AN AGENT-BASED APPROACH TO IMPROVE URBAN VEHICLE ROUTING OPERATIONS

Antonio G.N. Novaes
Federal University of Santa Catarina, Florianópolis, SC, Brazil
Bernd Scholz-Reiter
BIBA – University of Bremen, Germany
Edson Tadeu Bez
Univali – Itajaí Valley University, São José, SC, Brazil
Paulo Juliano Burin
Federal University of Santa Catarina, Florianópolis, SC, Brazil

RESUMO
Recentemente, um novo enfoque para tratar problemas de transporte e de logística em nível tático e operacional tem sido tratado na literatura. Dentro desse enfoque o sistema é visto como um conjunto de agentes formados por softwares inteligentes, cada um deles responsável por uma ou mais atividades e interagindo de forma autônoma entre eles. Neste artigo, esse conceito é usado para analisar um problema de coleta urbana e dinâmica de produtos, onde parte das tarefas planejadas e alocadas a um veículo pode ser eventualmente transferida para outros veículos sempre que a demanda excessiva de serviço seja prevista pelo sistema computacional de bordo. Com esse procedimento, a ocorrência de tarefas não cumpridas num ciclo diário pode ser consideravelmente reduzida. A metodologia dinâmica e estocástica inserida no modelo é baseada na Análise Sequencial.

ABSTRACT
In recent years, a new approach to treat transportation and logistics tactical and operational problems has emerged. It views the system as a set of intelligent software agents, each responsible for one or more activities and interacting autonomously among them. In the paper this concept is used to analyze an urban pick-up dynamic vehicle problem (*DVRP*) where part of the planned tasks assigned to a truck can eventually be transferred to other vehicles whenever an excessive service load is foreseen by the on-board computer system. With this procedure, the occurrence of unperformed tasks within a daily cycle can be dramatically reduced. The dynamic and stochastic methodological procedure inserted into the model is based on Sequential Analysis theory.

1.INTRODUCTION

The explosive growth in computer, communication, and information technology in recent years, together with dramatic changes in organizations and markets, have opened new forms of operating manufacturing and logistics activities in an integrated and collaborative way (Goel, 2008). To optimize performance supply chain functions must operate in a coordinated manner, but the dynamics of the participants' operations make it difficult in many instances. Truck breakdown, road traffic congestions, labor absences, customer's cancel or postponement of orders, among other drawbacks generate deviations from the basic plan. Thus, the management of these integrated systems must actuate in a dynamic way, revising the plans and schedules whenever it becomes necessary.

In recent years, a new form for managing integrated logistic services at the tactical and operational levels has emerged. The supply chain is viewed as a set of intelligent agents (software), each responsible for one or more activities, and each interacting with other agents in planning and performing their tasks. In this new form of acting, an agent is an autonomous goal oriented software process that operates asynchronously, communicating and coordinating with other participant agents as needed (Fox *et al*, 2000; Davidsson *et al*, 2005; Berger and Bierwirth, 2010). The theory of computational agents originated about twenty years ago when research in distributed artificial intelligence had been initiated. The modern agent concept made the real breakthrough a decade ago when the mainstream of Artificial Intelligence (*AI*) research shifted from direct goal-seeking to rational behavior, and from the single to multiple cognitive entities acting in communities (Monostori *et al*, 2006). These developments also coincided with the evolution of networked-based computing, as well as novel, human-oriented software engineering methodologies (Monostoni *et al*, 2006). All these achievements led to what is considered now the agent paradigm, which can be resumed in (Wooldridge, 2000; Monostori *et al*, 2006):

- An agent is a computational system that is situated in a dynamic environment and is capable of showing autonomous and intelligent behavior;
- An agent may have an environment that includes other agents. The community of interacting agents, as a whole, operates as a multi-agent system.

A multi-agent system is a loose aggregation of agents, each with clearly defined roles, responsibilities, and functionality (Becker *et al*, 2006). The smallest controlling entity in this approach (an agent) is described as anything that is able to "perceive its environment through sensors (hardware and software) and act upon that environment through actuators" (Russel and Norvig, 2003). In this paper, it will be presented and analyzed an agent-based approach to reduce the number of unperformed tasks in an urban freight distribution system. The rationale involved uses statistical Sequential Analysis to infer, during the servicing process, if traffic conditions will not permit the accomplishment of the planned tasks during the working day, thus transferring part of the jobs to other vehicles and leading to a collaborative servicing process among them.

2. THE PROBLEM

Dynamic Vehicle Routing Problems (*DVRP*) are receiving increased attention among researchers in the areas of transportation and logistics (Larsen, 2000; Ribeiro e Lorena, 2005; Larsen *et al*, 2007; Golden *et al*, 2008; Novaes e Burin, 2009). Such problems are usually related to efficiently assigning vehicles to tasks, such as picking-up and delivering cargo, or accomplishing other services in a previously defined order so that tasks are completed within a certain time limit and vehicle capacities are not exceeded (Figliozzi, 2007, 2010). In large cities such as São Paulo, Brazil, freight operators that deliver or pick-up cargo in such congested urban areas, tend to assign larger numbers of visits to their vehicles in order to increase revenue. This often leads to non-performed tasks at the end of the daily cycle-time, impairing the logistics service level and postponing the service to next day, or even later. This happens because, due to the volatile traffic conditions and the great number of random variables along the route, the vehicle cycle-time usually shows great variability. But even assuming that the fleet of vehicles has been well dimensioned, there are situations in which the traffic becomes exceptionally over congested due to severe accidents, unpredictable public transport strikes, abnormal weather conditions, etc. The objective of this paper is to present an agent-based dynamic model of assigning visiting tasks to a pick-up vehicle, in which one seeks to reduce as much as possible the number of unperformed tasks at the end of the working day. To attain this objective, an integrated working scheme is developed where part of the planned tasks assigned to a vehicle may be transferred to auxiliary vehicles whenever an excessive service load is foreseen by the on-board computer system.

A Multi-Agent System (*MAS*) is a system consisting of independent intelligent control units linked to physical or functional entities such as vehicles, orders, etc. (Mes *et al*, 2007). Agents act autonomously by pursuing their own objectives and interact with each other using informational exchange and negotiation mechanisms (Mes *et al*, 2007). In this application the agents are the vehicles which perform the on-route tasks (pick-up, maintenance services, etc.) plus the central depot which has supplementary vehicles that can be eventually assigned to the routes in case the other agents do not reach agreement to exchange tasks. In a more ample scheme, to be investigated further, other agents are to be added into the model, specifically suppliers and customer firms, each one with their own specific goals as, for example, on-time pick-up and deliveries, the lowest possible costs, etc. The dynamic methodological procedure inserted into the *DVRP* presented in this paper is taken from Sequential Analysis (Wald, 1947; Lai 2001). Assuming a standard composite statistical hypothesis to be sequentially tested, the approach is sufficiently robust and can be extended to more complex situations.

3. MEASURING TRAFFIC CONGESTION

Traffic congestion is seen as a condition of traffic delay (i.e., when vehicle flow is slowed below reasonable speeds) because the number of vehicles trying to use a road exceeds the capacity of the network to handle it (Weisbrod *et al*, 2003). In addition to speed reduction, congestion also introduces variability in traffic conditions, which is known as *travel time reliability* (Cambridge Systematics, 2005). The resulting traffic slowdowns and travel time reliability produce negative effects on supply chain activities, including impacts on vehicle traveling costs, air quality and noise, labor efficiency, industrial and commercial productivity, customer service level, etc. The severity and pattern of congestion, as well as the effectiveness of alternative policies and interventions to address it, vary widely from place to place. That can depend on the size and layout of the urban area, its available transportation options, and the nature of its traffic generators (Weisbrod *et al*, 2003). Congestion is usually the result of seven root causes, often interacting with one another (Cambridge Systematics, 2005): (1) physical bottlenecks; (2) traffic incidents; (3) work zones (temporally reserved for construction and repair activities on the roadway); (4) weather conditions (5) traffic control devices (railroad grade crossings, poorly timed light signals, etc.); (6) special events; and (7) fluctuations in normal traffic.

Another important traffic congestion classification is due to Brownfield *et al* (2003). The first type is *recurrent congestion*, which can be anticipated by road users that are acquainted with the route. The other type is *non-recurrent congestion*, which occurs at non-regular times at a site. It is unexpected and unpredictable by the driver. In our analysis, it is assumed that the logistics entity in charge of the urban distribution service is aware of all programmed events, i.e. it is fully prepared to cope with recurrent congestion. Thus, from the seven factors previously listed, causes 1, 3 and 5 are not considered in our application. Conversely, it is assumed that over-congested situations are originated by causes 2, 4, 6 and 7. Although there is no existing, universally accepted, quantitative definition for traffic congestion, its analysis must rely on easy to measure elements if its impacts are to be evaluated and compared across the range of situations considered in the investigation (Brownfield *et al*, 2003). One frequent assumption is to assume that an urban road link is congested if its *average speed* is below a given percentage of the local speed limit. In addition to average speed reduction to travelers, the sources of congestion also produce time variability known as *travel time reliability* (Cambridge Systematics, 2005), which can be defined in terms of how travel times vary within a pre-defined period. In practical terms, it is useful to fit statistical frequency distributions to travel time, to see how much variability exists in critical sites of the road network. Besides calculating the average time necessary to travel over a route segment, it is also useful to estimate the extra time needed by travelers to ensure a high rate of on-time arrival at their destinations. This extra period, called "*buffer time*", is statistically estimated from historical traffic data considering, for example, a 95th percentile travel time (Cambridge Systematics, 2005). Suppose we take a typical route segment with extension *d*, which can be traveled by a commercial vehicle in time t_0 under normal circumstances. Historical data on that route, taken on a day-to-day basis, and covering the same road segment *d*, furnish the 95% percentile travelling time $t_{95\%}$. The buffer time is defined as the difference

$$\tau_B = t_{95\%} - t_0 , \tag{1}$$

and the buffer index

$$\beta = \frac{t_{95\%} - t_0}{t_0} . \tag{2}$$

Exceptionally, unpredictable and heavy traffic congestions, caused by severe accidents, public transport strikes, heavy storms, etc., may occur during certain working days. In these situations the travelling time increases sharply, with an average value $\hat{t} \gg t_{95\%}$. Let t be the average travelling time in a route in a generic working day. Then, one could say that the traffic conditions are normal if $t \le t_{95\%}$, and over-congested if $t \ge \hat{t}$. Moreover, if $t_{95\%} < t < \hat{t}$ one would not decide immediately for any one of the alternatives, waiting for more information to take a decision. Of course, this is a typical statistical hypothesis testing. Nevertheless, an instantaneous travelling time increase is not, in itself, an indication of an over congested situation. In fact, many non-recurrent events have short duration, and their effects dissipate more or less rapidly. Furthermore, some recurrent events have local impact only, and their effects do not extend to other parts of the served region. Over-congested situations that are of interest in our analysis are the ones with broader geographical extension and longer duration, although in most cases they are no longer than 24 hours. Thus, travel time reliability covering an expressive subset of the urban region, seems to be a good judgmental criterion to evaluate it. And, in order to measure travel time reliability it is necessary to sequentially collect and analyze traffic data. With today's onboard telematics and computing devices it is not difficult to collect and analyze real-time information on traveled distance, time and speed with satisfactory accuracy (Goel, 2008). In our study, the statistical inference process to detect an over-congested condition follows a sequential analysis methodology (Wald, 1947), which is described in the next section.

4. DYNAMIC DETECTION OF OVER-CONGESTED TRAFFIC

Day-to-day traffic flow variability in urban networks produces typical traffic patterns, but unexpected events occasionally cause surges in traffic volumes that overwhelm the road system. Such events, of a "hectic" pattern, are generated by severe accidents with lasting traffic interruptions, extensive public transportation strikes, long duration storms, among others. Strong changes in some characteristic properties of a system may occur occasionally in both technological and natural worlds. And due to today's availability of information processing systems, complex monitoring algorithms have been developed and implemented (Basseville and Nikiforov, 1993). The key difficulty in detecting a fault occurrence through the observation of some properties of a system is to separate noise from the relevant factors. In addition, some failures have a catastrophic nature, leading to an abrupt change in the control variables. But some faults occur with gradual modifications along time, and are often represented by additive changes in the corresponding stochastic model. One way of tackling such kind of problems is *Sequential Analysis* (Wald, 1947; Basseville and Nikiforov, 1993; Lai, 2001).

Classical techniques of statistical inference and hypotheses testing adopt a fixed sample size. With this kind of approach one seeks to minimize the error probabilities for a given sample size. The size of the sample is defined beforehand, and following its statistical analysis, one of two possible actions is taken: accept the null hypothesis H_0, or accept the alternative hypothesis H_1. The null hypothesis represents in our analysis the standard or basic situation, whereas the alternative hypothesis indicates the occurrence of an abnormal condition. Another way to solve hypotheses testing problems when the sample size is not fixed a priori but depends upon the data that have already been observed, is *Sequential Analysis*. Now the problem is the following: for given error probabilities, try to minimize the sample size, or equivalently, make the decision with as few observations as possible. Contrary to the fixed sample size approach, a third possible course of action may occur in sequential analysis when the evidence is ambiguous: take more observations until the evidence strongly favors one of the two hypotheses. Thus, sequential analysis follows a dynamic sequence of observations in such a way that the decision to terminate or not the experiment depends, at each stage, on the previous test results.

Although some authors date the rudiments of sequential analysis to the works of Huyghens, Bernoulli, and Laplace, it was effectively born in response to demands for more efficient testing of anti-aircraft gunnery during World War II, culminating with the development of the *Sequential Probability Ratio Test* (*SPRT*) by Wald, in 1943 (Lai, 2001). A typical case of sequential estimation arises when only one unknown parameter μ is required to define the distribution of the random variable x object of our analysis. Let $f(x,\mu)$ denote the probability density function of x, when x is continuous. Conversely, if x is discrete, $f(x,\mu)$ represents its probability. Let $x_1, x_2, \ldots, x_m$ be a set of m sequential and independent observations on x. Because of the independence of the observations, the joint probability density function is

$$f(x_1,\mu)\, f(x_2,\mu) \ldots f(x_m,\mu). \tag{3}$$

4.1 Sequential test of simple hypothesis

Suppose that it is desired to test the simple hypothesis that $\mu = \mu_0$. This hypothesis is the null hypothesis denoted by H_0. The alternative hypothesis that $\mu = \mu_1$ will be denoted by H_1. Thus, we shall deal with the problem of testing the simple hypothesis H_0 against the alternative simple hypothesis H_1 , on the basis of a sample of m independent observations $x_1, x_2, \ldots, x_m$ on x. According to the developments of Neyman and Pearson, errors of two kinds are present when one accepts or rejects hypothesis H_0. We commit an error of the first kind if we reject H_0 when it is true. On the other hand, we commit an error of the second kind if we accept H_0 when H_1 is true. We denote the probability of an error of the first kind by α, and the probability of an error of second kind by β.

To apply the *Sequential Probability Ratio Test* (*SPRT*) developed by Wald (1947) for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1$, two positive constants A and B ($B < A$) are computed

$$A = (1-\beta)/\alpha \qquad \text{and} \qquad B = \beta/(1-\alpha). \tag{4}$$

Suppose one draws *m* samples, leading to the independent observations $x_1,...,x_m$ on the random variable *x*. At this stage of the experiment the *SPRT* (Wald, 1947; Basseville and Nikiforov, 1993; Lai, 2001) is computed

$$\pi_m = \frac{f(x_1,\mu_1)\,f(x_2,\mu_1)\;...\;f(x_m,\mu_1)}{f(x_1,\mu_0)\,f(x_2,\mu_0)\;...\;f(x_m,\mu_0)}. \tag{5}$$

Three situations may occur:

(a) If $B < \pi_m < A$, the experiment continues by taking an additional observation;

(b) If $\pi_m \geq A$, the experiment is terminated with the rejection of H_0;

(c) If $\pi_m \leq B$, the experiment is terminated with the acceptance of H_0.

For purposes of mathematical simplification it is more convenient to compute the logarithm of the ratio π_m. Let

$$z_i = \ln\left(\frac{f(x_i,\mu_1)}{f(x_i,\mu_0)}\right). \tag{6}$$

Define

$$\pi_m^* = \ln(\pi_m) = z_1 + z_2 + ... + z_m . \tag{7}$$

The test is addictive now. The experiment continues if $\ln B < \pi_m^* < \ln A$ by taking an additional observation; the process terminates with the rejection of H_0 if $\pi_m^* \geq \ln A$; and it terminates with the acceptance of H_0 if $\pi_m^* \leq \ln B$.

4.2 Sequential tests of composite hypothesis

In practical cases, composite hypothesis can occur. One way to solve sequential analysis problems with composite hypothesis is the method of a weighting function associated with the generalized likelihood ratio algorithm (Basseville and Nikiforov, 1993). To do this, two weighting probability distributions, with density functions $g(H_0)$ and $g(H_1)$, depending on H_0 and H_1 respectively, are introduced into the model. The *SPRT* is transformed now into a *weighted likelihood ratio test* (Basseville and Nikiforov, 1993). But, in order to do this, it is necessary to fit distributions $g(H_0)$ and $g(H_1)$ to real data, which depends on detailed information not commonly available in developing country environments. In the application considered in this paper, a more tractable composite hypothesis test can be adopted. This composite hypothesis testing is represented by $H_0^{'}: \mu \leq \mu_0$ versus $H_1^{'}: \mu \geq \mu_1$, such that $\mu_1 > \mu_0$. This model is usually sufficient for practical purposes (Lai, 2001). Assuming that the probabilities of the errors of first and second kind also do not exceed α and β, one can use the *SPRT* of the simple hypothesis $H_0: \mu = \mu_0$ versus $H_1: \mu = \mu_1$ with the same error probabilities α and β. However, while this *SPRT* has minimum expected sample size at $\mu = \mu_0$ and at $\mu = \mu_1$, its maximum expected sample size over μ can be considerably larger than the optimal fixed sample size (Lai, 2001). This means that sometimes the sequential test will not be sufficient to detect hypothesis H_1 during the daily tour, generating unperformed tasks at the end of the working day. But the rate of unperformed tasks will be drastically reduced when compared with the static alternative, as it will be shown in Section 6, a fact that justifies its adoption in our model.

In this application, the variable that commands the decision whether to seek help from another agent or to proceed along the planned routing process is the vehicle displacement time. This is because the route within the district is fixed beforehand, and the stopping times to serve the clients are supposed not to depend on traffic conditions. Thus, the average displacement time reflects reasonably well the traffic conditions in our application, as it was analyzed in Section 3. But, since link lengths vary along the route, and consequently the displacement time, speed is a more appropriate variable to measure traffic conditions. The renewal epochs of the sequential decision process are the instants when the vehicles are ready to depart from one client location to the next stop. At such an instant, the onboard computer evaluates the displacement time *t* over the traveled segment linking the last visit to the present one. The corresponding speed *s* is simply obtained by dividing the travelled segment extension by the respective time, both elements available on the onboard computer. For the analyzed district there are historical values for the travelling times and for the travelling time $\hat{t}$ related to the over-congested condition. This information, together with the series of data collected up to that point, will serve as the basis for inferring whether the traffic condition is normal or over-congested, thus leading to the appropriate operational decision. As discussed in Section 4.1, it is necessary to define a probabilistic distribution $f(x,\mu)$ to represent the variable that commands the decision process. Suppose it was gathered a sample of travelling times in the route inside the district under analysis, which is 15.8 km long. The resulting travelling time frequencies are exhibited in Figure 1. The mean travel time is 33 minutes, corresponding to an average speed of 28.7 km/hr. In the sequel, the cumulative speed frequencies are determined, generating the graph of Figure 2. Two different Erlang distributions were fitted to the data, one with $\theta = 2$ and the other with $\theta = 3$. Here the variable *x* is represented by the speed displaced by 10, i.e., $x = (s-10)$ km/hr. It is seen in Figure 2 that the Erlang distribution with $\theta = 3$ fits better to the data mainly in the lower extreme, a region of more importance for our analysis where the critical speeds occur. The Erlang distribution of a continuous variable *x*, of order θ, has the following probability density function

$$f(x,\mu) = \frac{\mu^{\theta}}{(\theta-1)!}\, x^{\theta-1}\, e^{-\mu x}, \text{ with } \theta = 1,2,...,\infty,\ \mu > 0,\ x \ge 0, \qquad (8)$$

with $\mu = \theta / E[x]$ and $\mathrm{var}[x] = \theta/\mu^2$. In this application, hypothesis H_0 corresponds to $E[x_0] = 28.7 - 10 = 18.7$ km/hr, with $\mu_0 = \theta / E[x_0] = 3/18.7 = 0.16043$. Over-congested situations, on the other hand, occur when $s \le 15$ km/hr, leading to $E[x_1] = 15 - 10 = 5$ km/hr and $\mu_1 = 3/5 = 0.6$.

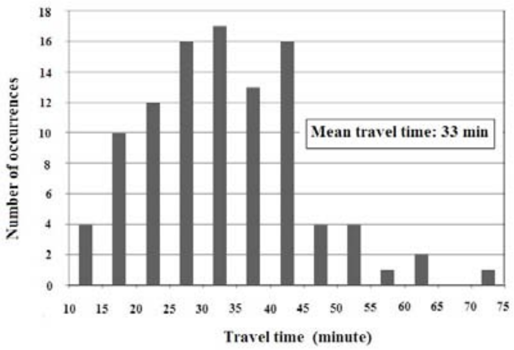

Figure 1 - Travel time distribution on a 15.8 km long route segment

Now, substituting (8) into (6) and (7), and simplifying, one has

$$\pi_m^* = \theta\ m\ \ln\left(\frac{\mu_1}{\mu_0}\right) - (\mu_1 - \mu_0)\sum_{i=1}^{m} s_i\ , \tag{9}$$

where m is the number of sequential tests, s_i is the speed observed in the last route segment, and π_m^* is the control variable of the *SPRT*.

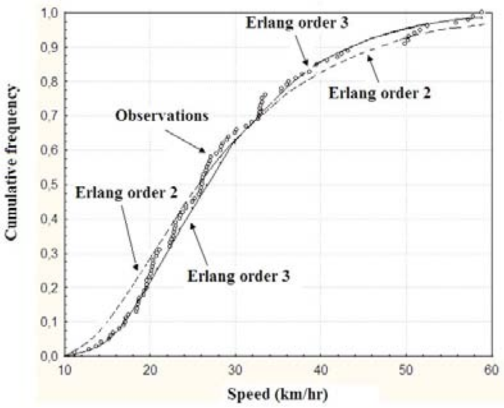

Figure 2 – Erlang distribution fitting to the travelling speed

5. STATIC SOLUTION

Let us take a district (Figure 3), containing 25 suppliers. The district is located 9.5 km from the central depot. The logistics operator has been contracted to collect components from those suppliers, take them to the central depot, transfer the cargo to long-haul trucks, and deliver the components to client's factories located in another towns. On-time delivery is an important logistics attribute, meaning that unperformed pick-ups along the route causes large safety inventory costs and manufacturing delays. Therefore, the number of undone tasks during a daily cycle is an important control variable. It is assumed that the limiting factor in the routing process is time. The time to go from the depot to the district and vice-versa, the displacement time between two successive suppliers in the route, as well as the servicing time (stop time at the supplier premises) vary according to log-normal distributions (Larsen, 2000). The line-haul displacement time has a coefficient of variation $CV_H = 0.2$, with $CV_z = 0.35$ for the displacement time within the district, and $CV_t = 0.45$ for the picking-up time at the suppliers. The average pick-up time at one supplier is 11 minutes. Taking the 25 points shown in Figure 3, it was applied a 3-opt algorithm (Syslo *et al*, 2006) to solve the corresponding Travelling Salesman Problem (*TSP*), resulting in the route sequence depicted in that figure.

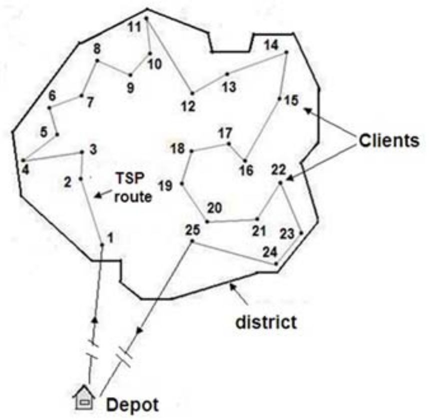

Figure 3 - **A district containing 25 suppliers**

A simulation was performed with 10,000 replications and considering the time sequence along the route, comprising: (a) the displacement time from the depot to the first supplier (number 1, Figure 1); (b) the service time at the first stop; (c) the displacement to the next point, etc., up to the last visit or when there is no time available to perform additional visits, also considering the return time to the depot. The maximum daily working time was assumed to be 8 hours. Considering days of normal traffic only (average speed of 28.7 km/hr), the resulting expected cycle time was 6.99 hr, with expected maximum cycle time equal to 7.9 hr (98% confidence), within the working time limit. For this situation it was observed a 0.19% rate of unperformed tasks. No more than 3 returns were observed for one same truck, with 3 returns occurring in only 0.004% of the cases, which is a quite negligible situation.

Next, considering over-congested days, with $s \leq 15$ km/hr, the situation changes significantly. Now, 4.3% of the tasks are not performed during the cycle time, being postponed to the next day or even later. There will be days when up to six pick-up tasks will not be performed. Even if the occurrence of over-congested traffic situations is scarce, the performance disruption is generally not acceptable under present day logistic standards, requiring an improved solution.

6. DYNAMIC SOLUTION

After completing a visiting task in the tour, the vehicle on-board system applies the *SPRT* test through relation (9) in order to infer which hypothesis is binding, or if it is necessary to proceed further with the test. Figure 4 shows a schematic representation of the vehicle routing sequence and the decision stages where the *SPRT* is performed. Assume that the vehicle agent *A* (Figure 4) left the depot with the assignment of 25 visits. Suppose the sequential test indicates the acceptance of hypothesis H_1 at stage 3, as shown in Figure 4. At this point, the on-board computer checks how many of the remaining visits should be transferred to another vehicle agent *B*. Let *k* be the number of visits to be transferred. Upon negotiation, agent *B* agrees to perform the *k* tasks. Of course, if there are two or more visits to be transferred, more than one agent can be involved in the transference.

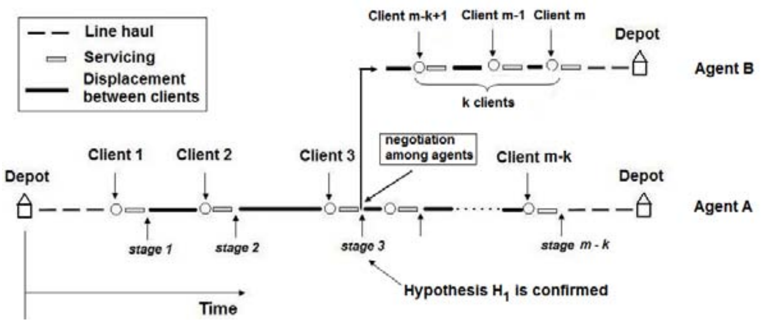

Figure 4 – The vehicle routing sequence and the decision stages

Let us analyze one specific day with over-congested traffic conditions. It is assumed $\alpha = \beta = 0.01$, leading to $\ln A = 4.59512$ and $\ln B = -4.59512$. Table 1 shows the sequential analysis process up to stage 7, when hypothesis H_1 has been detected. It is seen in Table 1 that the controlling coefficient π_m^* is in the interval $\ln B < \pi_m^* < \ln A$ for stages 1 to 6. This indicates that the sequential test must proceed further. At stage 7, the parameter π_m^* surpasses the value of $\ln A$, meaning hypothesis H_1 was detected, and thus leading to the exchange of tasks among agents. In this application, the occurrence of the transition stages varies from stop 6 to stop 18, as shown in Figure 5, with greater concentration around stage 9. When the transition stage occurs much later in the tour sequence, the transference of visits to other agents may not be accepted due to the higher risk of miss accomplishment of tasks. But a good part of the tasks to be transferred are usually situated in less critical points along the tour. The simulation of the agent-based dynamic model, with 10,000 replications and for the critical over-congested situation, resulted in 11,8% tasks transferred to other vehicles agents in average, and only 0.004% unperformed visits. Compared with the percentage of 4.3% unperformed visits that occurred in the corresponding static alternative, this result is quite significant. But the transference of tasks is somewhat large since, for a 25-client route, there would be an average of almost three visits transferred per cycle. One possibility is to improve the inference model, adding information (intelligence) from other sources, and anticipating the decision points.

Table 1 – Sequential analysis test for one specific working day

Stage m (equation 9)	Speed s_i in the last route segment (km/hr)	Cumulative speed $\sum_{i=1}^{m} s_i$	π^*_m (equation 9)	Decision to be taken
1	11.0	11.0	0.34690	Continue testing
2	6.7	17.7	1.29084	Continue testing
3	8.9	26.6	1.93109	Continue testing
4	12.5	39.1	2.06101	Continue testing
5	11.2	50.3	2.36947	Continue testing
6	4.8	55.1	3.57418	Continue testing
7	4.5	59.6	4.82893	H_1 confirmed

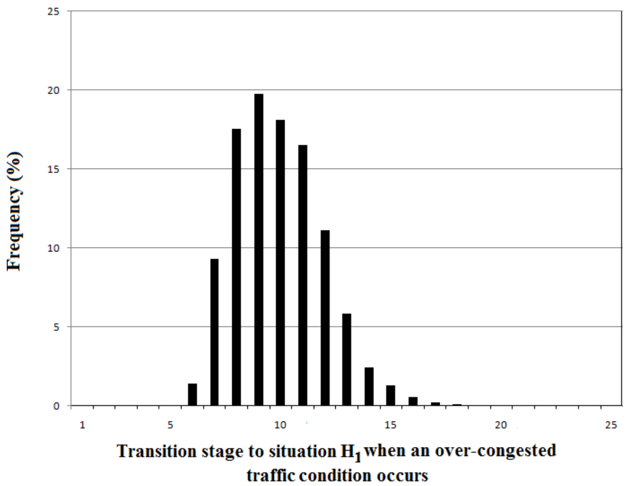

Figure 5. Sequential analysis: the frequency of transition stages

7. TASK TRANSFER AMONG AGENTS

At the present stage of the study, three types of collaborative agents participate in the process (Goel and Gruhn, 2006): (a) *vehicle agents*; (b) *task assignment agent*; (c) *auction manager agent* (Mes *et al*, 2007). Each vehicle constitutes a separate *vehicle agent*, analyzing the performance of its route tasks, verifying whether it is necessary to transfer part of its jobs to other vehicle agents or not, and evaluating the possible acceptance of tasks proposed by other vehicle agents. The *task assignment agent* prepares in advance the sequence of tasks for each vehicle agent, modifies them whenever external requests make it necessary (requests from clients or suppliers), keeps a register of failures, driver's performance, etc., for further analysis and corrections, and distributes the performance results (positive and negative) to the participating agents. Finally, the *auction manager agent* supervises the bidding process of tasks that occurs among vehicle agents. It evaluates all bids and interferes in the process in case the vehicle agents cannot reach a reasonable agreement. For example, it is possible that the vehicles assigned to the routes are not able to perform themselves all the tasks in excess at a certain moment. Then, the *auction manager agent* transfers those tasks, partially or totally, to the *task assignment agent*, which allocates one or more vehicles from the depot to perform the exceeding jobs.

The acceptance of an additional task by a vehicle agent will lead to a new vehicle scheduling route, for which there will be several sequential alternatives. For example, one may insert the new job at various positions in the current route sequence, or one may shuffle the entire schedule to find a new optimum route (Mes *et al*, 2007). Also, the vehicle agent that is offering tasks to be transferred to other vehicles can choose the jobs to be submitted to the auction, and will analyze the ones that suit him best to transfer. Of course, the *auction manager agent*, which has access to the information originating from all the vehicle agents, may interfere in the process if the individual vehicle agent propositions result in schemes too far from the overall optimum. Another important consideration is that the transfer instants along the route vary substantially (see Figure 5), and therefore it might happen that the other vehicle agents will not bid for the offered tasks when the auction occurs late during the cycle time. Another important point is the definition of fixed periods between auctioning of tasks. At the beginning of the daily cycle, there will not be enough information to justify an auction. So, depending on the specific characteristics of the problem, the starting point of the auctions and their periodic intervals must be defined, based on a simulation analysis.

8. CONCLUSIONS

The core of an agent-based decision structure intended to define a transferring task process among vehicle agents has been described and analyzed. Its objective is to effectively reduce non-performed tasks along a daily vehicle routing cycle, thus improving the logistic service level in an urban picking-up process subject to occasional over-congested traffic conditions. In a subsequent research, the authors will analyze in detail the routing process of the transferred tasks, plus the auctioning rationale and the reward/penalty modeling, the latter intended to give the economic and operational support to the agent's bidding process.

Acknowledgments
This research has been supported by the Brazilian Capes Foundation and by DFG - Deutsche Forschungs-gemeinschaft (German Research Foundation), Bragecrim Project nº 2009-2.

REFERENCES

Basseville, M. and Nikiforov, I.V. (1993). *Detection of Abrupt Changes: Theory and Application*, Prentice-Hall, New Jersey, USA.

Becker, M.; Wenning, B.-L.; Görg, C.; Allee, O.-H.; Gehrke, J.D.; Lorenz, M. and Herzog, O. (2006). Agent-based and Discrete Event Simulation of Autonomous Logistic Processes, In: Borutzky, W. Orsoni, A.; Zobel, R. (eds.) *20th European Conference on Modelling and Simulation 2006*, p. 566-571.

Berger, S. and Bierwirth, C. (2010). Solutions to the Request Reassignment Problem in Collaborative Carrier Networks, *Transportation Research Part E*, v. 46, p. 627-638.

Brownfield, J.; Graham, A.; Eveleigh, H.; Maunsell, F.; Ward, H.; Robertson, S.; Allsop, R. (2003). Congestion and Accident Risk, *Road Safety Research Report Nº 44*, Department of Transport, UK.

Cambridge Systematics (2005). *Traffic congestion and reliability*, Federal Highway Administration, Washington, DC.

Davidsson, P.; Henesey, L.; Ramstedt, L.; Törnquist, J. and Wernstedt, F. (2005). An Analysis of Agent-Approaches to Transport Logistics, *Transportation Research Part C*, V. 13, p. 255-271.

Figliozzi, M.A. (2007). Analysis of the efficiency of urban commercial vehicle tours: Data collection, methodology, and policy implications, *Transportation Research Part B*, **41**, p. 1014-1032.

Figliozzi, M.A. (2010). The impacts of congestion on commercial vehicle tour characteristics and costs, *Transportation Research Part E*, **46**, p. 496-506.

Fox, M.S.; Barbuceanu, M. and Teigenm R. (2000). Agent-Oriented Supply Chain Management, *The International Journal of Flexible Manufacturing Systems*, V. 12, p. 165-188.

Goel, A. and Gruhn, V. (2006). Collaborative Dispatching of Commercial Vehicles, *IEEE 1-4244-0023-6/06*.

Goel, A. (2008) *Fleet Telematics*, Springer, New York.

Golden, B.; Raghavan, S. and Wasil, E. (Eds.) (2008) *The Vehicle Routing Problem: Latest Advances and New Challenges*, Springer, New York.

Lai, T.L. (2001). Sequential analysis: Some classical problems and new challenges, *Statistica Sinica*, V. 11, p. 303-408.

Larsen, A. (2000). *The Dynamic Routing Problem*, Ph.D. Dissertation, Technical University of Denmark, Lyngby, Denmark.

Larsen, A., Madsen, O., Salomon, M. (2007) Classification of Dynamic Vehicle Routing Systems. In: Zeimpekis, V., Tarantilis, C.D., Giaglis, G., Minis, I. (eds.), *Dynamic Fleet Management*. Springer, New York.

Mes, M.; van der Heijden, M.; van Harten, A. (2007). Comparison of Agent-Based Scheduling to Look-ahead Heuristics for Real-Time Transportation Problems, European Journal of Operational Research, V. 181, pp. 59-75.

Monostori, L.; Váncza, J. and Kumara, S.R.T. (2006). Agent-Based Systems for Manufacturing, *Annals of the CIRP*, v 55/2/2006, p. 697-720.

Novaes, A.G. and Burin, P.J. (2009). A dynamic vehicle routing problem (in Portuguese), *XXIII ANPET*, Vitória, ES, Brazil, November 2009.

Novaes, A.G.; Frazzon, E.M. and Burin, P.J. (2011). Dynamic Vehicle Routing in Over Congested Urban Areas, In: Kreowski, H.-J., Scholz-Reiter, B. and Thoben, K.-D. (eds.), *Dynamics in Logistics*, p. 49-58, Springer: Berlin Heidelberg.

Ribeiro, G. M. and Lorena, L. (2005) Dynamic Vehicle Routing Using Genetic Algorithms (in Portuguese) *Proceedings, XVI Conference on Transport Research and Education - ANPET*, Recife, PE.

Russel, S. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach*, Prentice Hall, New Jersey, NJ, USA.

Syslo, M.M.; Deo, N. and Kowalik, J.S. (2006). *Discrete Optimization Algorithms with Pascal Programs*, Dover Books on Mathematics: Mineola, NY.

Wald, A. (1947). *Sequential Analysis*, Dover Publications: New York.

Weisbrod, G.; Vary, D. and Treyz, G. (2003). Measuring the economic costs of urban traffic congestion to business, *Transportation Research Record nº 1839*, Transportation Research Board, Washington, DC.

Wooldridge, M. (2000). *Reasoning about Rational Agents*, MIT – Massachusetts Institute of Technology, Cambridge, MA.

Corresponding author:
Prof. Antonio G.N. Novaes
Post-Graduate Program in Production Engineering
Federal University of Santa Catarina
Florianópolis, SC, Brazil
88040-900
Tel: (55) 48 37217081, E-mail: novaes@deps.ufsc.br