., POWELL, W. B. (2002a) An adaptive dynamic programming algorithm for dynamic fleet management, I: single period travel times. *Transportation Science*, Baltimore, v. 36 (1), pp. 21-39.

GODFREY, G. A., POWELL, W. B. (2002b) An adaptive dynamic programming algorithm for dynamic fleet management, II: multiperiod travel times. *Transportation Science*, Baltimore, v. 36 (1) pp. 40-54.

HAGHANI, A. (1989) Formulation and solution of a combined train routing and makeup, and empty car distribution model. *Transportation Research*, v. 23B (6), pp. 433-452.

HANE, C., BARNHART, C., JOHNSON, E., MARSTEN, R., NEMHAUSER, G., SIGISMONDI G. (1995) The fleet assignment problem: solving a large-scale integer program. *Mathematical Programming*, Amsterdan, v. 70 , pp. 211-232.

POWELL, W. B. (1986) A stochastic model of the dynamic vehicle allocation problem. *Transportation Science*, v. 20 (2), pp. 117-129.

POWELL, W. B. (2003) *Dynamic models of transportation operations*. In: Handbooks in Operations Research and management science, 11: Supply Chain Management: design, coordination and operation. A.G. de Kok e Stephen C. Graves (eds.), Amsterdam; Boston: Elsevier.

POWELL, W. B. (2009) What You Should Know About Approximate Dynamic Programming. *Naval Research Logistics*, v. 56, pp. 239-249.

POWELL, W. B., CARVALHO, T. A. (1998) Dynamic control of logistics queuing networks for large-scale fleet management. *Transportation Science*, v. 32 (2), pp. 90-109.

POWELL, W. B., CARVALHO, T. A., GODFREY, G. A., SIMÃO. H. P. (1995) Dynamic fleet management as a logistics queuing network. *Annals of Operations Research*, v. 61, pp. 165-168.

POWELL, W. B., JAILLET, P., ODONI, A. (1995) *Stochastic and dynamic networks and routing*. In: Handbook in Operations Research and Management Science, 8: Network Routing. M.O. Ball et al. (eds.), Amsterdam; Boston: Elsevier, pp. 141-295.

POWELL, W. B., SHEFFI, Y. (1989) Design and implementation of an interactive optimization system for network design in the motor carrier industry. *Operations Research,* v. 37 (1), pp. 12-29.

SIMÃO, H. P., DAY, J., GEORGE, A. P., GIFFORD, T., NIENOW, J., POWELL, W. B. (2008). An Approximate Dynamic Programming Algorithm for Large-Scale Fleet Management: A Case Application. *INFORMS Transportation Science, Articles in Advance, pp.1-20*.

TOPALOGLU, H., POWELL,W. B. (2006) Dynamic programming approximations for stochastic time-staged integer multicommodity-flow problems. *INFORMS Journal on Computing,* v. 18 (1), pp. 31-42.