A Learning Approach to the School Bus Routing Problem

Kristof Van Moffaert [a] Bert Van Vreckem [a] Mihail Mihaylov [a]
Ann Nowé [a]

[a] *CoMo, Vrije Universiteit Brussel, Pleinlaan 2, Elsene, Belgium*

Abstract

In this paper, we introduce a solution to the School Bus Routing problem (SBRP), using a reinforcement learning technique that we previously applied in the domain of transportation logistics. This approach consists of bundling transportation requests between several locations in order to construct combinations of items in a cost-efficient way. We investigate how the combination of reinforcement learning and our novel bundling algorithm can be used to increase the efficiency in logistics and how it can be applied to other domains. In particular, we discuss how the SBRP can be transformed to resemble a problem in transportation logistics and thus solve the SBRP using our combined technique. We obtain results comparable to those presented in literature, namely from a cost minimization approach that is specifically tailored to the SBRP. We conclude that our reinforcement learning and bundling algorithms are both flexible enough to be applied in different domains and efficient in terms of the cost reduction that they offer.

1 Introduction

Supply chain management and transportation are part of daily operations of many companies and are crucial aspects of the delivery process of goods. Organizing the transportation activities of a manufacturer for any kind of good it produces, is of great influence to its turnover. In the following sections, we will briefly cover some trends in transportation logistics and present some relatively new ideas of cost-optimization. One of these proposals has been covered in the Master's thesis of the first author [1]. Our idea consists of increasing the transparency and professional organization of the stakeholders of the supply chain by learning how to bundle complementary items for transportation together. Thus, we construct bundles that have more value to the companies who will physically transport them, than the individual items. This approach is considered promising, because of its flexibility and the distributed nature of the solution. Nevertheless, this setting has not been tested on a large scale and on real-life data as logistics providers are not eager to publish their data and work methods. Therefore, we elaborate our ideas in [1] to solve an optimization problem, called the School Bus Routing Problem or SBRP. In this paper, we will tackle the problem using real-life data from a case study and compare our results to those in literature. More precisely, we show how similar results can be achieved using our flexible, distributed approach, compared to standard solution proposals for the SBRP.

2 Transportation logistics

Logistics is the part of supply chain management that plans, implements, and controls the efficient, effective, forward, and reverse flow and storage of goods services, and related information between the point of origin and the point of consumption in order to meet customers' requirements. In the following sections, we will focus on the two main stakeholders in the domain of transportation logistics, which are third-party logistics and fourth-party logistics providers.

2.1 Third-party logistics

As manufacturers require that their goods are located at the right time at the right place, they often choose to outsource transportation to consulting firms for the physical transportation of their products. In the literature, these companies are called third-party logistics, 3PLs or carriers. They own the necessary transportation equipment (e.g. trucks) to perform the transportation activity. The manufacturers are also called shippers in the supply chain. Over the world, there are thousands of carrier companies that compete with each other to obtain the transportation requests that fit their company's profile best. This process is called load allocation and is considered a very challenging part of the supply chain as improved coordination and distribution can be achieved by optimization techniques to improve the performance. For example, as cited by Robu [2], in the Netherlands in 2008, the average transportation performance was only optimal in about 40% to 60% of the cases. The main reason behind this performance issue lies in the fact that coordination and orchestration between the stakeholders of transportation logistics is missing and this implies a bad allocation of the carriers' services to transportation requests. As transportation needs become even more complex, another type of logistics company started to appear that solely focuses on orchestrating and coordinating the supply chain. This is the fourth-party logistics provider.

2.2 Fourth-party logistics

A fourth-party logistics or 4PL is not asset-based, e.g. it does not provide any physical transportation activity, but it is concerned with the orchestration of the supply chain. More precisely, it relies on the services of 3PLs, but enriches them by using computer systems and intellectual capital to optimize their operations and increase their cost-efficiency. The pressure on these few companies is very high, as the shippers' demands on fast, cost-efficient solutions are always present. Proposals in the literature on how the 4PL should operate and in what form these solutions should be provided are very minimal. Section 3 summarizes our solution approach to solve the problems concerning the competition and communication in transport logistics.

3 The bundling approach

Our contribution is a new proposal to increase the flexibility and coordination between shippers and carriers and consists of one-to-many negotiations. To accomplish this negotiation technique, the 4PL works in close collaboration with an auction house. The idea, illustratively presented in figure 1, consists of agents representing carriers that can connect to a virtual auction house, where they can place bids on items. The items themselves are bundles of transportation requests. These transportation requests describe transportation activities demanded by the manufacturers, which want to have their products transported. These requests consist of all the necessary data (e.g. geographical coordinates, volumes and timing constraints) to conduct the physical transportation. The bids, on the other hand, are the requested prices the 3PLs wish to receive in order to perform the physical transportation activity of the requests in the bundle. In order to determine this price (or bid), the 3PL estimates its costs of executing the delivery by examining the contents of the bundle, which is being auctioned. In many cases, a pick-up and delivery problem (PDP) solver is used to estimate the route(s) the company's truck(s) would have to drive to deliver all items in the bundle. Using this PDP solver, which finds its origin in solving Vehicle Routing Problems (VRP), the cost of the route can be determined, by taking into account the distance of the route, the number of trucks and other criteria such as an expected profit margin.

The profit of the 4PL depends on the money it receives from its customers, e.g. the shippers, that wish to communicate with carriers in an indirect way and be able to address multiple 3PLs at the same time. This amount is referred to as P_{cust} in equation 1. Also, the bids (B_{3PL}) in the auction house are recorded and the lowest bid, b, for a given bundle is determined the winner. The utility function of the 4PL, U_{4PL}, is presented below.

$$U_{4PL} = P_{cust} - b \tag{1}$$

$$\text{where } b = \min B_{3PL}$$

From equation 1, it is obvious that the lower the winning bid of the 3PL, the higher the 4PL's profit. An idea, which we elaborate in [1], consists of learning which bundles are considered "attractive" for carriers and which are not. A bundle's attractiveness can reflect on many aspects of the bidder's profile. For example, a bundle of transport items might fit the bidder's preferences on volume and distance criteria. Volume characteristics of a bundle can be interesting when the bidder's company can reduce its empty truck space and ideally, can fill its trucks up to 100% capacity utilization. Also, the distance specifications of the bundle can be attractive when the requests in the bundle can be covered in one working day by a single truck driver. Thus, avoiding the company to arrange expensive sleeping accommodation for the truck driver.

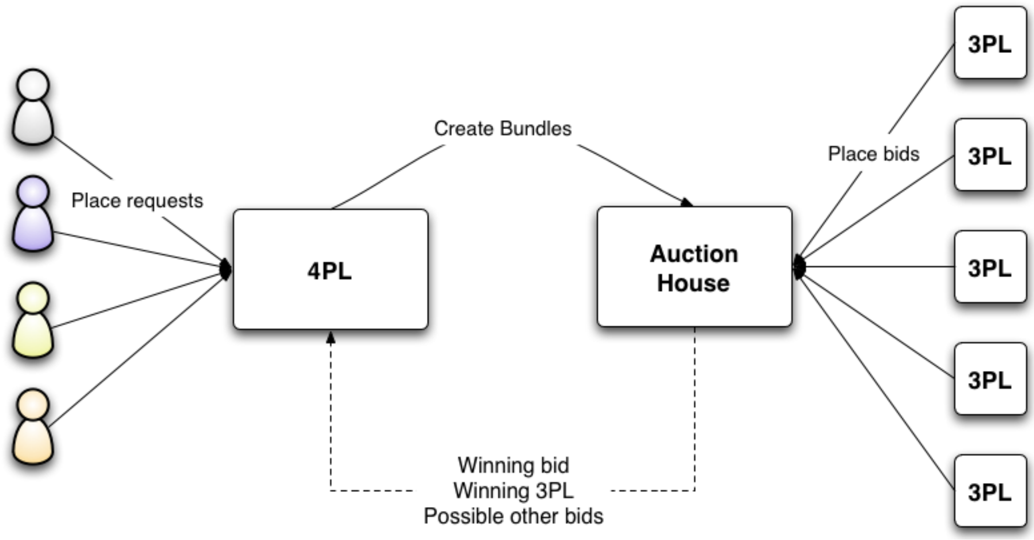

Figure 1: An illustrative representation of our proposal, elaborated in [1]. On the left hand side, the customers (shippers) provide the 4PL with thousands of transportation requests, for which it needs to find carriers. On its turn, the 4PL learns to create desirable bundles out of these requests and relies on an auction house that allows agents, representing interested 3PLs, to connect to and place bids.

The main difficulty in learning which properties of bundles significantly influence the 3PL's expenses and indirectly also its bids, consists of the fact that very limited information is available on these preferences. In fact, the only information that is public and can be retrieved by the 4PL is bidding information, as depicted in figure 1. More precisely, only the winning bidder and its bid are recorded, together with some possible other bids from other companies. No information is available on the properties the 3PLs take into account while calculating their bids because it is private. It is obvious that details on the number of available trucks, their capacity and fuel consumption specifications heavily influence the expenses of the company when transporting goods. Also the location of the depots of the carriers is not published. If this would not be the case and the carriers make this information public, the 4PL could rather easily bundle items close to each depot, and these bundles would contain excellent distance characteristics. Unfortunately, this is not the case and one will have to rely on other techniques to extract these preferences.

A possible solution approach, to still learn the preferences of the 3PLs, only relying on the minimal bidding information, consists of using learning mechanisms from the area of artificial intelligence and machine learning. In [1], we constructed a knowledge base, consisting of three properties of bundles. These properties are location, distance and volume characteristics. Location is an important criterion of a bundle, as dividing the world into regions can provide information on so-called *fruitful* areas. These fruitful areas contain a high variety of other transportation requests in the neighborhood, and make it easier to find complementary requests that can reduce the total distance to be covered, specified in the bundle. Distance criteria also represent a significant part of a transport company's costs, as some carriers specify a maximum threshold and once this restriction has been breached, its costs increase exponentially. As specified earlier, possibly, sleeping accommodation needs to be arranged when the truck driver can not return home at the end of the day. Recording a bundle's volume is interesting to deduct capacity information of the trucks. This information can be used to construct bundles up to this threshold and not beyond it, as then, the bidder would have to arrange multiple trucks for the same bundle which increases the costs, of course.

We conclude that the advantage of this approach lies in the fact that the auction house allows a lot of flexibility, as it uses an open market where carriers can easily connect to and exit. Thus, positively influencing the distributed characteristics of the solution. Also, the rivalry between the competitors is not reduced by negotiating with each carrier individually. Instead, in the auction house, carriers compete with each other in obtaining interesting bundles, with competitive bids as a result. In the following section, we describe the details on the model which is used to deduct the location, distance and volume preferences of the bidders.

3.1 Reinforcement learning

We propose to use three reinforcement learning systems that each focus individually on the deduction of one particular criteria of a bundle. Each of these systems use a stateless, simple reinforcement learning technique to calculate rewards for each criterion. The bidding information. provided by the auction house, is the only feedback used in the system. The learning rule, which calculates the reward of a bundle, is identical in each learning system and consists of several statistical measures to deduct information from the bidding data. The winning bid is providing information on how desirable the bundle is for the winning bidder, where the mean, standard deviation of the available bids can be used to deduct information on the overall interest of the bidders on the bundle. Lastly, we include the the number of transportation requests in the reward signal in order to stimulate creating larger bundles. The rule is listed below.

$$\begin{aligned} reward = Weight_a \times (1 - \text{Win bid}) + Weight_b \times \text{Requests} \\ Weight_c \times (1 - \text{Avg bid}) + Weight_d \times (1 - \text{Std dev}) \end{aligned} \quad (2)$$

The values of $Weight_a$, $Weight_b$, $Weight_c$ and $Weight_d$ are selected on an experimental basis and assigned to 0.6, 0.2, 0.1 and 0.1, respectively. The update rule, similar in each individual system, is based on single-state Q-learning algorithm [4]. Each system updates the quality on a certain aspect f of a bundle using the previously calculated quality (Q'_f) for criterion f and the newly obtained $reward$. On its turn, f can be instantiated by either $location$, $distance$ or $volume$. The update rule listed in equation 4:

$$Q_f = (1 - \alpha) \times Q'_f + \alpha * reward \quad (3)$$

In our experiments, the learning rate α was determined empirically and set to 0.1. To obtain a total quality of bundle, we combine the hypothesis of each individual learning system by using weights.

$$Q^b = \beta \times Q_{location} + \gamma \times Q_{distance} + \delta \times Q_{volume} \quad (4)$$

In our experiments, β, γ and δ are assigned weights of 0.6, 0.3 and 0.1, respectively. The next part of the algorithm is the actual bundling itself, based on the information gathered from the bidding. This part is covered in section 3.2.

3.2 Bundling behavior

In section 3.1, we have provided insights how knowledge on the 3PLs' preferences can be extracted in an indirect way. We will now focus on how one can use these Q-values to bundle. Because we are dealing with a highly noisy system and the knowledge base is constructed in an indirect way, we propose to extensively rely on artificial intelligence methods that balance between exploitation and exploration. In the exploration phase we create bundles regardless of the advices of the knowledge base. Thus, constructing random bundles. In the exploitation phase, we search for each transportation request its so-called 'best matching bundle'. This is the bundle that fits best for the current request. By consulting the knowledge base, we recover the Q-values from equation 4 and greedily search for the bundle that yields the highest Q-value if the transportation request would be added. To counter the problems of adopting a full-greedy approach, we incorporated an ϵ-greedy action selection method [4].

Another problem concerning the creation of bundles is the mechanism to decide when a bundle is 'full enough'. This is an important aspect of the bundling method as the bundling system tends to keep adding items to bundles in a greedy manner. Thus, there is need for a discrimination factor to determine when it is opportune to create a new bundle out of the item. The two possibilities, e.g. adding the item to the best existing bundle or creating a new one, are considered by estimating their future reward. If the reward of the bundle, when the item was added is not significantly larger or improving than before, it might not be interesting to assign the item to this bundle, although the matching system chose it. The rule used to discriminate between these two possibilities is given below.

$$\text{Bundle's reward with item} > \text{Bundle's reward without item} \times \rho \quad (5)$$

As value for the ρ value, we used 1.5. A final outline of our proposed bundler is depicted in 1.

```
input  : S: a set of N transportation requests
output: B: a set of bundles of transportation requests
for j ← 1 to maximum_Simulation_Days do
    Mode ← determineMode(j);
    Bundle according to mode;
    if Mode == Exploration then
        B = makeRandomBundles (S);
    else
        Make smart bundles;
        for i ← 1 to Numberofrequests do
            X = findBestMatchingBundle (S[i]);
            if (simulateEffect (X, S[i])) then
                addToBundle (X, S[i]);
            else
                startNewBundle (S[i]);
            end
        end
    end
    return B;
end
```

Algorithm 1: Outline of the bundling algorithm

3.3 Results

Our bundling approach, elaborated earlier, promises to increase transparency and coordination between the stakeholders of the supply chain, but is not yet incorporated in the real world. This, because both valorization and evaluation of new and innovative cost-reduction methods in the supply chain is a difficult task. In [1], we were unable to obtain real-life data from logistics providers, that could be used in our experiments. To test the bundling proposal, we focused on the Li and Lim [1] instance set, commonly used to represent pick-up and delivery optimization problems (PDP). Pick-up and delivery locations represent origin and destination locations of items that need to be transported. Therefore, problem instances of a PDP instance set are similar to transportation requests, defined by the manufacturers in the supply chain. Some results on this setting are presented in figure 2. For 20 simulation days, we have collected the cost per kilometer for the bidders in the auction house, which we use as a performance measure. Over time, our goal is to see a reduction in this cost. Several bundling schemes next to our learning method have been compared, such as random bundling, single-item bundling and no bundling at all. The result denote that learning the preferences and bundling in accordance to them can significantly reduce the bidders' expenses.

Despite the promising results, a thorough experimental evaluation of the bundling approach using real-life data is essential to proving the practical value of our technique. Therefore, an attempt has been made to apply our solution approach from the field of learning how to bundle complementary items as a cost-reduction technique, to another domain of optimization. This problem is the School Bus Routing Problem, which we will explain in the following section. We will provide the reader insights in the similarities between the bundling problem in supply chain management and the SBRP.

4 The School Bus Routing Problem

The School Bus Routing Problem (SBRP) is a very practical problem but has not been tackled that often in the field of computer science. The problem is closely related to the standard Vehicle Routing Problem or VRP, which has been a popular research area for the last three decades. VRP is a problem which searches the optimal routes that a vehicle travels in order to serve customers residing in a geographically dispersed area. The SBRP has same characteristics with VRP in several ways; however, there are noticeable differences. While a typical VRP mostly deals with freight transportation, the SBRP is related to student transportation.

[1] http://www.sintef.no/Projectweb/TOP/Problems/PDPTW/

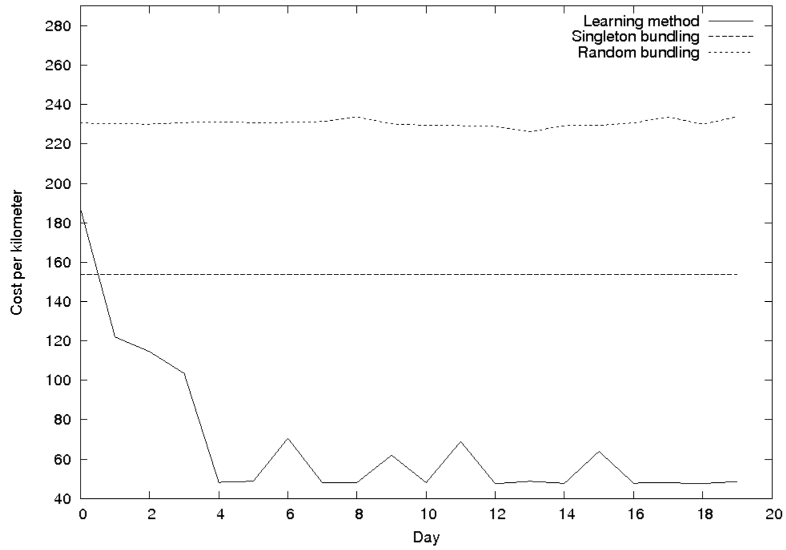

Figure 2: Learning the preferences of all bidders and bundling consequently significantly reduces the average cost per kilometer. The naïve bundling schemes such as random and single-item bundling were clearly outperformed by the learning method.

The School Bus Routing Problem (SBRP) can be specified as follows: a group of spatially distributed students must be provided with public transportation from their residencies to the school. Three factors make school bus routing unique: efficiency (the total cost to arrange a fleet of buses), effectiveness (how well the demand for service is satisfied) and equity (fairness of the school bus for each student). In this experiment we will focus on routing the buses to the bus stops.

In previous work, by Spasovic et al.[3], several attempts have been made to optimize bus routes from different neighborhoods to a single school. The case study discussed in the paper involves the Riverdale elementary school in New Jersey, where 199 students are bussed to the school in the morning. The school clustered the students into groups associated with 24 bus stops.

The school bus routing problems can be presented conceptually as a cost minimization problem, in which the objective function is the total operating cost. The following parameters are used in the mathematical formulation.

1. T_{max}, maximum time available for the bus to pick up students on a route
2. s_{ij}, distance between node i and j (in miles)
3. t_d, dwell time of the bus at a node (in hrs) per student. The time specified for one student to enter the bus, show his identification and to take seat
4. d_i, demand (number of students) to be picked up at node i
5. $\delta_{k,t}$ = 1 if bus route k has bus type t, 0 otherwise
6. O_t, operating cost for type t bus (in \$/bus-hr)
7. Cap_t, seat capacity for type t bus (in seats/bus)
8. V_t, average speed for bus of type t (in miles per hour)
9. $X_{ij,k}$ = 1 if nodes i and j are catered consequently by bus k.

The objective is to minimize total operating cost Z that is formulated as

$$Minimize\ Z = \sum_{k=1}^{l} t_k \left(\sum_{t} \delta_{k,t} O_t \right) \tag{6}$$

Here l represents the total number of bus routes, t represents the different types of buses and t_k is the time taken for bus k to pick up students on k_{th} bus route and drop students at node 0, e.g. the school. t_k is computed as

$$t_k = \sum_i \sum_j \left(X_{ij,k} s_{ij} / \sum_t \delta_{k,t} V_t \right) + 2 \times \sum_i \left(\sum_j X_{ij,k} \right) d_i t_d \tag{7}$$

In the case of the Riverdale elementary school, three types of buses are used, with a capacity of 54, 20 and 16, respectively. Their operational cost is \$60, \$50 and \$45 per hour, respectively.

Their results consist of three solution approaches, a time-savings heuristic, a mixed-integer linear program (MILP) called Router and a sweep method. These approaches consist of route optimization techniques and analytical methods. The results obtained are depicted in table 1. For more insights on the details of these three methods, we refer to the original paper [3].

Table 1: The results obtained by Spasovic et al. for the School Bus Routing Problem

Time-savings heuristics	ROUTER	Sweep method
\$112.90	\$109.50	\$113.65

The goal of this paper consists of applying the bundling method, explained in section 3, to the SBRP. We want to determine if the advantages of the bundling approach, consisting of flexibility and a distributed factor, can also be applied to a problem from another domain. In the following sections, we will present insights in the problem transformation from the bundling problem to the SBRP and discuss the results.

5 Problem transformation

In the bundling problem of the 4PL, shippers provide the 4PL with transportation requests that have to be assigned to carriers. In the SBRP, the transportation requests are represented by bus stops, each with a fixed amount of students to be picked up. These students will be picked up at their assigned bus stop and be transported towards the school. The carriers are represented by bus companies, each with a fleet of buses of a certain type with different characteristics, such as capacity and costs per hour. The 4PL, on its turn, learns over time which groups of bus stops are interesting to be grouped together, on location, distance and volume properties, as explained in section 3. The auction house will determine the winner of each bundle or route and notify the 4PL. This scenario is depicted in figure 3.

We recreated the setting by Spasovic et al. which consists of three bus types with a capacity of 54, 20 and 16 seats and a cost of \$$60$, \$$50$ and \$$45$ per hour, respectively. As in Spasovic's work, the 4PL receives the total lot of 24 bus stops.

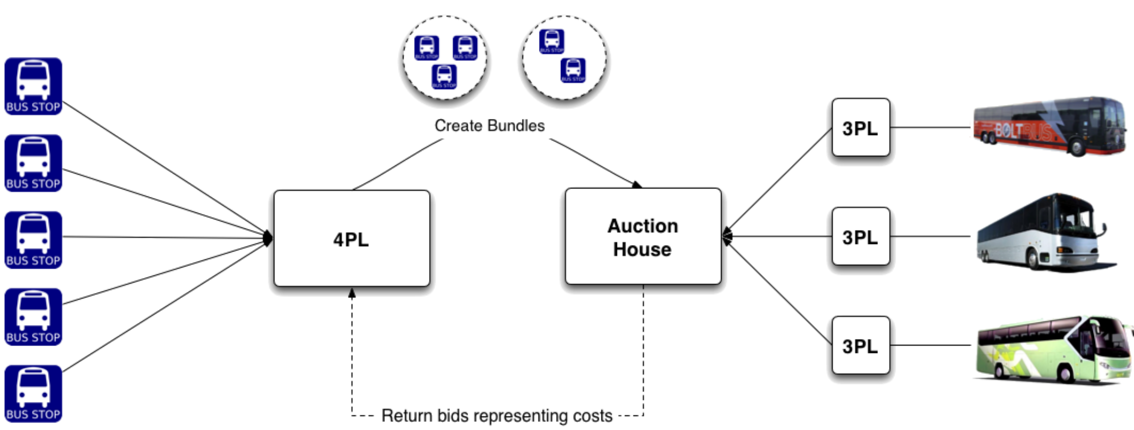

Figure 3: An illustrative representation of integration of the 4PL bundling solution to the School Bus Routing Problem. Each 3PL represents a different bus type, each with different characteristics and cost. The bids they place in the auction house represents the costs of assigning their bus type to the route in the bundle.

In our newly created setting, the 4PL applies its reinforcement learning algorithm to learn over time promising and interesting location, distance and volume characteristics of bundles of bus stops. Reinforcements are calculated for these criteria, based on the bids of the bus companies, calculated by equation 6. We ran the algorithm in the simulation environment and used its knowledge to create bundles, while relying on simple exploration techniques such as ϵ-greedy action selection [4]. The best result obtained consist of 6 routes for the different buses, with a total cost of \$109.86, which is very comparable to the solution obtained with Spasovic's experiments. We believe these results can be further improved by refining the PDP solvers, used by the 3PLs in their cost calculations, by adopting heuristic techniques, such as for example the time-savings heuristic, presented in Spasovic's work.

6 Discussion

In the solution approach by Spasovic, the school is the central component which runs the different PDP solvers on the benchmark with 24 bus stops. Each algorithm's goal consists of finding the best allocations of bus stops to three types of buses. Each of these mechanisms relies on specific information on the bus companies and tries to further optimize the routes, based on a certain objective function. Their traditional approach requires the school to acquire the information on the buses and the complete objective function to be known beforehand.

In the approach specified in this paper, the school does not act as a central component, but is considered the 4PL, provided with 24 transportation requests, representing bus stops. The school acts as the coordinator and is responsible for contacting interested transportation firms, in cooperation with the auction house. The auction house allows flexibility in contacting and negotiating with several transportation suppliers at the same time, as it uses an open market. Therefore, the approach, proposed in this paper is more distributed and does not require a central component to acquire information on all aspects a transportation company could take into consideration while computing its costs. Because the transportation companies are often not willing to publish detailed information on their work methods in the first place, we believe our proposal is much more practical to solving a real-life problem instance of the SBRP.

7 Conclusions

General PDP solvers require specific information to calculate and optimize routes between pick-up and delivery locations, using services of different transportation companies. This specific information contains details on the companies' trucks and cost. In a real-life setting, the transportation firms are not eager to publish this information and optimal service allocations are not possible.

In this paper, we have elaborated on a solution approach for the School Bus Routing Problem, based on a bundling technique we covered in [1]. We propose an indirect form of communication between the school and the transportation companies. Using an auction system, an open market is created where transportation companies can connect to and compete with other service providers. Thus, allowing a flexible and distributed setting, where a learning system is used to bundle the bus stops in accordance to the preferences of the transportation providers. We believe our bundling technique can also be applied to other problems that require the distribution of items into groups, such as for example the Set Covering Problem.

References

[1] Kristof Van Moffaert. Load allocation in transportation logistics using machine learning techniques. Master's thesis, Vrije Universiteit Brussel, 2011.

[2] Valentin Robu, Han Noot, Han La Poutré, and Willem-Jan van Schijndel. An interactive platform for auction-based allocation of loads in transportation logistics. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: industrial track*, AAMAS '08, 2008.

[3] Lazar Spasovic, Steven Chien, and Cecilia Kelnhofer-Feeley. A methodoly for evaluating of school bus routing - a case study of riverdale, new jersey. *TRB*, 2001.

[4] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, 1998.