

ON THE USE OF BIASED RANDOMIZATION AND SIMHEURISTICS TO SOLVE VEHICLE AND ARC ROUTING PROBLEMS

Sergio Gonzalez-Martin
Barry B. Barrios
Angel A. Juan
Daniel Riera

Computer Science Department
IN3 – Open University of Catalonia
156 Rambla Poblenou
Barcelona, 08018, SPAIN

ABSTRACT

This paper reviews the main concepts and existing literature related to the use of biased randomization of classical heuristics and the combination of simulation with meta-heuristics (Simheuristics) in order to solve complex combinatorial optimization problems, both of deterministic and stochastic nature, in the popular field of Vehicle and Arc Routing Problems. The paper performs a holistic approach to these concepts, synthesizes several cases of successful application from the existing literature, and proposes a general simulation-based framework for solving richer variants of Vehicle and Arc Routing Problems. Also examples of algorithms based on this framework successfully applied to concrete cases of Vehicle and Arc Routing Problems are presented.

1 INTRODUCTION

There is an emerging interest on introducing randomization into combinatorial optimization problems as a way of describing new real problems in which most of the information and data cannot be known beforehand. This tendency can be observed in Van Hentenryck and Bent (2010), which provides an interesting review of many traditional combinatorial problems with stochastic parameters. Thus, those authors studied Stochastic Scheduling, Stochastic Reservations and Stochastic Routing in order to make decisions on line, i.e., to re-optimize solutions when their initial conditions have changed and, therefore, are no longer optimal. This type of analysis has designed the Online Vehicle Routing Problems, in which re-optimization is needed apart from a previous situation. This set of routing problems seems to be well analyzed with the use of stochastic hypothesis in their definitions, thus providing more reality in their formulation. Another routing field in which randomness has also been developed is the resolution of inventory routing problems where the product usage is stochastic (Hemmelmayr et al., 2010). Bianchi et al (2009) have written an interesting survey of the appropriate meta-heuristics to solve a wide class of combinatorial optimization problems under uncertainty. The aforementioned survey is a good reference for obtaining an appropriate list of articles regarding the use of meta-heuristics for solving stochastic combinatorial optimization problems (SCOPs) in different application fields.

In this paper we discuss how simulation can be combined with heuristics and meta-heuristics in order to efficiently solve two family of SCOPs, the Vehicle Routing Problems (VRPs) and the Arc Routing Problems (ARPs). The paper is structured as follows: Section 2 offers an overview of both VRPs and ARPs; Section 3 presents a savings heuristics which can employ a biased randomization component for solving both families of problems; Section 4 shows how this heuristic can be applied in deterministic

problem; Section 5 introduces the concept of Simheuristics; Section 6 shows examples on which this methodology can be used; Section 7 presents a framework based on Simheuristics for solving non-deterministic problems; and finally, Section 8 summarizes the main conclusions of this work.

2 OVERVIEW OF VEHICLE AND ARC ROUTING PROBLEMS

VRPs is a family of well known problems which has long been tackled by researchers for several decades now, not only because its potential applications but also due to the fact that they can be used to test efficiency of new algorithms and optimization methods. VRPs are known to be NP-hard problems (Prins 2004). The most basic example of VRP is the Capacitated Vehicle Routing Problem (CVRP). In the CVRP (Juan et al. 2010) a set of customer demands have to be served with a fleet of vehicles from a depot. Each vehicle has the same limited capacity and each customer has a certain demand that must be satisfied and is known beforehand. Additionally, there is a cost matrix that measures the costs associated with moving the vehicle from one node to another. These costs usually represent distances or traveling times. Under these circumstances, the goal of the CVRP is to find the set of routes for serving all the customer's demands, considering that:

1. Each route starts and ends in the depot.
2. Each customer is supplied by a single vehicle.
3. A vehicle cannot stop twice at the same customer.
4. The demand served by a vehicle cannot surpass the vehicle capacity.

The objective function of this problem is defined by the sum of the costs between the nodes visited by each vehicle, taking all the vehicles into account.

ARPs is the counterpart to the VRPs where the customer demands are located in (some) of the edges in the graph instead of the nodes. In this case, the most basic example is the Capacitated Arc Routing Problem (CARP). Similarly to the CVRP, in the CARP (Golden and Wong 1981) a set of customer demands which are located in the edges of a graph need to be served by a homogeneous fleet of vehicles with limited capacity. The main difference of the CARP with respect to CVRP is that it is defined over a non-complete graph, which means that only some of the nodes have a direct connection. Also, every edge in the graph can be traversed as many times as required by any vehicle but it can only be served by a single vehicle.

3 BIASED RANDOMIZATION OF THE SAVINGS HEURISTIC

One of the most cited heuristic to solve the VRPs is the Clarke and Wright's Savings (CWS) constructive algorithm (Juan et al. 2011a). It is an iterative method that starts by considering an initial dummy solution in which each customer is served by a dedicated vehicle. Next, it starts an iterative process for merging some of the routes in the initial solution. Merging routes can improve the expensive initial solution so that a unique vehicle serves customers from the merged routes. The merging criterion is based on the concept of savings. Roughly speaking, given a pair of nodes to be served, a savings value can be assigned to the edge connecting these. This savings value is given by the reduction in the total cost function due to serving both customers with the same vehicle instead of using a dedicated one. This way, the algorithm constructs a list of savings, one for each possible edge connecting two customers.

The CWS heuristic can be modified to be used for the ARPs, by defining the Savings Heuristic for the Arc Routing Problem (SHARP). The adaption of the heuristic for the CARP is not trivial and is described in Gonzalez et al. 2012a. It starts by computing the shortest paths for all the pair of nodes in the graph, so it can define an equivalent complete graph including the costs of these paths. Having a complete graph allows computing the savings associated which each arc where the customers are located, in a similar way it is done for the CVRP. The edges are then sorted in a list according to their associated

savings, and the initial dummy solution is generated. From this, the iterative procedure is started which tries to merge routes in order to obtain a less costly solution.

This heuristic can be randomized, in both versions, to construct a powerful algorithm. The randomized algorithm benefits from the fact that the CWS has very low execution times, and also with the fact that, a different solution can be generated every time the algorithm is executed with a possibility of outperforming the best solution found so far with the use of pseudo-random number generators during the construction phase of the base heuristic. The randomization consists of using a skewed random number generator (e.g. geometric probability distribution) to guide the constructive process while keeping the heuristic criterion. Unlike a uniform randomization of the savings list, which will destroy the greedy behavior of the heuristic and therefore the output of the randomized algorithm will unlikely provide a good solution, a biased approach constitutes an efficient way to randomize the priority list without losing the greedy behavior of the heuristic. In the case of CWS (and SHARP), the constructive phase is based on a sorted list of candidates. These candidates have an associated savings value, which are sorted and processed starting with the best candidates. In the randomized version, instead of selecting always the best candidate available, all the candidates are considered, with items having a better value (according to the criterion) having better odds of being selected. By integrating this process inside a multi-start schema, many different solutions are obtained, some of them outperforming the solution obtained with the deterministic classical heuristic. With this, the algorithm is executed many times through the multi start procedure, and the best solution is selected as the solution obtained by the randomized algorithm. The described approach is able to considerably improve the results obtained by the classical heuristic. For example, in medium-sized instances in which the classical heuristic is able to obtain solutions with a cost within a gap of 20% to 40% with respect to the best know solution, the proposed approach is able to find solutions with a gap below 1%, being the best known solution in some instance.

3.1 Examples of applications in deterministic routing

We can find in the literature several examples of the application of the methodology described in the previous section. In Juan et al. 2010 we can find one example of the randomization of the CWS for the CVRP with the SR-GCWS algorithm. The algorithm introduces a skewed random behavior within the CWS heuristic in order to start an efficient search process inside the space of feasible solutions. Each of these feasible solution consists of a set of roundtrip routes from the depot that, altogether, satisfy all the demands of the nodes by visiting and serving all of them exactly once.

In Gonzalez et al. 2012a the RandSHARP algorithm is introduced for the CARP. This algorithm is a biased randomized version of the SHARP algorithm. The randomized algorithm uses solutions produced by the savings-based SHARP heuristic, and then iteratively it generates a new randomized solution by introducing a probabilistic criterion when selecting edges from the savings list. The algorithm uses a geometric probabilistic criterion. In Dominguez et al. 2014, the authors propose an algorithm to solve the 2L-HFVRP, which can be seen as a variant of the heterogeneous VRP where two-dimensional loading constraints have been incorporated to deal with 'large-size' items – which are usually, assumed to be rectangular shaped. The problem represents an important challenge since it combines a heterogeneous VRP with vehicle packing problems. The combination of these two classical problems is found in practical applications of some real-world transportation activities. Their approach relies on the biased randomization of routing and packing heuristics, which are integrated inside a multi-start framework.

4 SIMHEURISTICS

As shown in Figure 1, a Simheuristic approach is a particular case of simulation-optimization which combines a heuristic/meta-heuristic algorithm with simulation methodologies—e.g. Monte-Carlo, discrete-event, agent-based, etc.— in order to efficiently deal with the two components in a SCOP instance: the optimization nature of the problem and its stochastic nature. Some examples of Simheuristics applications to different fields can be found in the optimization-simulation literature. Thus, for instance,

Juan et al. (2011b) and Gonzalez et al. (2012b) combined Monte Carlo simulation with routing meta-heuristics in order to solve, respectively, the vehicle routing problem with stochastic demands and the arc routing problem with stochastic demands; Peruyero et al. (2011) combined Monte Carlo simulation with a scheduling meta-heuristic in order to solve the permutation flow-shop problem with stochastic processing times; and Caceres et al. (2012) combined Monte Carlo simulation with a routing meta-heuristic in order to solve the inventory routing problem with stock-outs and stochastic demands. Also, as illustrated later in this paper, discrete-event simulation can be used in combination with a meta-heuristic to solve other SCOPs where the stochastic behavior is conditioned by the time factor.

Typically, given a SCOP instance, a heuristic/meta-heuristic algorithm is run in order to perform an oriented search inside the solution space. This iterative search process aims at finding a feasible solution with the best possible value, which is expected to be near-optimal as well. During the iterative search process, the algorithm must deal with the stochastic nature of the SCOP instance. One natural way to do this is by taking advantage of the capabilities simulation methods offer to manage randomness. Of course, other approaches can also be used instead of simulation –e.g. dynamic programming, fuzzy logic, etc. However, under the presence of historical data on stochastic behavior, simulation allows the development of both accurate and flexible models. Specifically, randomness can be modeled throughout a best-fit probability distribution –including parameterization– without any additional assumptions or constraints. Thus, simulation is usually integrated with the heuristic/meta-heuristic approach and it frequently provides dynamic feedback to the searching process in order to improve the final outcome. In some sense, simulation allows to extend already existing and highly efficient meta-heuristics –initially designed to cope with deterministic problems– so that they can also be employed to solve SCOPs.

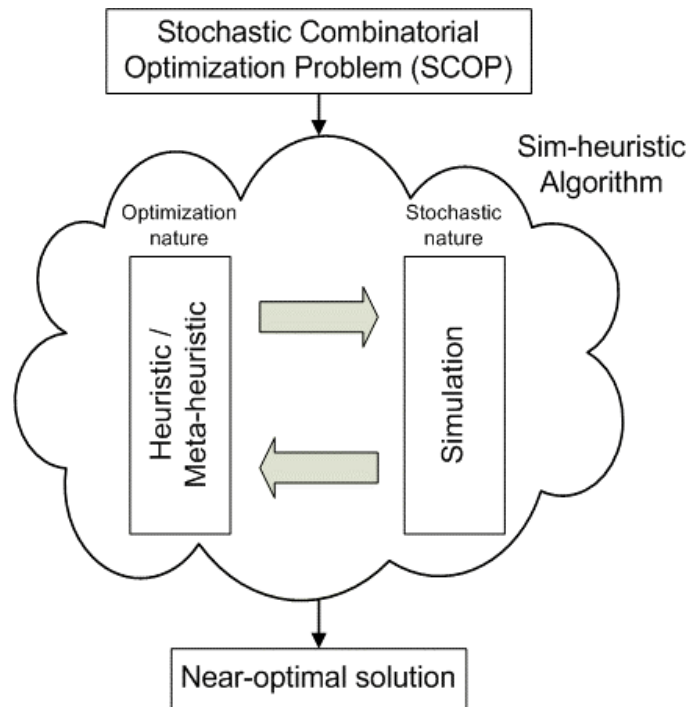


Figure 1: Overview scheme of a Simheuristics approach.

Obviously, one major drawback of this approach is that the results are not expected to be optimal anymore, since Simheuristics are combining two approximate methodologies. Nevertheless, real-life problems are complex enough and usually NP-hard even in their deterministic versions. Therefore, sim-

The random behavior of customers' demands could cause an expected feasible solution to become an unfeasible one if the final demand of any route exceeds the actual vehicle capacity. This situation is referred to as *route failure*, and when it happens some corrective actions must be introduced to obtain a new feasible solution. Thus, 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. As discussed in Juan et al. (2011b), one possible methodology to deal with this problem is to design reliable solutions, i.e., 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 vehicle capacity surplus (safety stock) while designing the routes, so that if final routes' demands exceed their expected values up to a certain limit, they can be satisfied without incurring in a route failure. Using safety stocks not only contributes to reduce variable costs due to route failures but, related to that, it also increases the reliability or robustness of the planned routes, i.e.: as safety stock levels increase, the probability of suffering a route failure diminishes. Notice, however, that employing safety stocks also increases fixed costs associated with aprioristic routing design, since more vehicles and more routes are needed when larger buffers are considered. Therefore, when minimizing the total expected cost a tradeoff exists between its two components, fixed costs and expected variable costs. Thus, the challenge relies on the selection of the appropriate buffer size.

Given a VRPSD instance, Juan et al. (2011b) consider different levels of this buffer size and then solve the resulting scenarios. This is performed by employing Monte Carlo simulation, which allows estimating the variable costs associated with each candidate solution. Thus, among the multiple solutions generated for each scenario, the ones with lowest total expected costs are stored as the best-found result associated with the corresponding safety-stocks level. Once the execution of the different scenarios ends, the corresponding solutions are compared to each other and the one with the lowest total expected costs is selected as the best-found routing plan.

In regards of ARPSD, Gonzalez et al. (2012b) proposed and implementation based on the same framework for the ARPSD problem. The authors also proposed to solve the problem by solving the equivalent deterministic problem where the customer demands were equal to the mean value and the vehicle capacity was restricted due to the consideration of a safety stock. Also, Monte Carlo simulation was used for simulate realizations of the problem instance and evaluate the total expected cost of the obtained solution. Additional details on this simulation-based framework follow in next section.

6 AN INTEGRATED SIMULATION-BASED FRAMEWORK

The solutions proposed for the VRPSD in Juan et al. (2011b) and to the ARPSD in Gonzalez et al. (2012b) were based on the same framework, which will be described in this section. This methodology is based in two facts: (a) the stochastic problem can be seen as a generalization of the deterministic problem where the random value has zero variance; and (b) efficient meta-heuristics already exist for the deterministic problems while the stochastic problems are an emerging field. Accordingly, the key idea behind this framework is to transform the stochastic instance into a new problem which consists of solving several *conservative* instances of the deterministic problem, each of the characterized by a specific risk of showing routing failures. The term conservative refers to the fact that only a certain percentage of the vehicle capacity is considered during the route design phase, so the rest of capacity is considered as *safety stock*. So, this remaining capacity of the vehicle will be available in case the actual demand of the route is greater than expected.

The methodology consists on the following steps (see Figure 3):

1. Consider a problem instance with a set of customers. Each customer has associated a stochastic demand, D_i , characterized by its mean and probability distribution.

2. Select a value for the risk level percentage (k) and compute the value of the vehicle usable capacity (VMC^*) given by a percentage of the vehicle capacity (VMC). This percentage will be a parameter defined for the algorithm.
3. Consider the deterministic instance of the problem, consisting on the same problem instance than the stochastic version, but where the demands are deterministic and their value equal to the mean of the stochastic demand ($d_i^* = E[D_i]$), and the vehicle capacity is restricted to the value computed on step 2 (VMC^*).

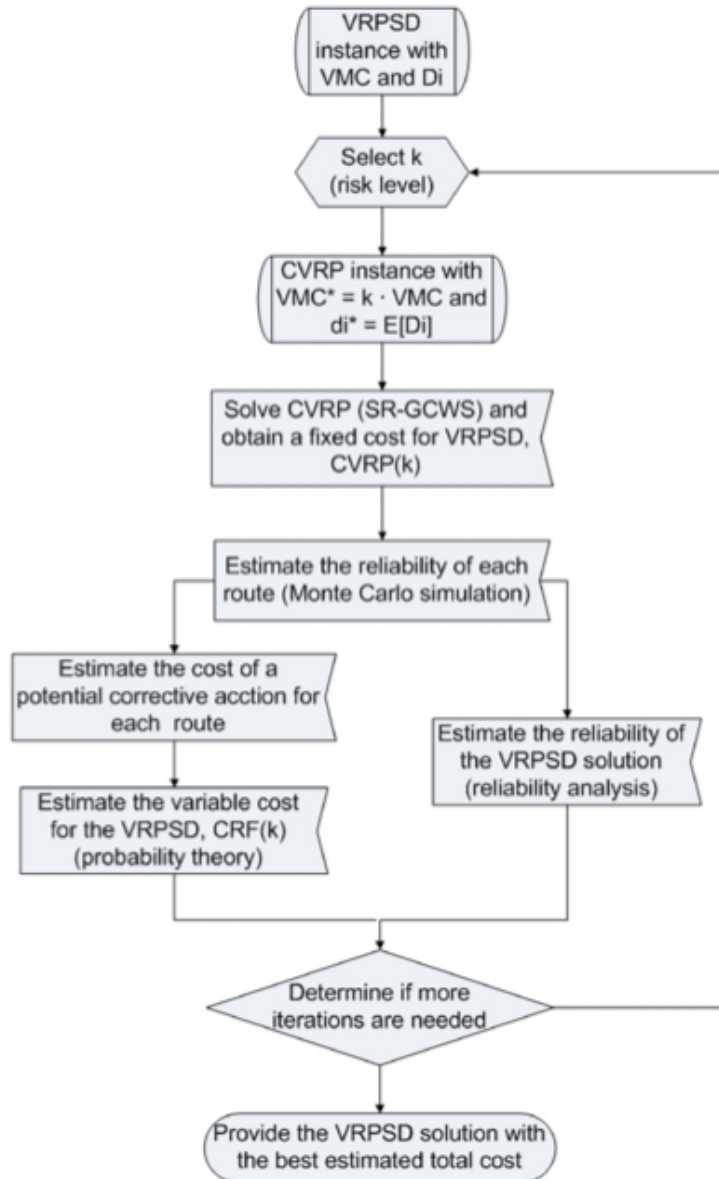


Figure 3: Flow diagram for the proposed methodology applied for the VRPSD.

4. Solve previous instance using a biased randomized algorithm: e.g. the Simulation in Routing Generalized Clarke and Wright Savings (SG-GCWS) for CVRP or Randomized Savings Heuristic for Arc Routing Problem (RandSHARP) for CARP. This solution will be an aprioristic solution for the original problem instance. Furthermore, it will be a feasible solution as long as there are not any route failures.
5. Using the previous solution, estimate the expected cost due to possible failures on any route ($CRF(k)$). This is done by using Monte Carlo simulation. To this end, random demands are generated and, whenever a route occurs a repair action is applied, registering the associated cost of this action in the total cost of the solution. The repair action consists of a round-trip from the depot to the failing customer so the vehicle capacity is reloaded. After iterating this process some thousands of times, a random sample of costs is obtained, and an estimate for its expected value can be calculated. Then, the expected total costs due to possible route failures can be computed by addition of these variable costs and the costs of the deterministic solution obtained during the design phase.
6. Using the deterministic solution, obtain an estimate for the reliability of each route of the solution. In such context, the reliability index is defined as the probability that a route will not fail. This route reliability index is computed by direct Monte Carlo simulation using the probability distributions that model customer demands in each route. Remark that in each route, over-estimated demands could sometimes be compensated by under-estimated demands.
7. The reliability index for the total solution is computed as the multiplication of the value for each route in the solution. This value can be considered as a measure for the feasibility for the solution in the context of the stochastic problem.
8. Depending on the total costs and the reliability indices associated with the solutions already obtained, repeat the process from Step 1 with a different value of the value used for the safety stock.
9. Finally, the best solution found so far is returned as solution to the problem, as well as its corresponding properties such cost or reliability index.

7 CONCLUSIONS

In this paper we have discussed how biased randomization and Simheuristics can be used for solving problems in the family of Vehicle Routing Problems and Arc Routing Problems. We have shown how, for deterministic problems, classical heuristics like the CWS can be combined with biased randomization to obtain state of the art solutions for the problems. Also, for stochastic problems (SCOPs), we have shown how Simheuristics can be used for solving the problem. The basic idea of Simheuristics is to combine the biased randomized algorithm for the deterministic version of the problem with Monte Carlo simulation in order make the algorithm be able to react to greater values than expected during the phase of resolution of the deterministic problem. As the described methodology is quite flexible, it can be easily adapted for other families of problems in the field of routing or scheduling.

ACKNOWLEDGMENTS

This work has been partially supported by the Spanish Ministry of Economy and Competitiveness (TRA2013-48180-C3-3-P).

REFERENCES

- Bianchi, L., M. Dorigo, L.M. Gambardella, and W.J. Gutjahr. 2009. "A Survey on Metaheuristics for Stochastic Combinatorial Optimization". *Natural Computing: An International Journal* 8(2): 239-287.

- Caceres, J., A., Juan, S. Grasman, T. Bektas, J. Faulin. 2012. "Combining Monte Carlo Simulation with Heuristics for Solving the Inventory Routing Problem with Stochastic Demands". *Proceedings of the 2012 Winter Simulation Conference*, 1-9. Berlin, Germany.
- Dominguez, O, Juan, A., Barrios, B., Faulin, J., and Agustin, A. 2014. "Using Biased Randomization for Solving the Two-Dimensional Loading Vehicle Routing Problem with Heterogeneous Fleet". *Annals of Operations Research*.
- Fleury, G., Lacomme, P., Prins, C., and Ramdane-Chérif, W. 2005. "Improving Robustness of Solutions to Arc Routing Problems". *Journal of the Operational Research Society*, 56(5): 526-538.
- Golden, B.L., and Wong, R.T. 1981. "Capacitated arc routing problems". *Networks*, 11(3): 305-315.
- Gonzalez, S., Juan, A., Riera, D., Castella, Q., Munoz, R., and Perez, A. 2012a: "Development and Assessment of the SHARP and RandSHARP Algorithms for the Arc Routing Problem". *AI Communications*, 25: 173-189
- Gonzalez, S., A. Juan, D. Riera, M. Elizondo, P. Fonseca, P. 2012b. "Sim-RandSHARP: A Hybrid Algorithm for solving the Arc Routing Problem with Stochastic Demands". *Proceedings of the 2012 Winter Simulation Conference*, 1-11. Berlin, Germany.
- Hemmelmayr, V., K.F. Doerner, R.F. Hartl, and M.W.P. Savelsbergh. 2010. "Vendor Managed Inventory for Environments with Stochastic Product Usage". *European Journal of Operational Research* 202: 686-695.
- Juan, A., Faulin, J., Ruiz, R., Barrios, B., and Caballe, S. 2010: "The SR-GCWS hybrid algorithm for solving the capacitated vehicle routing problem". *Applied Soft Computing*, Vol. 10, No. 1, pp. 215-224
- Juan, A., Faulin, J., Jorba, J., Riera, D., Masip; D. and Barrios, B. 2011a: "On the Use of Monte Carlo Simulation, Cache and Splitting Techniques to Improve the Clarke and Wright Savings Heuristics". *Journal of the Operational Research Society*, 62(6): 1085-1097.
- Juan, A., Faulin, J., Grasman, S., Riera, D., Marull, J. and Mendez., C. 2011b. "Using Safety Stocks and Simulation to Solve the Vehicle Routing Problem with Stochastic Demands". *Transportation Research Part C*, 19: 751-765.
- Novoa, C., and R. Storer. 2009. "An Approximate Dynamic Programming Approach for the Vehicle Routing Problem with Stochastic Demands." *European Journal of Operational Research*, 196: 509-515.
- Peruyero, E., A. Juan, and D. Riera. 2011. "A Hybrid Algorithm Combining Heuristics with Monte Carlo Simulation for Solving the Stochastic Flow Shop Problem". *Proceedings of the 2011 ALIO/EURO Workshop*, 127-130. Porto, Portugal.
- Prins, C. 2004. "A Simple and Effective Evolutionary Algorithm for the Vehicle Routing Problem". *Computers and Operations Research*, Vol. 31, pp. 1985-2002.
- Van Hentenryck, P., and R. Bent. 2010. *Online Stochastic Combinatorial Optimization*. The MIT Press. Boston. USA.

AUTHOR BIOGRAPHIES

SERGIO GONZALEZ-MARTIN is a Ph.D. candidate of Applied Optimization and Simulation in the Computer Science Department at the IN3-Open University of Catalonia (Barcelona, Spain). Mr. González holds a M.S. in Telecommunication Engineering and a M.S. in Software Libre. His research interests include Applied Optimization and Randomized Algorithms with focus on its Telecommunication application. His e-mail is sgonzalezmarti@uoc.edu.

BARRY B. BARRIOS obtained his B.S. degree in Physics from Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts. His research interests are in the area of applied optimization and discrete-event simulation. He has recently been working in healthcare problems, applying industrial and

systems engineering techniques to the improvement of health systems. Previously, he has worked at the Open University of Catalonia (UOC) on logistics and transportation problems. His research interests are in the area of rich and real-life vehicle routing problems, supply chains and stochastic discrete-event simulation. His email address is barry.brian.barrios@gmail.com.

ANGEL A. JUAN is an Associate Professor of Optimization & Simulation in the Computer Science Department at the Open University of Catalonia (UOC). Dr. Juan holds a Ph.D. in Applied Computational Mathematics. He completed a predoctoral internship at Harvard University and a postdoctoral internship at the MIT Center for Transportation & Logistics. He has been invited researcher at the University of Southampton (UK), at LAAS-CNRS (France), at the University of Natural Resources and Life Sciences (Austria), and at the University of Portsmouth (UK). His research interests include applications of Randomized Algorithms and Simheuristics in Logistics, Production, and Internet Computing. He has published over 140 peer-reviewed papers regarding these fields. His website address is <http://ajuanp.wordpress.com>.

DANIEL RIERA is Lecturer in the Department of Computer Science, Multimedia and Telecommunication (EIMT). He is academic director of Master in Bioinformatics (UOC), and also of Computer Science Engineering (EIMT). Dr. Daniel holds a Ph.D. in Computer Science (UAB), a M.S. in Advanced Techniques of Processes Automation (UAB), and a M.S. in Computer Science (UAB). His main research scopes include the model of discreet systems using Petri nets, and the optimization using Constraint Programming techniques. He has been researcher of the LOGISIM, Centre of simulation and optimization of logistic systems, from the Network of Centers of Support to Technological Innovation (XIT) of the CIDEM. Currently, he belongs to the research group in Software Engineering of GRES-UOC where investigates the verification of UML+OCL models by means of Constraint Programming. His email address is driera@uocl.edu and his web page is <http://www.uoc.edu/webs/drierat>.