A Simulation-based algorithm for solving the Vehicle Routing Problem with Stochastic Demands

Angel Juan * Javier Faulin † Daniel Riera * Jose Caceres * Scott Grasman ‡

* Open University of Catalonia - IN3
Barcelona, Spain
{ajuanp, drierat, jcaceresc}@uoc.edu

† Public University of Navarre
Pamplona, Spain
javier.faulin@unavarra.es

‡ Missouri University of Science & Technology
Rolla, MO, USA
grasmans@mst.edu

ABSTRACT

This paper proposes a flexible solution methodology for solving the Vehicle Routing Problem with Stochastic Demands (VRPSD). The logic behind this methodology is to transform the issue of solving a given VRPSD instance into an issue of solving a small set of Capacitated Vehicle Routing Problem (CVRP) instances. Thus, our approach takes advantage of the fact that extremely efficient metaheuristics for the CVRP already exists. The CVRP instances are obtained from the original VRPSD instance by assigning different values to the level of safety stocks that routed vehicles must employ to deal with unexpected demands. The methodology also makes use of Monte Carlo Simulation (MCS) to obtain estimates of the expected costs associated with corrective routing actions (recourse actions) after a vehicle runs out of load before completing its route.

Keywords: Metaheuristics, Routing, Scheduling

1. INTRODUCTION

The Vehicle Routing Problem with Stochastic Demands (VRPSD) is a well-known NP-hard problem in which a set of customers with random demands must be served by a fleet of homogeneous vehicles departing from a depot, which initially holds all available resources. There are some tangible costs associated with the distribution of these resources from the depot to the customers. In particular, it is usual for the model to explicitly consider costs due to moving a vehicle from one node -customer or depot- to another. These costs are often related to the total distance traveled, but they can also include other factors such as number of vehicles employed, service times for each customer, etc. The classical goal here consists of determining the optimal solution (set of routes) that minimizes those tangible costs subject to the following constraints: (i) all routes begin and end at the depot; (ii) each vehicle has a maximum load capacity, which is considered to be the same for all vehicles; (iii) all (stochastic) customer demands must be satisfied; (iv) each customer is supplied by a single vehicle; and (v) a vehicle cannot stop twice at the same customer without incurring in a penalty cost.

Notice that the main difference between the Capacitated Vehicle Routing Problem (CVRP) and the VRPSD is that in the former all customer demands are known in advance, while in the latter the actual demand of each customer has a stochastic nature, i.e., its statistical distribution is known beforehand, but its exact value is revealed only when the vehicle reaches the customer. For the CVRP, a large set of efficient optimization methods, heuristics and metaheuristics have been already developed ([1]). However, this is not yet the case for the VRPSD, which is a more complex problem due to the uncertainty introduced by the random behavior of customer demands. Therefore, as suggested by Novoa and Storer [2], there is a real necessity for developing more efficient and flexible approaches for the VRPSD. On one hand, these approaches should be efficient in the sense that they should provide optimal or near-optimal solutions to small and medium VRPSD instances in reasonable times. On the other hand, they should be flexible in the sense that no further assumptions need to be made concerning the random variables used to model customer demands, e.g., these variables should not be assumed to be discrete neither to follow any particular distribution. To the best of our knowledge, most of the existing approaches to the VRPSD do not satisfy this flexibility requirement.

The random behavior of customer demands could cause an expected feasible solution to become infeasible if the final demand of any route exceeds the actual vehicle capacity. This situation is referred to as "route failure", and when it occurs, some corrective actions must be introduced to obtain a new feasible solution. For example, after a route failure, the associated vehicle might be forced to return to the depot in order to reload and resume the distribution at the last visited customer. Our methodology proposes the construction of solutions with a low probability of suffering route failures. This is basically attained by constructing routes in which the associated expected demand will be somewhat lower than the vehicle capacity. Particularly, the idea is to keep a certain amount of surplus vehicle capacity (safety stock or buffer) while designing the routes so that if the final routes' demands exceed their expected values up to a certain limit, they can be satisfied without incurring a route failure.

2. BASIC NOTATION

The Stochastic Vehicle Routing Problem (SVRP) is a family of well-known vehicle routing problems characterized by the randomness of at least one of their parameters or structural variables [3]. This uncertainty is usually modeled by means of suitable random variables which, in most cases, are assumed to be independent. The Vehicle Routing Problem with Stochastic Demands

(VRPSD) is among the most popular routing problems within the SVRP family. There are two other classical problems belonging to that family: the Vehicle Routing Problem with Stochastic Customers (VRPSC) which was solved by Gendreau et al. [4] using an adapted Tabu Search, and the Vehicle Routing Problem with Stochastic Times (VRPST), but their applications are rather limited in comparison with the VRPSD, which is described in detail next.

Consider a complete network constituted by $n+1$ nodes, $V = \{0,1,2,\ldots,n\}$, where node 0 symbolizes the central depot and $V^* = V\backslash\{0\}$ is the set of nodes or vertices representing the n customers. The costs associated with traveling from node i to node j are denoted by $c(i,j)\ \forall i,j \in V$, where the following assumptions hold true: (i) $c(i,j) = c(j,i)$ (i.e., costs are usually assumed to be symmetric, although this assumption could be relaxed if necessary); (ii) $c(i,i) = 0$, and (iii) $c(i,j) \leq c(i,k) + c(k,j)\ \forall k \in V$ (i.e., the triangle inequality is satisfied). These costs are usually expressed in terms of traveled distances, traveling plus service time or a combination of both. Let the maximum capacity of each vehicle be $VMC >> \max_{i \in V^*}\{D_i\}$, where $D_i \geq 0\ \forall i \in V^*$ are the independent random variables that describe customer demands -it is assumed that the depot has zero demand. This capacity constraint implies that the demand random value never will be greater than the VMC value, which allows us an adequate performance of our procedure. For each customer, the exact value of its demand is not known beforehand but it is only revealed once the vehicle visits. No further assumptions are made on these random variables other than that they follow a well-known theoretical or empirical probability distribution -either discrete or continuous- with existing mean denoted by $E[D_i]$. In this context, the classical goal is to find a feasible solution (set of routes) that minimizes the expected delivery costs while satisfying all customer demands and vehicle capacity constraints. Even when these are the most typical restrictions, other constraints and factors are sometimes considered, e.g., maximum number of vehicles, maximum allowable costs for a route, costs associated with each delivery, time windows for visiting each customer, solution attractiveness or balance, environmental costs, and other externalities.

3. OUR SIMULATION-BASED APPROACH

Our approach is inspired by the following facts: (a) the VRPSD can be seen as a generalization of the CVRP or, to be more specific, the CVRP is just a VRPSD with constant demands –random demands with zero variance–; and (b) while the VRPSD is yet an emerging research area, extremely efficient metaheuristics do already exists for solving the CVRP. Thus, one key idea behind our approach is to transform the issue of solving a given VRPSD instance into a new issue which consists of solving several "conservative" CVRP instances, each characterized by a specific risk (probability) of suffering route failures. The term conservative refers here to the fact that only a certain percentage of the vehicle total capacity will be considered as available during the routing design phase. In other words, part of the total vehicle capacity will be reserved for attending possible "emergencies" caused by under-estimated random demands during the actual distribution (routing execution) phase. This part can be considered as a safety stock since it reflects the level of extra stock that is maintained to buffer against possible route failures. Next, the specific steps of our methodology are described in detail:

1. Consider a VRPSD instance defined by a set of customers with stochastic demands, where each demand is a random variable following a given statistical distribution –either theoretical or empirical as long as its mean exists.

2. Set a value k for the percentage of the maximum vehicle capacity that will be used as safety stock during the routing design stage.

3. Consider the CVRP(k) defined by: (a) the reduced total vehicle capacity, and (b) the deterministic demands given by the expected value of the real stochastic demands.

4. Solve the CVRP(k) by using any efficient CVRP methodology. Notice that the solution of this CVRP(k) is also an aprioristic solution for the original VRPSD. Moreover, it will be a feasible VRPSD solution as long as there will be no route failure, i.e., as long as the extra demand that might be originated during execution time in each route does not exceed the vehicle reserve capacity or safety stock. Notice also that the cost given by this solution can be considered as a base or fixed cost of the VRPSD solution, i.e., the cost of the VRPSD in case that no route failures occur. Chances are that some route failures occur during the execution phase. If so, corrective actions -such as returning to the depot for a reload before resuming distribution- and their corresponding variable costs will need to be considered. Therefore, the total costs of the corresponding VRPSD solution will be the sum of the CVRP(k) fixed costs and the variable costs due to the corrective actions.

5. Using the solution obtained in the previous step, estimate the expected (average) costs due to possible failures in each route. This can be done by using Monte Carlo simulation, i.e., random demands are generated and whenever a route failure occurs (or just before it happens), a corrective policy is applied and its associated costs are registered. In the experimental section of this paper, every time a route fails we consider the costs of a round-trip from the current customer to the depot; but, since we are using simulation, other alternative policies and costs could also be considered in a natural way. After iterating this process for some hundred/thousand times, a random sample of observations regarding these variable costs are obtained and an estimate for its expected value can be calculated.

6. Depending on the total costs associated with the solutions already obtained, repeat the process from Step 1 with a new value of k -i.e., explore different scenarios to check how different levels of safety stock affect the expected total cost of the VRPSD solution.

7. Finally, provide a sorted list with the best VRPSD solutions found so far as well as their corresponding properties: fixed costs, expected variable costs, and expected total costs.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In the CVRP literature, there exists a classical set of very well-known benchmarks commonly used to test their algorithm. However, as noticed by Bianchi et al. [5], there are no commonly used benchmarks in the VRPSD literature and, therefore, each paper presents a different set of randomly generated benchmarks. Thus, we decided to employ a natural generalization of several classical CVRP instances by using stochastic demands instead of constant ones. So, for each instance, while we decided to keep all node coordinates and vehicle capacities, we changed d_i, the deterministic demands of client i ($\forall i \in \{1,2,\ldots,\#nodes-1\}$) to stochastic demands D_i following an exponential distribution with $E[D_i] = d_i$.

For each instance, a total of 16 scenarios were simultaneously executed using a cluster of 16 personal computers Intel®Core^{TM}2 Quad Q8200 at 2.33GHz and 2GB RAM. The 16 scenarios were obtained by varying the available vehicle capacity (i.e., the complementary of the safety-stocks level) from 100% to 85% during the routing-design stage. Table 1 shows the complete results obtained for all 55 classical instances we generalized and tested.

The first column in Table 1 contains the name of each instance, which includes the number of nodes and also the number of routes of the 'standard' solution, e.g. B-n78-k10 is an instance of class B with 78 nodes and able to be solved with a 10-route solution.

Columns 2 to 4 are related to solutions obtained by our algorithm when a 100 % of the vehicle maximum capacity is considered during the design stage. Notice that this strategy always provides pseudo-optimal solutions in terms of fixed costs (Column 2), since they can be directly compared with the CVRP best-known solution. However, since no safety stock is used, there is a chance that these solutions can suffer from route failures. In turn, route failures might imply high expected variable costs (estimated in Column 3 by Monte Carlo simulation), thus increasing the total expected costs, which is estimated in Column 4. Here is where using safety stocks can be of value: by not necessarily using all vehicle maximum capacity during the design stage, some route failures can be avoided. Hopefully, this might lead to new solutions with slightly higher fixed costs but also with lower expected variable costs. At the end, these alternative solutions might present lower total expected costs, which are the ones to be minimized. On the one hand, columns 5 to 9 show the results obtained with our algorithm. Notice that fixed costs in Column 7 are always higher or equal to those in Column 2. However, total expected costs in Column 9 are always lower or equal to those in Column 4. Notice also that sometimes the best-found strategy (for this set of benchmarks) is to use a 100 % of the vehicle maximum capacity (i.e. no safety stocks at all) when designing the routes (Column 5).

5. CONCLUDING REMARKS

We have presented a hybrid approach to solving the Vehicle Routing Problem with Stochastic Demands (VRPSD). The approach combines Monte Carlo simulation with well-tested metaheuristics for the Capacitated Vehicle Routing Problem (CVRP). One of the basic ideas of our methodology is to consider a vehicle capacity lower than the actual maximum vehicle capacity when designing VRPSD solutions. This way, this capacity surplus or safety stocks can be used when necessary to cover route failures without having to assume the usually high costs involved in vehicle restock trips. Another important idea is to transform the VRPSD instance to a limited set of CVRP instances -each of them defined by a given safety-stocks level-, to which efficient solving methods can be applied. Our approach provides the decision-maker with a set of alternative solutions, each of them characterized by their total estimated costs, leaving to him/her the responsibility of selecting the specific solution to be implemented according to his/her utility function. Although other previous works have proposed to benefit from the relationship between the VRPSD and the CVRP, they usually require hard assumptions that are not always satisfied in realistic scenarios. On the contrary, our approach relaxes most of these assumptions and, therefore, it allows for considering more realistic customer demand scenarios. Thus, for example, our approach can be used to solve CVRPSD instances with hundreds of nodes in a reasonable time and, even more important, it is valid for virtually any statistical distribution –the one that best fits historical data on customer demands.

6. ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Ministry of Science and Innovation (TRA2010-21644-C03) and by the Navarrese and Catalan Governments (IIQ13172.RI1-CTP09-R2, 2009 CTP 00007 and Jerónimo de Ayanz network). This work has been developed in the context of the CYTED-IN3-HAROSA Network (http://dpcs.uoc.edu).

7. REFERENCES

[1] G. Laporte, "What you should know about the vehicle routing problem," *Naval Research Logistics*, vol. 54, pp. 811–819, 2007.

[2] C. Novoa and R. Storer, "An approximate dynamic programming approach for the vehicle routing problem with stochastic demands," *European Journal of Operational Research*, no. 196, pp. 509–515, 2009.

[3] C. Bastian and A. R. Kan, "The stochastic vehicle routing problem revisited," *European Journal of Operations Research*, vol. 56, pp. 407–412, 2000.

[4] M. Gendreau, G. Laporte, and R. SÈguin, "A tabu search heuristic for the vehicle routing problem with stochastic demands," *Operations Research*, vol. 44(3), pp. 469–477, 1996.

[5] L. Bianchi, M. Birattari, M. Chiarandini, M. Mastrolilli, L. Paquete, O. Rossi-Doria, and T. Schiavinotto, "Hybrid metaheuristics for the vehicle routing problem with stochastic demands," *Journal of Mathematical Modelling and Algorithms*, vol. 5, pp. 91–110, 2006.

	Using 100% of the Capacity			Using a Percentage P of the Capacity						
Instance	Fixed	Variable	Total (1)	P	Routes	Fixed	Variable	Total (2)	Time (s)	Gap (1) - (2)
A-n32-k5	787.08	179.49	966.57	100%	5	787.08	179.49	966.57	1	0.00%
A-n33-k5	662.11	159.77	821.88	97%	5	676.10	135.80	811.90	1	1.21%
A-n33-k6	742.69	162.45	905.14	100%	6	742.69	162.45	905.14	1	0.00%
A-n37-k5	672.47	134.43	806.89	97%	5	692.53	109.47	802.00	1	0.61%
A-n38-k5	733.95	157.48	891.43	93%	6	761.25	117.97	879.22	1	1.37%
A-n39-k6	835.25	178.10	1,013.35	94%	6	842.92	150.35	993.27	1	1.98%
A-n45-k6	944.88	254.68	1,199.55	94%	7	979.31	197.70	1,177.01	1	1.88%
A-n45-k7	1,154.39	325.68	1,480.07	100%	7	1,154.39	325.68	1,480.07	2	0.00%
A-n55-k9	1,074.96	304.33	1,379.28	100%	9	1,074.96	304.33	1,379.28	1	0.00%
A-n60-k9	1,362.19	395.42	1,757.61	100%	9	1,362.19	395.42	1,757.61	2	0.00%
A-n61-k9	1,040.31	288.01	1,328.32	95%	10	1,073.86	241.57	1,315.43	1	0.97%
A-n63-k9	1,632.19	518.31	2,150.50	100%	9	1,632.19	518.31	2,150.50	4	0.00%
A-n65-k9	1,184.95	341.43	1,526.37	99%	10	1,213.73	304.73	1,518.46	1	0.52%
A-n80-k10	1,773.79	548.84	2,322.63	100%	10	1,773.79	548.84	2,322.63	7	0.00%
B-n31-k5	676.09	169.46	845.54	95%	5	680.98	158.07	839.05	1	0.77%
B-n35-k5	958.89	267.77	1,226.66	99%	5	978.51	239.61	1,218.12	3	0.70%
B-n39-k5	553.20	142.48	695.68	100%	5	553.20	142.48	695.68	1	0.00%
B-n41-k6	834.92	248.30	1,083.22	96%	7	856.76	224.13	1,080.89	1	0.22%
B-n45-k5	754.23	146.48	900.71	100%	5	754.23	146.48	900.71	1	0.00%
B-n50-k7	744.23	202.85	947.07	93%	7	754.26	186.11	940.37	1	0.71%
B-n52-k7	754.38	204.83	959.21	92%	7	771.02	164.87	935.88	1	2.43%
B-n56-k7	716.42	211.94	928.36	88%	8	757.68	140.32	898.00	1	3.27%
B-n57-k9	1,602.28	559.89	2,162.17	96%	9	1,623.27	515.53	2,138.80	1	1.08%
B-n64-k9	868.40	277.39	1,145.79	100%	9	868.40	277.39	1,145.79	10	0.00%
B-n67-k10	1,039.46	316.59	1,356.05	100%	10	1,039.46	316.59	1,356.05	1	0.00%
B-n68-k9	1,283.16	442.17	1,725.33	97%	9	1,303.09	388.54	1,691.63	8	1.95%
B-n78-k10	1,245.82	367.24	1,613.06	98%	10	1,252.38	357.03	1,609.41	9	0.23%

Table 1: Results for instances A and B using exponentially distributed demands with $E[D_i] = d_i$