

# School Bus Routing and Scheduling: a Real Case Study

Pedro Oliveira and Ana Paula Barbosa-Póvoa\*  
Centro de Estudos de Gestão, CEG-IST  
Instituto Superior Técnico, Universidade Técnica de Lisboa  
Av Rovisco Pais  
1049-001 Lisboa, Portugal

## Abstract

This paper presents a model to schedule school buses transportation routes. Student's daily routes defined between student's homes, schools and extra curricula activities are analysed while accounting for three different objectives: number of vehicles, route total time and average passenger travelling time. A combined multiobjective function is defined. As solution method to this problem, a heuristic algorithm was developed. This combines a set of heuristic procedures with a simulated annealing algorithm. The developed algorithm applicability and efficiency is studied on the solution of a real case-study of a students transportation company that operates in the city of Lisbon in Portugal. As results, it was found an improvement on the economics and service level up to 3% and 28%, respectively, when compared to the real current routes.

**Keywords:** routing problem; school bus; simulated annealing; multiobjective optimization; real case study

---

\* Author to who correspondence should be addressed; apovoa@ist.utl.pt

## Introduction

The well known Vehicle Routing Problem (VRP), initially studied in Dantzig and Ramser (1959) and widely discussed in Laporte (1992), is the source of the School Bus problem. The school Bus problem can then be viewed as one of the VRP special applications, where the objective involves the definition of student's transportation routes. A route is defined from student's homes to schools and extra-curricula activities.

In the literature some works exist on the treatment of this problem that differ on the type of problem definition, on the objectives treated, on the model details, and on the algorithms used to find a solution.

As to the problem type, two main types can be identified. A first one involves the study of a single school scenario (Bennett and Gazis, 1972) while the second one treats a multi school planning scenario (Newton and Thomas, 1974). On the latter two different approaches can be distinguished based respectively on the school or on the students home. In the first one, a problem for each school is solved and the creation of mixed routes is not possible. Therefore students that attend different schools can not share the same vehicle (Bodin and Berman, 1979). On the other hand, the approach based on student's home, considers the resolution of the problem student by student and therefore the sharing of resources (buses) is allowed (Braca *et al*, 1994).

Regarding the objectives, several examples are mentioned on the literature. These go from economics factors to the level of service to customers. Braca *et al*. (1994) considers as only objective the minimization of number of vehicles while Bowerman *et al*. (1995) besides the minimization of vehicles also considers the minimization of the walking distances between the pick up points and the student's homes. The travelling time per student inside the bus, the total journey time as well as the balance between the number of passengers and the journey time are also considered. The last objectives are also studied in Lyo and Fu (2002) where they are formulated within a multiobjective formulation.

As to model details different restrictions have been considered by the published works. Newton and Thomas (1974) considered an upper limit on the travel per student. Braca *et al*, (1994) looked into limits on the capacities of the vehicle both as minimum and maximum values as well as to the existence of time windows. Bowerman *et al*, (1995) introduced the notion of maximum route duration.

Due to the complexity of the problem in study, several algorithms have been developed to solve it. For the problems modelled according to the school based approach, it can be

highlighted the work of Bennett and Gazis (1972) that used the methodology proposed by Clarke and Wright (1963). The authors attribute stopping points to each route and then re-link them in order to fulfil the problem restrictions. The Lin-3-opt (Lin, 1965) exchange procedure is then applied to improve the initial solution. Bodin and Berman (1979) used the method proposed by Newton and Thomas (1974), starting the algorithm with a TSP solution with the Lin-3-opt and then breaking the solution into several routes so as to satisfy the problem restrictions.

On the problems based on the student's home, Braca *et al.* (1994) solved it in a single step. Firstly, the authors build up routes randomly selected between homes and schools. After that pairs of homes and schools that minimize the total journey distance are inserted assuring that problems restrictions were satisfied.

Metaheuristics approaches such as scatter search, simulated annealing or tabu search are also widely used to solve School Bus problems. Examples of applications using such techniques can be listed in the works of Corberán *et al.* (2002) and Spada *et al.* (2003).

Following the above works this paper looks into the school bus scheduling problem and considers firstly the need to allocate customers to the transportation vehicles and then the definition of the sequence of nodes to be visited by each vehicle. A generic model is developed where several types of restrictions are considered. These involve limits on the travelling time per passenger; the transport capacity of each vehicle; the possibility of having pick ups and deliveries throughout the route and finally, the existence of time windows associated with pick up and delivery moments. As main objectives the service level and economic issues are both considered. Specifically the total travelling time spent on each journey per student, the number of vehicles and the total routing time of the entire fleet are modelled. Besides these objectives it is also considered and identified as secondary indicators, some performances measures related exclusively with the service level. To solve the proposed model a solution method is developed. This considers a heuristic algorithm that combines a set of heuristics, designed exclusively for the problem in study, coupled with a generic metaheuristics, the simulated annealing algorithm. As final result a set of "optimal" routes for each objective are obtained. This gives the possibility to the planner to choose which routes fit better his(her) preferences. The model and algorithm developed are applied to a real case-study of a students transport company located in Lisbon Portugal.

The paper is structured as follows. In the next section the case-study is characterised. Then in third section the mathematical formulation is presented. This is followed by the algorithm description and the analysis of the results obtained. The paper concludes with some generic conclusions and the identification of future lines of research.

## **Case Study**

The case-study studied in this paper is obtained from a real problem observed at a Portuguese company whose main activity is the daily transport of students to schools, extra-curricular activities and their return home. This activity and its proper characteristics are subject to a number of restrictions so as to guarantee the clients service level and attain certain objectives. These are described bellow.

### **Daily Operation**

The company operation, that is going to be closely analysed in the paper, can be broke in two time periods:

- Morning period: 7:00 am to 10:00 am;
- Afternoon period: 14:00 pm to 19:00 pm.

In the first period, the usual operation consists in the students pick up from their homes and on their delivery to schools or colleges, while in the second period the operation is mainly characterised by the pick up of students from schools to home or in after school extra activities. Nevertheless, in the latter is still possible to make some home pick ups and school deliveries (especially for the ones who have afternoon school schedules).

The definition and planning of the set of routes of this company, is made by a number of planners whose activity is entirely manually and is based on their experience achieved through the time. It consists on the simulation and effective experimentation of innumerable possible combinations. This results as very time consuming since it implies the full allocation of several resources to a large number of clients. Furthermore, has as main drawback the impossibility of a quick response to a possible client order.

### **Restrictions**

The definition of routes is subject to a number of restrictions that must be considered and fully respected. These involve:

- Transportation duration limits per passenger;
- Finite and differentiated transportation capacity limit per vehicle;
- Time windows associated to each customer (related to pick up and delivery node).

For formulation purposes, it is necessary to take in account the sequential relations between nodes to be visited, in order to assure that for each customer the pick up point is firstly considered and just after the delivery node appears within the route defined.

### Planners Objectives

From the transport company point of view, there are two basic objectives to be achieved: economics and level of service.

From the economics perspective, the target is to satisfy all the demand at a minimal cost. This implies the minimization of the necessary vehicles in operation (1) reducing the investments costs in equipment and wages with the vehicle crew. This can be written as follows

$$NV = \sum_{i=1}^K k_i \quad \text{in which } i \in K \{1, \dots, K\} \quad (1)$$

where  $K$  is the number of available vehicles to the operation and  $NV$  the total number of vehicles.

The evaluation of the variable costs is done by assuming that the main sources of costs are the travelling costs. These costs can be materialized by the value of the fuel necessary to the journeys and depreciation of vehicle on route. Knowing that travelling costs increase with the travelled distances and assuming that there is a constant average speed between the several nodes it is defined as indicator the total number of minutes to perform all operation (2), which should be minimized.

$$NTM = \sum_{k=0}^K \sum_{i=0}^N \sum_{j=0}^N c_{ij} x_{ij}^k \quad (2)$$

where  $C_{ij}$  is the time distance between nodes  $i$  and  $j$ , in minutes and  $NTM$  the total time in operation measured in minutes.

$$x_{ij}^k = \begin{cases} 1, & \text{case arch } (i,j) \text{ belongs to route made by vehicle } k, \\ 0, & \text{otherwise.} \end{cases}$$

The evaluation of the costumer service is measured taking into account the time spent in journey (3) by each passenger. A theoretical upper bound exists which should never be achieved. This is defines as the "impatience maximum limit" given by:

$$TTD = \sum_{j=0}^N \sum_{i=0}^N (T_j \cdot D_{ji}) - \sum_{i=0}^N \sum_{j=0}^N (T_i \cdot P_{ij}) \quad (3)$$

where:

$D_{ji}$  – delivery in node  $j$  having as origin  $i$ , in units;

$P_{ij}$  – pick up in node  $i$  and delivery in  $j$ , in units;

$T_i$  – spent time until node  $i$  (included  $i$ ) in minutes;

$T_j$  – spent time until node  $j$  (included  $j$ ) in minutes;

$V$  – set of customer pick up and delivery nodes;

$N = |V|$ .

Therefore and as metric for this criterion several possibilities may arise, as for example the maximum travelling time among all passengers. However in this work the chosen measure is the average travelling time (4). This choice is based on the fact that exist several transport services for which it will be very difficult to decrease the transport time since they are very far away for their destiny. This metric is then given by:

$$TDM = \frac{2 \cdot TTD}{N} \quad (4)$$

where:

$V$  - set of customer pick up and delivery nodes

$N = |V|$

Beyond the metrics presented above there are others indicators evaluated in the decision process in order to incorporate other perspectives in the planning process. These are the secondary indicators and can be defined as:

- Minutes in late arrivals;
- Minutes in early arrivals;
- Waiting time for vehicle;
- Vehicle immobilization time during course;
- Vehicle capacity utilization rate.

## Purpose of Work

The objective of this work is the development of a model that can be used by the planner as a tool to help his(her) decision making process where two distinguished perspectives are accounted for. These are respectively a strategic long run approach and an operational and short run purpose.

On the strategic perspective the planner should be able to plan in advance the long run supply capacity requirements based on historical data and new client's previsions. Different scenarios can then be defined and evaluated. On the operational point of view the planner should be able to execute the operational routing planning, considering scenarios where the effective demand is known. The process should be considered as wide as possible where the integration of primary and secondary performances indicators within a acceptable time horizon must be accounted for.

## Mathematical Formulation

Based on the problem described above a mathematical formulation is developed. The distance between the several nodes is evaluated using the time distance as the unit of measure. The average travelling time between two nodes, is given by the sum of the expectable course time (6) and the service time (5), which is the necessary time to allow the entry or exit of a customer to/from the vehicle.

Average Travelling Time (between nodes  $i$  and  $j$ )

$$C_{ij} = CS_{ij} + CP_{ij} \text{ (minutes)} \quad (5)$$

where,

Expectable Course Time (between nodes  $i$  and  $j$ )

$$CP_{ij} = \frac{d_{ij} * 60}{SP_{ij}} \text{ (minutes)} \quad (6)$$

$d_{ij}$  – is the geographical distance between nodes  $i$  and  $j$  (evaluated in Km);

$SP_{ij}$  – is the expected average speed of vehicle in normal conditions at a certain hour of day between nodes  $i$  and  $j$  (evaluated in Km/hour).

There are as many pick up and delivery points as the number of customers to be transported. In the cases in which the customer location (origin or destiny) is geographical identical, it is

considered that the course time is equal to zero and that the only distance between these two matching points is the service time.

### Notation

$i, j$  – nodes that identify the origin and destination of a route;  
 $k$  – transportation vehicle;  
 $O_k$  – set of vehicles origin points;  
 $V$  – set of customers pick up and delivery points;  
 $VO$  – set of customers pick up and delivery points plus set of vehicles origin points  $VO = V \cup O_k$ ;  
 $C_{ij}$  – distance between nodes  $i$  and  $j$ , in minutes;  
 $CP_{ij}$  – expectable course time (between nodes  $i$  and  $j$ , in minutes);  
 $CS_{ij}$  – expectable time to pick up or delivery customer in node  $j$  having as origin node  $i$ , in minutes;  
 $D_i$  – delivery in node  $i$  in units;  
 $D_{ij}$  – delivery in node  $i$  having as origin  $j$ , in units;  
 $d_{ij}$  – geographical distance between nodes  $i$  and  $j$ , in kilometres;  
 $K$  – maximum number of vehicle involved in operation, in units;  
 $L$  – maximum limit of transportation time per passenger, in minutes;  
 $N$  – total number of pick up and delivery points  $N = |V|$ ;  
 $P_i$  – pick ups in node  $i$  in units;  
 $P_{ij}$  – pick up in node  $i$  having as destiny  $j$ , in units;  
 $Q_k$  – transportation capacity of vehicle  $k$ , in units;  
 $RT_i$  – time window associated to node  $i$ , in minutes;  
 $SP_{ij}$  – average speed between nodes  $i$  e  $j$ , in Km/Hour.

### Variables

$k_k$  – number of vehicles  $k$  used;  
 $T_i$  – time spent until node  $i$  (included  $i$ );  
 $tp_{ij}^k$  – stopped time of vehicle  $k$  during arch  $(i, j)$ ;

$$x_{ij}^k = \begin{cases} 1, & \text{if arch } (i, j) \text{ belongs to route made by vehicle } k, \\ 0, & \text{otherwise.} \end{cases}$$

$Y_{ij}$  – number of pick ups made until node  $i$  (included  $i$ ) and transported in arch  $(i, j)$ ;  
 $Z_{ij}$  – number of deliveries made until node  $i$  (included  $i$ )



### Objective Function

As previously mentioned three objectives are evaluated: the minimization of the number of vehicles (7), the total course time in route (8) and the average travelling time per passenger (9). These are defined above:

$$\text{Min } NV = \sum_{k=1}^K k_k \quad (7)$$

$$\text{Min } NTM = \sum_{k=0}^K \sum_{i=0}^N \sum_{j=0}^N c_{ij}^k x_{ij}^k \quad (8)$$

$$\text{Min } TDM = \frac{2 * \left[ \sum_{j=0}^N \sum_{i=0}^N (T_j \cdot D_{ji}) - \sum_{i=0}^N \sum_{j=0}^N (T_i \cdot P_{ij}) \right]}{N} \quad (9)$$

### Multiobjective Function (FAM)

The objectives considered are aggregated into a single objective where their importance is weight based on three given parameters:

$$\text{Min}(FAM) = \beta \times NV + \lambda \times NTM + \delta \times TDM \quad (10)$$

Where:

$\beta$  - Reflects the importance of the number of vehicles in FAM;

$\lambda$  - Reflects the importance of total course time in route in FAM;

$\delta$  - Reflects the importance of average travelling time per passenger in FAM.

### Restrictions

Apart from the objective function and having in mind its characteristics different restrictions are defined. These are defined below:

$$\sum_{i=0}^N \sum_{k=1}^K x_{ij}^k = 1 \quad \forall j \quad (11)$$

$$\sum_{i=0}^N x_{ij}^k = \sum_{i=1}^N x_{ji}^k \quad \forall j \text{ and } \forall k \quad (12)$$

$$\sum_{O=1}^K x_{oj}^k = \sum_{k=1}^K k_k \quad \forall j \quad (13)$$

$$\sum_{k=1}^K k_k \leq K \quad (14)$$

$$CP_{ij} = \frac{d_{ij} * 60}{SP_{ij}} \quad (15)$$

$$C_{ij} = CS_{ij} + CP_{ij} \quad (16)$$

$$T_j = T_i + tp_{ij}^k + \sum_{j=1}^N C_{ij} x_{ij}^k \quad \forall j \quad (17)$$

$$T_i + tp_{ij}^k + \sum_{i=0}^N c_{ij} x_{ij}^k \leq RT_j \quad \forall j \quad (18)$$

$$T_j - tp_{ij}^k - \sum_{j=0}^N c_{ij} x_{ij}^k \leq RT_i \quad \forall i \quad (19)$$

$$T_j - tp_{ij}^k - \sum_{j=0}^N c_{ij} x_{ij}^k \geq RT_i \quad \forall i \quad (20)$$

$$\sum_{j=0}^N Y_{ij} x_{ij}^k = \sum_{j=0}^N Y_{ji} x_{ji}^k + P_i \quad \forall i \quad (21)$$

$$P_i = \sum_{j=1}^N P_{ij} \quad \forall i \quad (22)$$

$$\sum_{j=0}^N Z_{ij} x_{ij}^k = \sum_{j=0}^N Z_{ji} x_{ji}^k + D_i \quad \forall i \quad (23)$$

$$D_i = \sum_{j=1}^N D_{ij} \quad \forall i \quad (24)$$

$$Y_{ij} - Z_{ij} \leq \sum_{k=1}^K Q_k x_{ij}^k \quad \forall j \text{ and } \forall i \quad (25)$$

$$(T_j.D_{ji}) \geq (T_i.P_{ij}) \quad \forall j \text{ and } \forall i \quad (26)$$

$$(T_j.D_{ji}) - (T_i.P_{ij}) \leq L \quad \forall j \text{ and } \forall i \quad (27)$$

$$x_{ij}^k \in (0,1) \quad \forall j, \forall i \text{ and } \forall k \quad (28)$$

$$Y_{ij} \in \text{Integers} \quad \forall j \text{ and } \forall i \quad (29)$$

$$Z_{ij} \in \text{Integers} \quad \forall j \text{ and } \forall i \quad (30)$$

$$T_i \geq 0 \quad \forall i \quad (31)$$

Restriction (11) assures that each node is visited for just one vehicle, while restriction (12) guarantees that it is the same vehicle that arrives and departs from each node. The restriction (13) shows the numbers of started routes while (14) assure that only  $K$  vehicles are available.

Restrictions (15) and (16) specify that the time distances between the nodes result from the sum of course and service time.

The equation (17) is a stock equation where the route time is updated summing the course time and the time of vehicle immobilization corresponding to the waiting to reach next node. Restriction (18) guarantees the delivery time windows. Restriction (18) and (19) assure the fulfilment of the pick up time windows.

In order to ensure the pick up, delivery and the update of flow as well as stock functions, restrictions (20) and (22) are used. Equations (22) and (23) ensure the completeness of pick ups destinies and delivery origins. Equation (25) ensures the fulfilment of capacities restrictions associated to each vehicle.

Equation (26) assures the precedence of pick up concerning the delivery while restriction (27) limits the maximum travelling time per passenger.

Finally, equations (28), (29), (30) and (31) define the nature of the decision variables integrated in the problem.

## Solution Algorithm

One approach for the presented problem could pass by the enumeration of all admissible solutions and choosing the ones that achieved the best numerical results for the objectives in cause. However, this method is too inefficient in computational terms and unbearable for route planning with many passengers. This because the problem can be labelled as NP-hard

combinatory, in which the solution range increases at an exponential rate with the increase of variables in analysis (Lenstra and Rinnooy Kan, 1981).

In this paper we present a heuristic algorithm in which the achievement of the optimal solution it is not guaranteed. However, the algorithm produces as output a set of admissible solutions that have an expectable closure to the optimum for each objective and are obtained with a very efficient procedure in computational terms.

The algorithm is constructed considering two major blocks (Figure 1): Construction and Exploring.

### **Figure 1 – Algorithm Description**

The construction block is responsible for the creation of an initial solution via an *Initial Heuristic* module, where all passengers with origin and destiny nodes are allocated to a certain vehicle. In this initial allocation, all restrictions are respected except the number of available vehicles. As an extreme the result could have as many routes as the number of passengers. This restriction will be recovered in the *Route Linking* module.

The exploring block has as unique module the *Solution Optimization*, in which a Simulated Annealing algorithm is used as solution technique to optimize the actual solution. Exchanges and transfers between the passengers' positions among the routes are tested. The possibility of non admissible solutions could be accepted, as current solutions, and a penalization in objective function evaluation is then considered.

#### **Initial Heuristic**

Within the initial heuristic the initial allocation is preceded by passengers and vehicles sorting. The passengers arrangement has as criteria the following aspects:

- Destiny time window (Ascending);
- Time distance between origin and destiny (Descending);
- Difference between destiny and origin time windows (Ascending).

As to the vehicles sorting, the only criterion is their passenger transportation capacity.

After obtaining the separate lists it is necessary to proceed the initial allocation (Figure 2). The first node to be placed is the origin of the first passenger in the list. For this introduction it is verified the fulfilment of all restrictions and if this is not possible the next position will be tested (Step 1). After origin introduction, the associated destiny is tested as close right neighbour in the same route. Every restriction is validated in order to proceed to the allocation

and if any is not respected the allocation in the next position is tested (Step 2). After testing every position and if the problem is still impossible, the allocation of origin and destiny is tested in a next route (Step 3). An illustrative example is shown in (Figure 2).

### **Figure 2 – Initial heuristic example**

As output from this block will result a set of routes combining a solution, respecting all restrictions but the number of available buses. The respect of the origin time window will always be done and can be responsible for the creation of stopping moments during the route.

### **Route Linking**

Whenever the initial solution uses virtual vehicles, that is, when the used vehicles are more than the available, it is necessary to link so many routes as the exceed number. This module is also used whenever necessary to evaluate a new solution set with a less vehicle.

Before linking virtual routes created in the previous module, it is necessary to sort the vehicles. This is done through the following criteria:

- Vehicles capacity (Ascending);
- Number of passengers on route (Ascending);
- Route time length (Ascending);
- Passengers travelling time (Ascending).

The route linking procedure is shown in Figure 3, where the first two routes from the sorted list are selected (Step 1) and the first from the list will be collated to the end of the second one, resulting in a single route (Step 2). This procedure will be repeated until the number of routes is equal to the number of available vehicles, and the capacity of this merged vehicle is equal to the higher capacity of the involved vehicles.

### **Figure 3 – Route linking example**

After this module, it's only guaranteed the fulfilment of number of vehicles restriction and the sequentially relations between origin and destiny nodes.

### **Solution Optimization**

As referred above the meta-heuristic Simulated Annealing, initially presented in Kirkpatrick *et al.* (1983) and in Cerny (1985), is used in order to optimize a set of routes for the established objectives: total course time (30) and average travelling time per passenger (31).

$$FO_1^k = FO_1 + k \sum (UnrespConst) \quad (30)$$

$$FO_2^k = FO_2 + k \sum (UnrespConst) \quad (31)$$

where:

$FO_1^k$  - Value of penalized objective function for total course time;

$FO_1$  - Value of objective function for total course time;

$FO_2^k$  - Value of penalized objective function for travelling average time per passenger;

$FO_2$  - Value of objective function for travelling average time per passenger;

$K$  - Penalization coefficient;

$UnrespConst$  - Number of non respected restrictions in solution.

For single objective problems, the acceptance of the candidate solution as current solution is given through the comparison of the value of  $p$  with a randomly generated number with a (0-1) uniform distribution (32).

$$p = \min \left\{ 1, \exp \left( \frac{F(Y^*) - F(Y)}{T} \right) \right\} \quad (32)$$

In this multi-objective problem the solution evaluation is done through the comparison of the value of  $p$  with a random number.

$$p1 = \min \left\{ 1, \exp \left( \frac{FO1^{k*} - FO1^k}{T} \right) \right\} \quad (33)$$

$$p2 = \min \left\{ 1, \exp \left( \frac{FO2^{k*} - FO2^k}{T} \right) \right\} \quad (34)$$

$$p = \frac{p1 + p2}{2} \quad (35)$$

Both objectives are evaluated (33 and 34) simultaneously in a single solution, and the acceptance probability arises from the average of individual probabilities of each objective (35) when evaluated in a single and individual way (Suman and Kumar, 2005).

The Simulated Annealing is based on two fundamental concepts that were worked and analysed on this application: diversification and intensification. In the algorithm, the diversification is achieved with the "temperature" evolution (Figure 4). A bigger value corresponds to a more probable acceptance of a worst solution, but more important, widely diverse solutions are considered. Whenever an admissible solution is reached, the algorithm temperature is increased to one thousand times its value (*partial reheating*) and by *total reheating*, it is understood the procedure of re-established the initial temperature with the intention of a completely new solution search. These procedures enable the algorithm to diversify the search through a wider range of potential solutions.

**Figure 4 – Temperature evolution**

The concept of intensification is worked using searching techniques with different neighbourhood universes (Figure 5). In a first phase it is possible to explore exchanges and transfers between every element in the system (*Pure Randomly Generation*). However, every time the number of iterations, since the last decrease of non respected restrictions, reaches the value of a certain parameter it can only explore the elements that do not respect the problems restriction (*Restricted Randomly Generation*).

**Figure 5 – Intensification strategies**

This strategy is used until a decrease on the number of non respected restrictions is verified or until an admissible solution becomes the current solution. In both cases the algorithm temperature gets a partial reheating and *Pure Randomly Generation* is used again.

After the definition of which randomly generation type is to be used, two neighbourhood structures are designed to search solutions. In each iteration only one structure can be used. This is randomly selected amongst the two following:

- **Exchange passengers** between routes or its position within a route;
- **Transfer passenger** for another route or position within a route without exchange.

The criteria chosen for **exchange neighbourhood structure** definition are the following:

- Vehicle(s) selection;
- Origin nodes selection;

- Destiny nodes localization.

The vehicle selection is completely randomly with the probability linked with the course duration of each vehicle:

$$\frac{(CourseDur_i)^\lambda}{\sum_i (CourseDur)^\lambda} \quad (36)$$

where:

$CourseDur$  – Course duration of vehicle  $i$ ;

$\lambda$  - Parameter that defines the probability of  $i$  selection, between  $0 \leq \lambda \leq 1$ .

The passenger selection would be made through the selection of origin nodes and is also linked with the spent time on travelling:

$$\frac{(T_{Destiny} - T_{Origin})^\alpha}{\sum (T_{Destiny} - T_{Origin})^\alpha} \quad (37)$$

where:

$(T_{Destiny} - T_{Origin})$  - Passenger travelling time;

$\alpha$  - Parameter that define the probability of *origin node* selection, between  $0 \leq \alpha \leq 1$ .

The destiny node must be allocated to the new route and must in its position be restricted to the right of its associated origin node.

$$\frac{1}{RighPnt + 1} \quad (38)$$

where  $RighPnt$  is the number of pick up or deliveries nodes in the right of exchanged node.

In Figure 6, it is shown an example of the exchange structure neighbourhood procedure, where origin nodes from passenger A and C are randomly selected to exchange position between routes. Positions of related destiny nodes are also randomly selected from right neighbourhood of the exchanged node.

**Figure 6 – Passengers exchange**



The criteria to **transfer neighbourhood structure** definition are the following:

- Vehicle selection;
- Origin node to transfer selection;
- Origin node localization on new route;
- Destiny node localization on new route.

The origin node and vehicle selection is similar to the procedure describe in the previous neighboured structure. This is randomly selection depending on travelling time and course duration, respectively. In this structure is also possible to select the same vehicle for the transfer, enabling the possibility of a certain vehicle being passenger “exporter” and “importer”.

As to origin node a new localization will be selected according to the following probability:

$$\frac{1}{2NumPax + 1} \quad (39)$$

where *NumPax* is the number of existing passengers in “importer” route before transfer.

As to the destiny node a new localization is selected according to the following probability:

$$\frac{1}{RighPnt + 1} \quad (40)$$

where *RighPnt* is the number of pick up or deliveries nodes in the right of transferred node.

In Figure 7, there is an example of the transfer structure neighbourhood procedure, where the origin node from passenger A is randomly selected to be moved to vehicle 2, right next to passenger’s C origin. Position of A’s destiny is also randomly generated from the right neighbourhood space of transferred node.

**Figure 7 – Passengers transfer**

### **Routes Re-Linking**

In the problem in study, the solution is based on the achievement of the best results for the three different objectives. Since in the previous module, two of then are closely worked out in the *Routes Linking module* only the vehicle number minimization is now considered.

As result of the Solution Optimization, the best solutions for each objective for a certain number of vehicles are saved. The best solution for each one will be used as seed for the creation of a new current solution with less one vehicle.

The rules for this module are similar to the ones explained before in the *Route Linking* module. This cycle will be used as long as admissible solutions are obtained in the previous *Solution Optimization* or when the current number of vehicles is equal to one.

## Experimental Results

The algorithm presented is applied to a real case-study of a Portuguese school bus company. Therefore, the collected data is based on real routes of an area of the city of Lisbon. The transport of 28 passengers through 4 vehicles, each one with a 6 person transport capacity is considered. The planning period is the morning block, between 6:30 and 10:00 hours.

The primary objectives values from the current real routes are presented in Table 1. It can be highlighted the 403 minutes for total course time and the 24,29 minutes per passenger average travelling time.

**Table 1 – Primary objectives values of the real problem**

The secondary indicators values collected are presented in Table 2. The vehicle capacity utilization rate is 28,12%, the average passenger waiting time is 7,39 minutes, the early arrivals is of 31,32 minutes and the vehicle stopping time during course is 65 minutes.

**Table 2 – Secondary objectives values of the real problem**

The application of the proposed algorithm resulted in better results that are presented Table 3 for the five best solutions for each objective achieved (course duration and average travelling time).

**Table 3 – Solutions for 28 passengers and 4 vehicles**

When comparing the Course Duration best result (391 min) with the current real solution (403 min), an improvement of 2,97% is observed. As to the Average Travelling Time it is possible to create 28,69% shorter routes spending on bus around 17 minutes.

Furthermore, for the same group of passengers it was identified the possibility of just using 3 vehicles to execute the transport. The five best results respectively for course duration and average travelling time are presented in Table 4.

**Table 4 – Solutions for 28 passengers and 3 vehicles**

According to results achieved (Table 4), it is possible to improve the Course Duration in 4,96% and the Average Travelling Time in more than 22% when compared with current practise. However, and more important than the objectives improvement is the fact that the actual transport could be done with just 3 vehicles, while guaranteeing the service level to customers and saving relevant costs.

It is clear to observe an evident trade off between the two analysed objectives. With an increase on Average Travelling Time there is a significant reduction in Course Duration and vice-versa (Figure 8).

**Figure 8 – Trade off between objectives**

To a better evaluation of the algorithm performance on a wider range of situations, it was defined a partition on collected routes and the values of the primary and secondary objectives are presented in Table 5:

- Nine passengers (Current route 3);
- Sixteen passengers (Current route 2 + route 3);
- Twenty two passengers (Current route 2 + route 3 + route 4).

**Table 5 – Current partition solution**

In Table 6 are presented the bests results for each objective and for each partition on initial sample. As can be observed the algorithm reach for every case better results than the current solution.

**Table 6 – Achieved partition solution**

## Results Discussion

In Cordeau *et al.* (2002) is described an analysis of heuristics for the VRP resolution compiled from the literature. Recovering this methodology, the developed algorithm is evaluated according to the following criteria: efficacy, efficiency, simplicity and flexibility.

As to the efficacy, it is relevant to compare the obtained results with the optimal solution to each tested data instance. However, there is no clear idea about which is the optimal solution, in this circumstance and therefore the comparison will be exclusively done with the actual routes practised by the company. For these cases the algorithm achieved, for every instance, better results for both objectives when compared with the solution presented by the company (Table 7).

**Table 7 – Collected values and achieved results**

The tests were executed in a computer with a Pentium IV with 2.8Ghz processor, which represent that for a problem with 28 passengers 1 000 iterations consume 7,4 seconds of CPU time.

**Figure 9 – Evolution of CPU time by problem dimension**

It is clear the quasi linear relation between CPU time and the dimension of problem (Figure 9). This relation reveals a certain inefficiency of algorithm when planning high complexity routes however it was also identified that very good solutions are achieved before 100 000 iterations (12 minutes of processing), which it is far more efficient when compared with the time spent by the actual manual procedure of planning routes (more than 12 men working hours).

As to the simplicity criterion the algorithm can be evaluated according to the number and complexity of parameters. In this case even that existing 11 parameters to define for each execution, it was detected an evident stability of the performance of algorithm to most of them.

Finally, the flexibility can be explained by the capacity of the algorithm to adapt to real data. The case study presented in this paper is referring to collected data from a company operating in city of Lisbon and it is also expectable that the algorithm will have good performances with other data instance as long as they respect the data input structure.

## Conclusions

A heuristics algorithm to model the routing and schedule of school buses has been developed along this paper. Three different objectives were considered: number of vehicles, route total time and average passenger travelling time. A combined multiobjective function is defined where different parameters are used to model each criterion importance.

The proposed heuristic algorithm combines a set of heuristic procedures with a simulated annealing algorithm. Being based in a metaheuristic, the best result achieved can not be catalogued as even a local optimum for a certain objective, however it is extremely efficient searching for admissible solutions, which in highly constrained problem is an added value specially. This was particular clear for the case-study solved when comparing the algorithm performance with the actual planning situation in the company.

The obtained results compared with the actual solutions represent huge improvements to each objective, and will serve as a useful tool to the decision aid process, complementing the planner personal experience. It is also relevant to highlight the compute efficiency that allows the quick comparison of multiple scenarios.

## References

- Bennet, B.T. and Gazis, D.C. (1972). School bus routing by computer. *Transport Research* 6, 317-326.
- Bodin, L.D. and Berman, L. (1979). Routing and scheduling of school buses by computer. *Transport Science* 13, 113-129.
- Bowerman R., Hall, B., and Calamai, P. (1995). A multi-objective optimization approach to urban school bus routing: formulation and solution method. *Transport Research* 29A, 107-123.
- Braca, J., Bramel, J., Poser, B. and Simchi-Levi, D. (1994). A computerized approach to the New York city school bus routing problem. Graduate School of Business, Columbia University, NY.
- Cerny, V. (1985). A thermodynamical approach to the travelling salesman problem: An efficient simulated algorithm. *Journal of Optimization Theory and Applications* 45, 41-51.

Clarke, G. and Wright, J.W. (1963). Scheduling of vehicles from a depot to a number of delivery points. *Operations Research* 12, 568-581.

Corberán, A., Fernández, E., Laguna, M. and Martí, R. (2002). Heuristic Solutions to the Problem of Routing School Buses with Multiples Objectives. *Journal of the Operational Research Society* 53, 427-435.

Cordeau, J.F., Gendreau, M., Laporte, G., Potvin, J-Y, and Semet, F. (2002). A guide to vehicle routing heuristics. *Journal of the Operational Research Society* 53, 512-522.

Dantzig, G.B. and Ramser, J.H. (1959). The truck dispatching problem. *Management Science* 6, 80-91.

Kirkpatrick, S., Gelatt Jr., C.D. and Vecchi, M.P. (1983). Optimization by simulated annealing. *Science* 220, 671-680.

Laporte, G. (1992). The Vehicle Routing Problem: An overview of exact and approximate algorithms. *European Journal of Operational Research* 59, 345-358.

Lenstra, J.K. and Rinnooy Kan, A. (1981). Complexity of vehicle routing and scheduling problems. *Networks* 11, 221-227.

Lin, S. (1965). Computer solutions of the travelling salesman problem. *Bell System Technology Journal* 44, 2245-2269.

Lyo, L. and Fu, Z. (2002). The school bus routing problem: a case study. *Journal of the Operational Research Society* 53, 552-558.

Newton, R.M. and Thomas, W.H. (1974). Bus Routing in a Multi-School System. *Computers & Operations Research* 1, 213-222.

Spada, M., Bierlaire, M. and Liebling, T.(2003). Decision-aid Methodology for the School Bus Routing and Scheduling Problem. *Swiss Transport Research Conference*.

Suman, B. and Kumar,P. (2005). A survey of simulated annealing as a tool for single and multiobjective optimization. *Journal of Operational Research Society*, 1-18.

## Figures

Figure 1

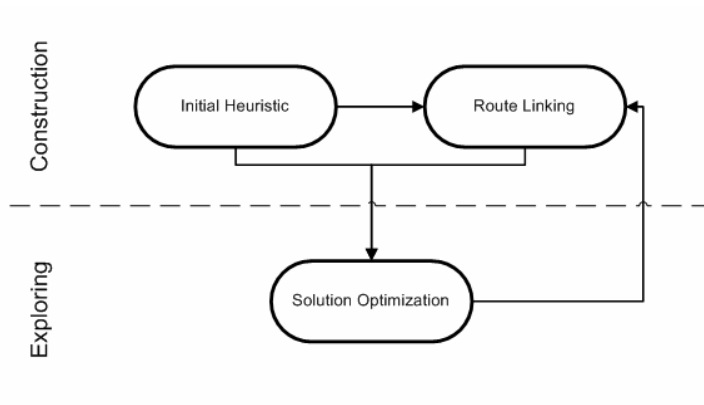


Figure 2

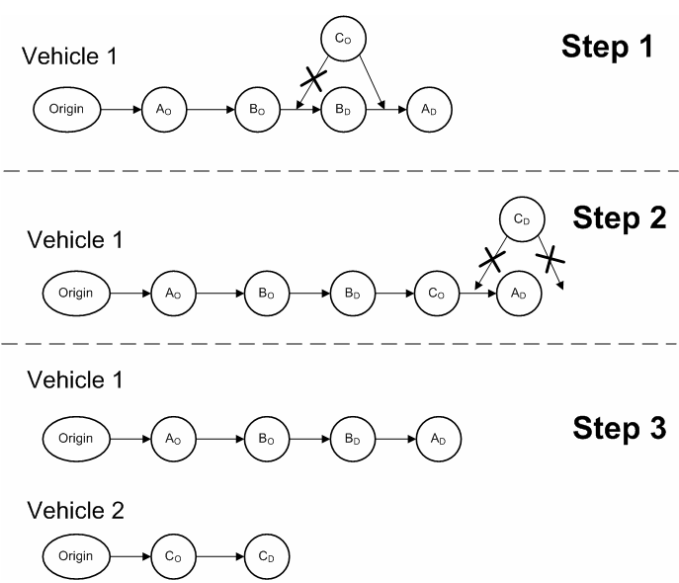




Figure 3

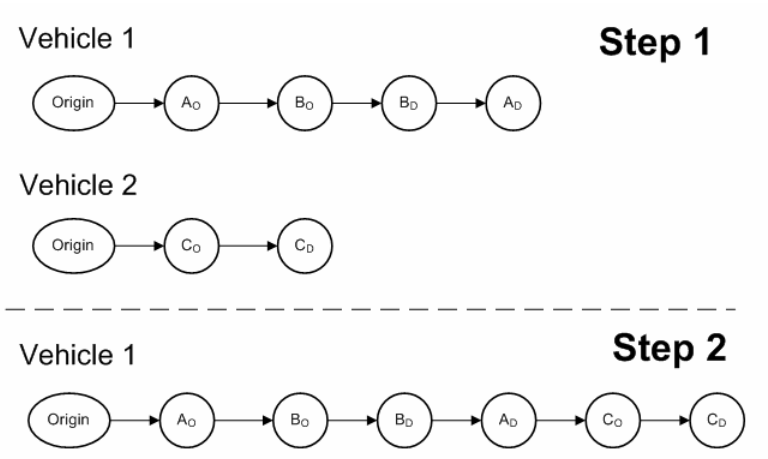


Figure 4

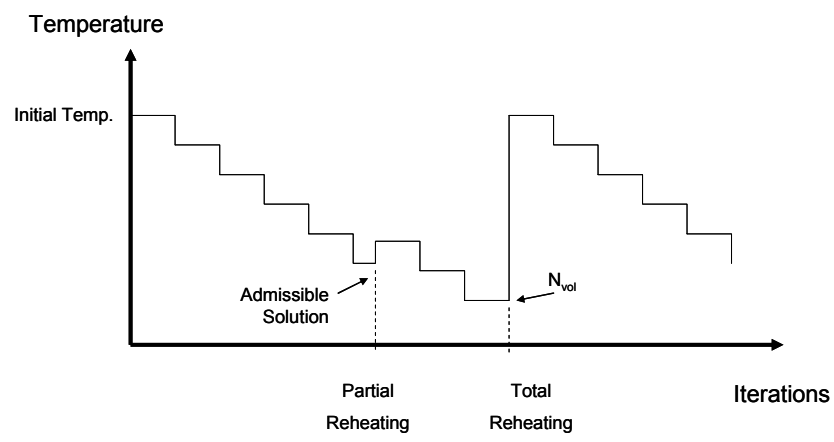


Figure 5

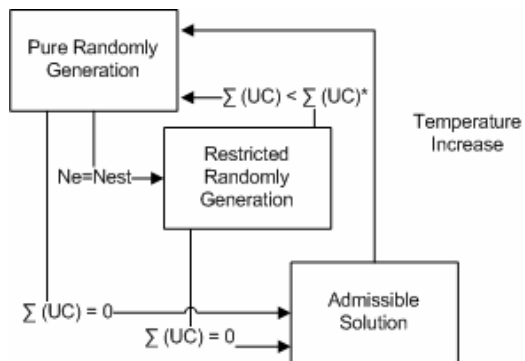


Figure 6

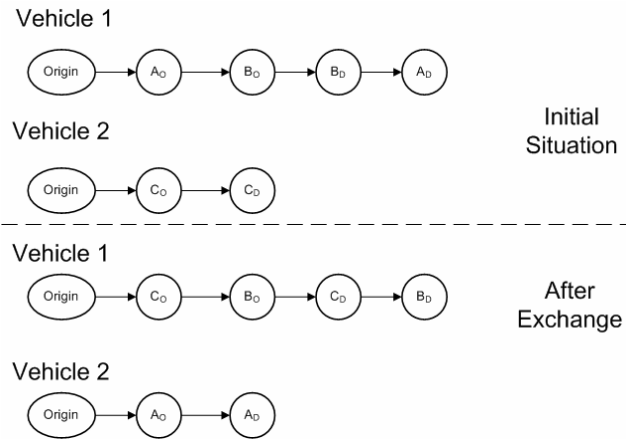


Figure 7

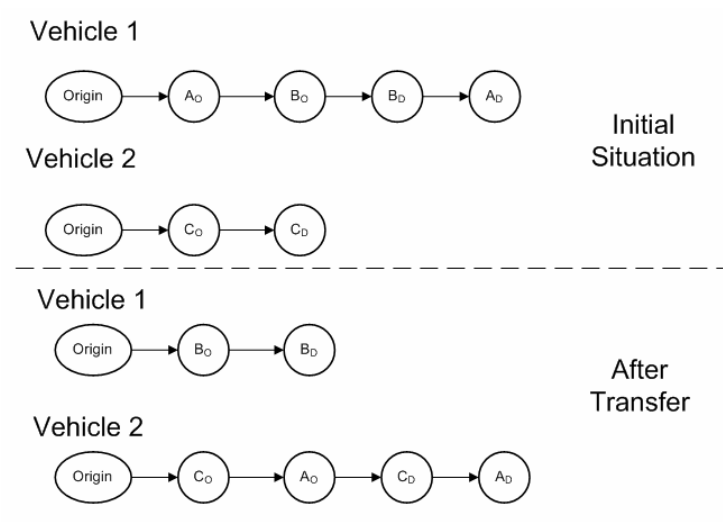


Figure 8

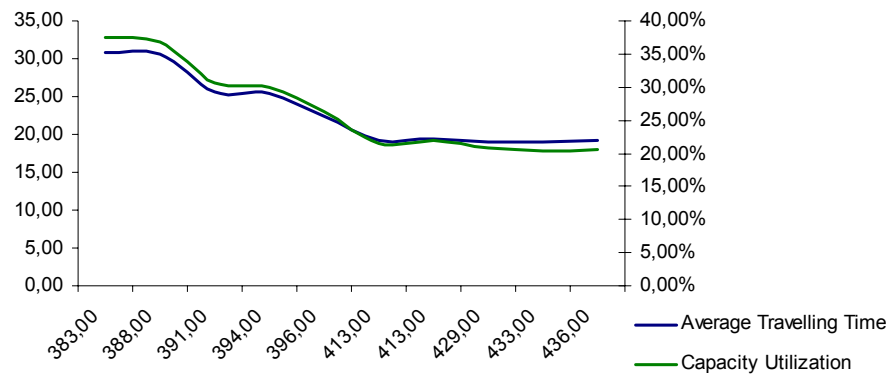
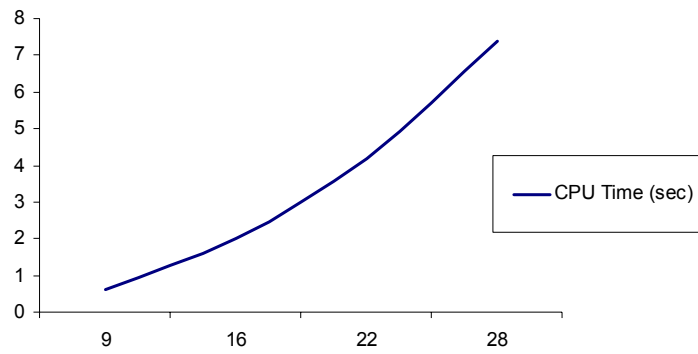


Figure 9



## Tables

Table 1

	<b>Course Duration (min)</b>	<b>Average Travelling Time (min)</b>
Route 1	87	12,83
Route 2	97	29,57
Route 3	130	29,67
Route 4	89	21,50
<b>Total</b>	<b>403</b>	<b>24,29</b>



Table 2

	<b>Passengers</b>	<b>Capacity Utilization</b>	<b>Waiting Time (min)</b>	<b>Early Arrivals (min)</b>	<b>Stopping Time (min)</b>
Route 1	6	14,75%	9,17	18,33	50
Route 2	7	35,57%	5,71	32,57	0
Route 3	9	34,23%	2,78	32,22	15
Route 4	6	24,16%	14,50	41,50	0
<b>Total</b>	<b>28</b>	<b>28,12%</b>	<b>7,39</b>	<b>31,32</b>	<b>65</b>

Table 3

<b>Course Duration (min)</b>	<b>Average Travelling Time (min)</b>	<b>Passengers</b>	<b>Utilization Capacity</b>	<b>Waiting Time (min)</b>	<b>Early Arrivals (min)</b>	<b>Stopping Time (min)</b>
391,00	21,18	28	25,28%	15,50	24,21	76,00
396,00	22,25	28	26,22%	15,14	23,50	76,00
399,00	25,68	28	30,03%	10,21	25,00	115,00
401,00	25,75	28	29,97%	10,29	24,86	115,00
406,00	21,18	28	24,34%	13,79	25,93	81,00
512,00	17,32	28	15,79%	12,71	30,86	115,00
486,00	17,36	28	16,67%	10,07	33,46	88,00
483,00	17,50	28	16,91%	10,07	33,32	91,00
525,00	17,86	28	15,87%	13,57	29,46	105,00
582,00	18,04	28	14,46%	14,86	28,00	33,00

Table 4

<b>Course Duration (min)</b>	<b>Average Travelling Time (min)</b>	<b>Passengers</b>	<b>Utilization Capacity</b>	<b>Waiting Time (min)</b>	<b>Early Arrivals (min)</b>	<b>Stopping Time (min)</b>
383,00	30,82	28	37,55%	9,54	20,54	51,00
388,00	30,64	28	36,86%	10,07	20,18	51,00
391,00	25,57	28	30,52%	14,36	20,96	25,00
394,00	25,36	28	30,03%	14,82	20,71	25,00
396,00	22,32	28	26,30%	9,32	29,25	48,00
433,00	18,93	28	20,40%	15,00	26,96	63,00
429,00	19,07	28	20,75%	15,50	26,32	60,00
413,00	19,11	28	21,59%	14,64	27,14	68,00
436,00	19,14	28	20,49%	16,36	25,39	53,00
413,00	19,46	28	21,99%	14,64	26,79	53,00

Table 5

Course Duration (min)	Average Travelling Time (min)	Passengers	Utilization Capacity	Waiting Time (min)	Early Arrivals (min)	Stopping Time (min)
130	29,67	9	34,23%	2,78	32,22	15,00
227,00	29,63	16	34,82%	4,06	32,38	15,00
316,00	27,41	22	31,91%	6,91	34,86	15,00

Table 6

Course Duration (min)	Average Travelling Time (min)	Passengers	Utilization Capacity	Waiting Time (min)	Early Arrivals (min)	Stopping Time (min)
105,00	36,67	9	52,38%	3,89	21,67	50,00
125,00	23,89	9	28,67%	3,33	35,00	25,00
197,00	33,06	16	44,75%	8,44	21,94	50,00
252,00	23,06	16	24,40%	14,81	25,56	10,00
307,00	27,09	22	32,36%	12,27	26,77	50,00
402,00	21,23	22	19,36%	15,32	29,59	20,00

Table 7

Passengers	Vehicles	Objectives	Collected Values	Achieved Results	Gap %
9	1	Course Duration (min)	130	105	19,23%
		Average Travelling Time (min)	29,67	23,89	19,47%
16	2	Course Duration (min)	227,00	197	13,22%
		Average Travelling Time (min)	29,63	23,06	22,16%
22	3	Course Duration (min)	316,00	307	2,85%
		Average Travelling Time (min)	27,41	21,23	22,54%
28	3	Course Duration (min)	-	383	-
		Average Travelling Time (min)	-	18,93	-
	4	Course Duration (min)	403	391	2,98%
		Average Travelling Time (min)	24,29	17,32	28,68%

## List of Captions

**Figure 1 – Algorithm Description**

**Figure 2 – Initial heuristic example**

**Figure 3 – Route linking example**

**Figure 4 – Temperature evolution**

**Figure 5 – Intensification strategies**

**Figure 6 – Passengers exchange**

**Figure 7 – Passengers transfer**

**Figure 8 – Trade off between objectives**

**Figure 9 – Evolution of CPU time by problem dimension**

**Table 1 – Primary objectives values of the real problem**

**Table 2 – Secondary objectives values of the real problem**

**Table 3 – Solutions for 28 passengers and 4 vehicles**

**Table 4 – Solutions for 28 passengers and 3 vehicles**

**Table 5 – Current partition solution**

**Table 6 – Achieved partition solution**

**Table 7 – Collected values and achieved results**