Defence Research and Development Canada | Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

A Hybrid Genetic Algorithm for the Vehicle Routing Problem with Time Windows

J. Berger
M. Salois
DRDC Valcartier

Defence R&D Canada – Valcartier
Technical Report
DRDC Valcartier TR 2002-162
June 2004

Canada

A Hybrid Genetic Algorithm for the Vehicle Routing Problem with Time Windows

J. Berger
DRDC Valcartier

M. Salois
DRDC Valcartier

Defence Research and Development Canada - Valcartier

Technical Report

DRDC Valcartier TR 2002-162

2004-06-03

Author

Jean Berger and Martin Salois

Approved by

Éloi Bossé

Head/DSS Section

Abstract

A variety of hybrid genetic algorithms has been recently proposed to address the vehicle routing problem with time windows (VRPTW), a problem known to be NP-hard. However, very few genetic-based approaches exploit implicit knowledge provided by the structure of the intermediate solutions computed during the evolutionary process to explore the solution space. This report presents a new hybrid genetic algorithm for the VRPTW. It investigates the impact of using explicitly domain knowledge and prior knowledge/characteristics about typical solutions expected from the recombination and mutation phases of the algorithm. Basic principles borrow from recent hybrid and standard genetic algorithms, and features of well-known heuristics to drive the search process. Designed to support time-constrained reasoning tasks, the procedure is intended to be conceptually simple, easy to implement, and allow fast computation of near-optimal solution. A computational experiment has been conducted to compare the performance of the proposed algorithm with similar and standard techniques.

Résumé

Une variété d'algorithmes génétiques a récemment été proposée pour résoudre le problème de tournées de véhicules avec fenêtres de temps (PTVFT), un problème NP-difficile. Cependant, peu d'approches génétiques exploitent la connaissance implicite liée à la structure des solutions intermédiaires émergeant au cours du processus évolutif, dans le but d'explorer l'espace de recherche. Ce rapport présente un nouvel algorithme génétique hybride pour le PTVFT. Il explore l'impact de l'utilisation explicite de connaissances du domaine et des connaissances/caractéristiques a priori de solutions typiques anticipées, résultant des phases de recombinaison et de mutation de la procédure. Les principes de base sont inspirés d'algorithmes génétiques récents et de méthodes heuristiques bien connues afin de guider la recherche. Conçue pour appuyer le raisonnement comportant des tâches avec contraintes temporelles, la procédure est simple, facile à implanter tout en permettant le calcul rapide de solution quasi optimale. Une expérience de simulation a été réalisée afin de comparer la performance de l'approche proposée à des méthodes bien connues.

This page intentionally left blank.

Executive summary

The need to provide automated capabilities to support mission planning is a primary and very challenging issue for the Canadian Air Command. This is motivated by mission demand, nature and diversity, and the heterogeneity and number of resources available to support planning and cost-effective resource allocation at different levels. An important aspect of mission planning involves the vehicle routing problem with time windows (VRPTW). The combinatorial complexity of this problem calls for further development and refinement of heuristic methods. A promising class of procedures to tackle this problem relies on the use of genetic algorithms. Even though a variety of genetic algorithms have been recently proposed to address VRPTW, very few genetic-based approaches exploit implicit knowledge provided by the structure of the intermediate solutions computed during the evolutionary process to further explore the solution space.

This document presents a new hybrid genetic algorithm for VRPTW. It investigates the impact of using explicitly domain knowledge and prior knowledge/characteristics about typical solutions expected from the recombination and mutation phases of the algorithm. Basic principles are borrowed from recent hybrid and standard genetic algorithms, and features of well-known heuristics to achieve directed-search. A computational experiment has been conducted to compare the performance of the proposed algorithm with similar and standard techniques. Simple and easy to implement, the stochastic procedure allows for the fast computation of a near-optimal solution, a suitable property to address time-constrained reasoning constraints. Simulation results show that even though the proposed hybrid genetic algorithm does not outperform some recent tabu search techniques yet, computed solutions remain very competitive and nearly match some of the best known solutions.

This work represents a significant contribution to tackle key resource management problems faced by the Canadian Air Force. Perfectly compatible with the objectives of the Air Force AFCCIS project to support resource planning, the proposed algorithm can be easily integrated into any advisory decision support system as a separate component or module. Targeted application domains include Army Aviation Fleet, Fighter Fleet, Air Transport Fleet and Maritime Air Fleet activities, at Wing and Squadron levels. These comprise transport, surveillance, identification, interception, intervention or destruction, contingency operations, tactical airlift, emergency management such as human relief and disaster operations, and other air mission management activities.

Berger, J., 2002. A Hybrid Genetic Algorithm for the Vehicle Routing Problem with Time Windows, TR 2002-162 Defence R&D Canada Valcartier.

Sommaire

Le besoin de fournir des capacités automatisées pour faciliter la planification de mission est primordial et constitue un défi majeur pour le commandement aérien canadien. Ceci est motivé par la demande, la nature et la diversité des missions, ainsi que par l'hétérogénéité et le nombre de ressources disponibles pour appuyer la planification et l'affectation efficace des ressources à différents niveaux. Un aspect important lié à la planification de mission concerne le problème de tournées de véhicules avec fenêtres de temps (PTVFT). La complexité combinatoire de ce problème exige le développement et le raffinement accrus de méthodes heuristiques. À cet effet, les algorithmes génétiques constituent une classe de procédures prometteuses pour résoudre ce problème. Malgré le nombre d'algorithmes génétiques récemment proposés, très peu d'approches exploitent la connaissance implicite liée à la structure des solutions intermédiaires émergeant du processus évolutif afin de mieux explorer l'espace de recherche.

Ce document présente un algorithme génétique hybride pour le problème de tournées de véhicules avec fenêtres de temps. Il explore l'impact de l'utilisation explicite de connaissances du domaine et des connaissances/caractéristiques a priori de solutions typiques anticipées, résultant des phases de recombinaison et de mutation de la procédure. Les principes de base sont inspirés d'algorithmes génétiques récents et de méthodes heuristiques bien connues afin de guider la recherche. Une expérience de simulation a été réalisée afin de comparer la performance de l'approche proposée à des méthodes bien connues. Simple et facile à implanter la procédure stochastique permet le calcul rapide de solution quasi optimale, une propriété fort souhaitable dans le raisonnement de tâches comportant des contraintes temporelles. Les résultats montrent que même si l'algorithme proposé ne surpasse pas encore les techniques de recherche "tabous" récentes, les solutions obtenues demeurent très compétitives et avoisinent certaines des meilleures solutions connues.

Cet effort représente une contribution significative afin d'aborder les problèmes d'affectation de ressources auxquels doivent faire face les Forces aériennes canadiennes. Alignée avec les objectifs du projet AFCCIS des Forces aériennes pour appuyer la planification de ressources, l'approche proposée peut être implantée dans un module indépendant facilitant son intégration à n'importe quel système d'aide à la décision. Les domaines d'application ciblés incluent la flotte d'avions de l'armée, la flotte des chasseurs, la flotte aérienne de transport et la flotte maritime, au niveau escadron et escadre. Ceux-ci comprennent des opérations de transport, surveillance, identification, interception, intervention ou destruction, contingence, tactique, urgence, désastre et autres activités de gestion de missions aériennes.

Berger, J., 2002. A Hybrid Genetic Algorithm for the Vehicle Routing Problem with Time Windows, TR 2002-162 Defence R&D Canada Valcartier.

Table of contents

List of figures

List of tables

1. Introduction

In the vehicle routing problem with time windows [1,2], a set of customers with specific demands is serviced by a homogeneous fleet of vehicles of limited capacity, initially located at a central depot. Routes are assumed to start and end at the central depot. Each customer provides a time period in which they require a service, consisting in a particular task to be completed such as repair work or loading/unloading the vehicle. Hard time window constraints impose customer service to be initiated within a predetermined specific time interval. It is worth noticing that the time window requirement does not prevent any vehicle from waiting at a customer location before servicing, even though opportunistic waiting implicitly incurs a penalty in route cost. The primary objective is to minimize the number of tours or routes, then the total route (travel and waiting) time such that each customer is serviced in its time window and, the total load on any vehicle associated with a given route does not exceed vehicle capacity.

A variety of algorithms including exact methods and efficient heuristics have already been proposed for VRPTW [1,2]. Recent performance shown by some metaheuristics and evolutionary approaches, such as genetic algorithms [3,4] and tabu search [5,6] relying on stochastic-based search techniques to explore the solution space, make these techniques extremely promising [7-11]. The problem-solving methodology of genetic-based algorithms has definitely been one of the most suitable approaches to tackle the VRPTW [7-10] so far. The algorithm proposed by Blanton and Wainwright [10] consists in encoding individuals by representing a chromosome as a sequence of customers. Search is driven toward suitable orderings of customers based upon precedence relationship (temporal, spatial, mixed) as well as a fixed global precedence order defined over customer time window lower bounds. Genetic operators are proposed to generate new customer sequences or orderings from previous ones. The decoding procedure relies on a greedy heuristics reminiscent of the bin-packing algorithm, in which ordered customers from the chromosome are successively inserted in the solution. On the other hand, the approach proposed by Thangiah [8,9] relies on an alternate two-phase process, namely customer clustering and cluster-based customer routing. Customer clustering uses a genetic-based method in which individuals encoding (chromosome) are represented by a sequence of neighboring sectors originating from the central depot and describing respectively clusters of customers. Local cluster-based customer routing is then achieved using a particular heuristic. As for the algorithm proposed by Potvin [7], it mainly relies on the basic principles of genetic algorithms disregarding explicit solution encoding issues for problem representation. Genetic operators are simply applied to a population of solutions as opposed to a population of encoded solutions (chromosomes). The search is mainly driven by heuristic information about the problem domain. The evolution of the population is achieved considering the best features shown by parent solutions.

The classical genetic algorithm paradigm primarily operates at the encoding level (chromosomes) and therefore do not exploit explicitly any useful information about a particular problem domain. In counterpart, hybrid genetic algorithms [4] attempt to make use of particular knowledge of the problem domain to drive the search process, as illustrated by some of the recent methods described earlier for VRPTW. However, we conjecture that most of the genetic-based algorithms proposed for this problem do not capture enough explicit domain knowledge and prior knowledge/characteristics about expected solutions that tend to possibly improve the search process. The contribution of such knowledge to the selection of favorable domain-dependent strategies might contribute to explore more efficiently the solution space, while speeding-up the generation of near-optimal solutions, a desirable property to achieve time-constrained reasoning.

In this document a new hybrid genetic algorithm for VRPTW is proposed. It investigates the impact of using explicitly domain knowledge and prior knowledge/characteristics about expected solutions, during the recombination and mutation phases of the algorithm. Basic principles are borrowed from recent hybrid and standard genetic algorithms and features of well-known heuristics to drive the search process. Designed to support time-constrained reasoning tasks, the procedure is intended to be conceptually simple, easy to implement, and allow fast computation of near-optimal solution. Despite the fact that the algorithms are neither admissible nor complete, we contend that the proposed technique is well suited to support time-constrained reasoning tasks, involving limited computational resources.

The report is outlined as follows. Section 2 introduces the basic concepts of genetic algorithms. It describes the general principles of the technique and then presents preliminary observations about an early investigation of the method for VRPTW. An overview of the proposed hybrid genetic algorithm is then given. The main features of the procedure are briefly presented with a special emphasis on the recombination and mutation operators. Then, in Section 3, we present the results of a computational experiment to assess the value of the proposed approach. The performance of the method is compared with similar and standard techniques, and the strengths and weaknesses of the algorithm are briefly discussed. Finally, we conclude with a short summary in Section 4.

2. Genetic Algorithm Approach

2.1 General Principles

A genetic algorithm [3,4] is an evolutionary computation technique inspired from the principles of natural selection to search a solution space. It evolves a population of individuals encoded as chromosomes by creating new generations of offspring through an iterative process until some convergence criteria or conditions are met. Such criteria might, for instance, refer to a maximum number of generations or the convergence to a homogeneous population composed of similar individuals. The best chromosome generated is then decoded, providing the corresponding solution. The underlying reproduction process is mainly aimed at improving the fitness of individuals; a measure of profit, utility, or goodness to be maximized while exploring the solution space.

The creation of a new generation of individuals involves primarily four major steps or phases: representation, selection, recombination (crossover), and mutation. The representation of the solution space consists in encoding salient features of a solution as a chromosome (chromosome encoding), defining an individual member of a population. Typically pictured by a bit string, a chromosome is made up of a sequence of genes, which capture the basic characteristics of a solution. The evolution of the encoded solutions or population members is driven by the selection, recombination and mutation phases, respectively. The selection phase consists in choosing randomly two parent individuals from the population for mating purposes. The probability of selecting a population member (parent) is generally proportional to its fitness in order to emphasize genetic quality while maintaining genetic diversity. Biased toward the best chromosomes, selection is aimed at propagating good solution features from one generation to the next. Recombination propagates genes of selected parents to produce offspring that will form the next generation. It combines characteristics of chromosomes (parent solutions) to potentially create offspring with a better fitness (intensification). Mutation consists in randomly modifying gene(s) of a single individual at a time to further diversify search of the solution space therefore ensuring genetic diversity. The occurrence of mutation is generally associated with a low probability. A new generation is created by executing repeatedly the selection, recombination and mutation processes until all chromosomes are replaced. A proper balance between genetic quality and diversity is therefore required within the population in order to support efficient search.

2.2 Preliminary Observations

A preliminary investigation of some variants of the known genetic-based techniques has been conducted for VRPTW in order to identify their strengths and weaknesses. An extension of the algorithm proposed in [10] to further exploit data distribution of the problem instance was first explored. It consisted in generating various customer orderings, considering several a priori precedence relationships based upon time windows' lower and upper bounds, as well as statistical distribution over the distance separating two customers for a problem instance. The generalized precedence-ordering scheme was randomly used in concert with specialized recombination operators. A particular ordering selection was based on a probability distribution reflecting the data distribution of the problem instance. As a result, the extended procedure outperformed the original algorithm, the introduction of additional domain and problem structure knowledge consistently improving the quality of the computed solution. The next observation relates to the importance of encoding schemes, as illustrated by the quality of results obtained using different approaches [7-10]. Comparative results show that encoding strategies can play a key role in the quality of the computed solution, and/or its convergence rate. However, given that suitable and proper encoding for VRPTW may be extremely complex to achieve, it appears that traditional explicit encoding might not necessarily represent an absolute requirement to design very efficient genetic-based techniques as shown in [7]. Similarly to the work reported in [7], the approach proposed in this document overlooks explicit encoding issues.

The last observation refers to Potvin's work [7] in which some genetic operators can be very costly, and therefore can significantly impair the performance of the algorithm. In fact, the attempt to build a feasible route by connecting two route segments from two parent solutions while successfully inserting unrouted customers during the recombination phase can be very time-consuming. The failure to insert unrouted customers being very likely, alternate parent solutions need to be repeatedly examined until a feasible child solution can be generated. Local search performed in the mutation phase may be quite prohibitive as well. In the latter case, an iterative procedure repetitively attempts to move around several sequences of customers with various cardinality, and resorts to a potentially expensive repair procedure to perform local optimization. Consequently, special attention must be paid to the cost of search and methods used, such as repair procedures and, in a lesser extent, local optimization techniques to find feasible solutions if a trade-off between solution quality and run-time computation is a primary concern for a targeted application.

2.3 Hybrid Genetic Algorithm

In this Section, a hybrid genetic algorithm to solve VRPTW is presented. It consists in combining a basic genetic algorithm, domain knowledge and local search heuristics to further improve solution quality. Designed to support time-constrained reasoning tasks related to VRPTW, the proposed procedure is intended to be conceptually simple, easy to implement and allow fast computation of a near-optimal solution.

The algorithm uses explicitly domain knowledge and prior knowledge/characteristics about expected solutions during the recombination and mutation phases. It borrows from recent hybrid and standard genetic algorithms, and also includes features of well-known heuristics to drive the search process. Typical prior knowledge about an expected solution involving route-based and customer-based characteristics, include:

- waiting time and total route distance are minimum
- average travel time and variance amongst route customers tend to be minimum
- nearest customers (aggregate) tend to be part of the same route
- customers with a large number of close neighbors can be inserted in multiple routes
- customers with large/narrow service time window are easier/harder to schedule
- early/late route customers are scheduled earlier/later over their service time window
- a lower bound on the number of routes necessary (based on vehicle capacity and customer demand) can be computed
- small routes (small number of customers) tend to be merged within larger routes
- remote route customers tend to be inserted in nearest routes (route centroids)
- average and variance (distribution) over service time, time window parameters (lower, upper bounds, and length) and, travel time (customer distance)
- average time window/average travel time (between two customers) ratio

This knowledge is used to derive/design new strategy knowledge, in order to select routes or customers of parent solutions to be recombined or modified, namely:

- select routes of parent solution with the largest waiting time
- select a route neighborhood from the second parent for mating considerations
- select route customers (to be moved) of parent solution with:
    - the largest waiting time
    - a large time window
    - a large slack time (deadline - effective service time)
    - the largest travel time to its immediate (route customer) neighbors

The proposed approach consists in defining a suite/family/sequence of route-based genetic operators based on this knowledge, as well as features of well-known heuristics to drive the search process. Restricted to customer nodes specified through the routes of the mating parent solutions, a heuristic is locally applied during the recombination and mutation phases, to incrementally build feasible routes at low cost. The use of a heuristic is coupled to a random customer removal procedure, reminiscent of the principles of the simulated annealing technique [12], to further explore the search space and escape local minima. Some of the heuristics being investigated include features of insertion-based techniques [13] and the nearest-neighbor procedure.

Representation. A solution is represented by a set of feasible routes, an individual being implicitly encoded as a chromosome formed of multiple segments. A chromosome segment (sequence of genes) represents a feasible route, referring to a sequence of customers to be visited by a vehicle. A segment is delimited by two separators to specify the related route. Admissible genes are therefore defined by indexed customers and separator symbols.

$$/C_3C_4C_2C_{11}C_5/\ C_1C_9C_8C_{10}/\ C_{12}C_6C_7/\ldots////$$

/---- Route 1 ----/-- Route 2 ---/- Route 3 -/...////

Figure 1: Chromosome Encoding

A chromosome encoding is shown in Figure 1. It illustrates a string of symbols involving indexed customers (C_i) as well as separator (/) occurrences. A maximum number of separator occurrences can easily be determined (upper bound on the maximum number of routes, e.g. the number of vehicles available) to ensure a fixed chromosome length for each individual (solution).

Selection. The selection process consists in choosing two individuals (parent solutions) within the population for mating purposes. The selection procedure is stochastic and biased toward the best solutions using a roulette-wheel scheme [3]. In this scheme, the probability to select an individual is proportional to its fitness. Delayed to the completion of a new generation, an individual fitness is computed as follows:

$$fitness_i = r_i - r_m + \gamma d_i \tag{1}$$

where

r_i: number of routes in solution i

r_m: number of routes of the best solution in the current population

d_i: total traveled distance associated with solution i

γ: relative user-defined weight parameter controlling the distance contribution. The value of γ is selected to privilege solutions having a smaller number of routes (e.g. $\gamma = 1/d_m$, where d_m refers to the maximum traveled distance over the individuals forming the initial population),

The proposed fitness expression indicates that better solutions include fewer routes, and then involve smaller total traveled distance.

Recombination. The insertion-based (IB_X) crossover operator creates an offspring by combining iteratively various routes r_1 of a parent solution P_1 with a subset of customers, formed by r_1 nearest-neighbor routes from parent solution P_2. The neighboring routes of r_1 are determined by routes of P_2 whose centroids are located within a certain range of r_1's centroid. This range is primarily determined by the average distance separating r_1 from its neighbor P_2 routes. A distance measure between two routes is defined by the Euclidean distance or the travel time separating their respective centroid. A route centroid corresponds to a virtual site or customer whose coordinates refer to the average position of its specific routed customers. Using criteria based on prior knowledge over domain and expected solution, a removal procedure is first applied to remove from r_1 key customer nodes susceptible (mostly subject to prior knowledge) to be migrated within alternate routes. The stochastic customer removal procedure involves three kinds of knowledge from which strategies are defined. The related approaches consist in removing either randomly specific customers, rather distant customers from their successors (or alternately from route centroid) based on average route successors distance, or customers with waiting times (or alternately large time windows) above route customer average. Strategies are chosen randomly by the removal procedure based on a user-defined prior probability distribution. The technique moves candidate customer nodes most likely to be relocated, and therefore creates better opportunities for alternate key node insertions. The removed customer nodes will further be visited for insertion in future route constructions of the child solution. Then an insertion-based routing heuristic inspired from I1 [13] is locally applied to incrementally build a feasible route at low cost, considering the modified partial route r_1 as the initial solution. The standard heuristic I1 is coupled to a random customer acceptance procedure, in order to insert one of the best candidate customers on the route. For each new customer insertion to the partial route, the random acceptance method successively examines the sequence of best candidate nodes determined during the current iteration of the standard heuristic, until one is inserted. The method includes principles somewhat reminiscent of the simulated annealing technique to further explore search space, and escape local minima. Once a route construction is completed, the overall process is repeated for a random number of different routes r_1 of P_1. The child then inherits the remaining altered routes of P_1, if necessary, and any unrouted customers simply form additional one-customer routes. The process can be reiterated to generate a second child by interchanging P_1 and P_2.

The proposed insertion-based (IB_X) crossover operator can be summarized as follows:

Step 1 Select a random number of routes of P_1 to be visited.
Step 2 Repeat
- 2.1 Select randomly a new route r_1 (chromosome segment) from P_1 (chromosome). Probability for route selection is proportional to its relative total waiting time.
- 2.2 Select a subset RS2 of routes from P_2 located in the neighborhood of r_1. Neighboring routes of r_1 are determined by routes of P_2 whose centroids are located within a certain range of r_1's centroid. This range primarily determines the number of routes of P_2 to be considered.
- 2.3 Remove some customers from r_1 using a specific customer removal procedure.
- 2.4 Build a route child solution: apply an insertion-based routing heuristic considering the modified route r_1 as the initial partial solution and the pool of customers formed by the routes of RS2 and possibly already visited (but unrouted) customers.
  Insertion of unrouted but already visited customers will be examined again when building the next route.
- 2.5 Inserted nodes are then removed from the remaining subset of routes of P_1 to be visited.

Until (the selected random number of routes of P_1 has been visited).

Step 3 Inherits the remaining unvisited routes of P_1 (if any) while eliminating customers already routed.
Step 4 For each remaining unrouted customer (if any), build a new one-customer route.

The insertion-based routing heuristic used in Step 2.4 mainly exploits some features of the method I1 proposed in [13]. Limited to the construction of a single route for the child solution, the procedure considers the modified route r_1 (Step 2.3) as the initial partial solution and restricts the set of customers to be explored to RS2 (Step 2.2) as well as already visited (but unrouted) customers. The procedure inserts one customer at a time. Accordingly, it determines for each candidate customer left, the best possible location to be inserted, and then the best one to be added next to the partial route. The best possible location for a candidate customer is evaluated using domain knowledge criteria characterized by various adjustable parameters [13]. Parameters of the routing heuristic are selected randomly for each crossover execution. As for the insertion of the best candidate customer per se, an embedded random acceptance procedure is used. The probabilistic method successively examines the sequence of best candidate nodes determined during the current iteration until one is inserted. The probability for node insertion follows a periodic pattern, in which a schedule governing a linearly non-decreasing function determines a probability value over a given number of generations (period). If the best customer node to be inserted is rejected, the second best is then considered and, then the third, and so forth. A successful customer is eventually inserted. The procedure tends to create room for alternate customers, by rejecting best customer nodes to be inserted, and therefore escape local minima. The whole process (next customer insertion) is then repeated until no more customer can be inserted.

Mutation. The mutation operator is aimed at reducing the number of routes of solutions having only a few customers, and/or, evolving routes or escaping local minima by locally reordering customers. Three mutation operators are proposed namely, IB_M, NNR_M and DCR_M.

The insertion-based (IB_M) mutation operator attempts to move customers from the smaller route(s) (small number of customers) of a solution, to alternate existing routes, using the insertion-based routing technique described earlier. Alternate routes for customer insertion are randomly visited. The operator can be summarized as follows:

Step 1 Select the smallest route(s) of a solution.
Step 2 For each customer of the selected route do
  2.1 Repeat
  2.1.1 An alternate route is randomly examined for customer insertion.
  2.1.2 Try to relocate/insert the related customer using the insertion-based routing technique described above.
  Until (customer insertion OR all available routes have been unsuccessfully examined).

The nearest neighbor-based reordering (NNR_M) mutation operator attempts to reorder customers by applying locally the nearest-neighbor procedure [14] to each route of the targeted solution. The scheme is aimed at evolving the solution on a local basis, using an alternate routing heuristic biased toward a specific domain attribute; namely distance.

Step 1 For each route of a solution do
  1.1 Reorder locally customers based on the nearest-neighbor procedure. If reordering is not possible the route remains unchanged.

The distant customer reordering (DCR_M) mutation operator consists, for a specific route of the targeted solution, in trying to move remote customers regarded as significantly distant from their route successor (immediate neighbor), to an alternate route whose centroid is located within its immediate vicinity. A customer is assumed distant from its route successor, if the travel distance between the two is larger than the average travel distance separating route successors over the route. The operator can be summarized as follows:

Step 1 For each route of a solution do
  1.1 Compute average travel distance separating customers over the route.
  1.2 Remove customers regarded as significantly distant from their route successor, that is customers whose travel distance to their successor (immediate neighbor) is larger than the average distance computed in Step 1.1.

Step 2 For each removed customer in Step 1.2 do

2.1 Examine insertion in neighboring routes whose centroid is located within its immediate vicinity. The nearest routes (centroid) are visited first. Attempts for customer re-insertion to a different route make use of the insertion-based routing technique introduced earlier for the recombination phase.

Step 3 Build new routes from remaining unrouted customers (if any) emerging from Step 1.2, by aggregating as much customers as possible on the same route using a nearest-neighbor routing heuristic.

3. Computational Experiment

A computational experiment has been conducted to compare the performance of the proposed algorithm, with similar and standard techniques, and therefore assess the value of the proposed approach. The algorithm has been applied to a standard set of VRPTW instances [13]. Each instance involves 100 customers randomly distributed over a geographical area. The travel time separating two customers corresponds to their relative Euclidean distance. Customer locations for a problem instance are either generated randomly using a uniform distribution (problem sets R1 and R2), clustered (problem sets C1 and C2) or mixed, combining randomly distributed and clustered customers (problem sets RC1 and RC2). The standard set defines two classes of problem instances, namely, (R1, C1, RC1) and (R2, C2, RC2). The former includes problem instances characterized by a narrow scheduling horizon and small vehicle capacity. The latter involves problem instances characterized by a large scheduling horizon and large vehicle capacity. A narrow scheduling horizon generally requires a large number of routes (vehicles) to timely satisfy customer demands, whereas a large horizon involves fewer routes as a larger number of customers can be serviced by the same vehicle. More details about these problems and related data set may be found in [13].

The experiment has been conducted under specific conditions. Typical parameters and settings for the investigated algorithm include:

- Population size: 50 (2 populations)
- Maximum number of generations: 100
- Migration: 5
- Population replacement scheme: elitism
- Recombination rate: 60%
  - IB_X:
  - customer removal procedure (from route r_1 of parent solution P_1):
    - random: 25%
    - distance-based: 25%
    - largest waiting time: 50%
  - customer insertion acceptance procedure:
    - scheduling period T: 20 generations
    - insertion probability (over T generations): min $\{1/2 + i/T, 1\}$, i=1..T

- Mutation rate: 60%
  - if best fitness improves from one generation to the next then
    - IB_M
  - else apply with a 50% probability an alternate mutation operator:
    - NNR_M: 70%
    - DCR_M: 30%

The proposed hybrid genetic algorithm has been built on top of the GAlib [15] genetic algorithm library and implemented in C++. In order to emphasize genetic diversity, the simultaneous evolution of two populations has been considered. The migration parameter, a feature provided by GAlib, refers to the number of (best) chromosomes exchanged between populations after each generation. Population replacement is based upon an elitic scheme, meaning that the best solution ever computed from a previous generation is automatically replicated and inserted as a member of the next generation, before the reproduction (selection) process even starts. Each initial population has been generated randomly using a nearest-neighbor procedure to construct feasible solutions.

The results for the 56 Solomon's problems are summarized in Tables 1-6. For each problem set, the performance of the proposed algorithm is compared with the best computed results reported in [11] and obtained from various similar (tabu search and genetic-based) and standard (operations research-based [16]) techniques. The second column presents the number of routes and total traveled distance, respectively, for the best computed solution, whereas the third column refers to the corresponding results for our algorithm. An entry with a star symbol indicates a problem instance in which the number of routes for a computed solution matches the best known solution. As published results failed to report run-time for many methods associated with best known solutions, computational effort has been deliberately omitted for comparison purposes.

Table 1. Performance comparison for R1 problem set

Problem Set R1	**Routes/Total Distance (best)**	**Routes/Total Distance (Hybrid GA)**
R101	18/ 1607.7	19/ 1688.4
R102*	17/ 1434.0	17/ 1567.1
R103	13/ 1207.0	14/ 1310.2
R104*	10/ 982.0	10/ 1081.1
R105*	14/ 1377.1	14/ 1406.7
R106*	12/ 1252.0	12/ 1367.6
R107	10/ 1159.9	11/ 1145.8
R108	9/ 980.9	10/ 1002.6
R109	11/ 1235.7	12/ 1230.5
R110	10/ 1080.4	11/ 1133.5
R111	10/ 1129.9	11/ 1174.3
R112*	10/ 953.6	10/ 1031.1

Table 2. Performance comparison for C1 problem set

Problem Set C1	Routes/Total Distance (best)	Routes/Total Distance (Hybrid GA)
C101*	10/ 827.3	10/ 828.9
C102*	10/ 827.3	10/ 837.3
C103*	10/ 828.1	10/ 848.4
C104*	10/ 824.8	10/ 852.4
C105*	10/ 828.9	10/ 828.9
C106*	10/ 827.3	10/ 828.9
C107*	10/ 827.3	10/ 828.9
C108*	10/ 827.3	10/ 828.9
C109*	10/ 828.9	10/ 828.9

Table 3. Performance comparison for RC1 problem set

Problem Set RC1	Routes/Total Distance (best)	Routes/Total Distance (Hybrid GA)
RC101	14/ 1669.0	15/ 1696.7
RC102*	13/ 1477.5	13/ 1638.7
RC103*	11/ 1110.0	11/ 1392.1
RC104*	10/ 1135.8	10/ 1238.7
RC105	13/ 1733.6	14/ 1652.9
RC106*	12/ 1384.9	12/ 1416.8
RC107*	11/ 1230.9	11/ 1303.0
RC108	10/ 1170.7	11/ 1191.9

Table 4. Performance comparison for R2 problem set

Problem Set R2	Routes/Total Distance (best)	Routes/Total Distance (Hybrid GA)
R201*	4/ 1281.6	4/ 1448.5
R202	3/ 1530.5	4/ 1248.3
R203*	3/ 948.7	3/ 1075.6
R204	2/ 869.3	3/ 821.2
R205*	3/ 1063.2	3/ 1162.6
R206*	3/ 833.0	3/ 1056.4
R207*	3/ 814.8	3/ 891.9
R208*	2/ 738.6	2/ 761.9
R209	2/ 855.0	3/ 996.0
R210*	3/ 967.5	3/ 1047.5
R211	2/ 949.5	3/ 820.2

Results show that computed solutions compete with most of the best known solutions. The method proves to be quite satisfactory for clustered problems (C1, C2), as computed solution quality (distance and number of routes) nearly match best published results. But the procedure slightly degrades in quality, as total distance may differ up to 8% for R1 and RC1. This observation becomes more apparent for R2 and RC2 showing some weaknesses in efficiently combining routes (or vehicles) having a large number of customers. The particular knowledge subset used and/or currently designed genetic operators are suspected to insufficiently reduce and confine the search space to good solution neighborhoods for these problem instances (R2, RC2). Nevertheless, it is worth noticing that for all standard test problems the computed number of routes mostly corresponds or slightly differs from the best reported solutions. On the other hand, the computation of a near-optimal solution is reasonably fast (1-10 minutes on a Sun SPARC 10) in comparison to alternate methods, due to low cost operator execution. In that respect, additional speed-up could easily be achieved by optimizing the current algorithm implementation.

Table 5. Performance comparison for C2 problem set

Problem Set C2	Routes/Total Distance (best)	Routes/Total Distance (Hybrid GA)
C201*	3/ 591.6	3/ 591.6
C202*	3/ 591.6	3/ 591.6
C203*	3/ 591.2	3/ 600.2
C204*	3/ 590.6	3/ 616.6
C205*	3/ 588.9	3/ 588.9
C206*	3/ 588.5	3/ 588.5
C207*	3/ 588.3	3/ 588.3
C208*	3/ 588.3	3/ 588.3

Table 6. Performance comparison for RC2 problem set

Problem Set RC2	Routes/Total Distance (best)	Routes/Total Distance (Hybrid GA)
RC201*	4/ 1249.0	4/ 1616.1
RC202*	4/ 1165.6	4/ 1380.7
RC203*	3/ 1079.6	3/ 1222.3
RC204*	3/ 806.8	3/ 903.3
RC205*	4/ 1333.7	4/ 1465.2
RC206	3/ 1212.6	4/ 1215.3
RC207*	3/ 1085.6	3/ 1510.4
RC208*	3/ 834.97	3/ 960.7

Even though recently developed tabu search [17,18] techniques slightly outperform the proposed algorithm, the latter has shown similar or better performance than any previously reported (published) genetic-based methods. The stochastic procedure is conceptually simple, easy to implement and allows for the fast computation of a near-optimal solution; a desirable real-time feature when dealing with time-constrained reasoning. Significant improvement is further expected in exploiting alternate key feature combinations of promising routing heuristics to better balance intensification and diversification while exploring the solution space.

4. Conclusion

This effort toward solving VRPTW aims at investigating the impact of explicitly using domain knowledge and prior knowledge/characteristics about expected solutions during the recombination and mutation phases of a genetic algorithm. The approach is based upon some principles of recent hybrid and standard genetic algorithms and features of well-known heuristics to drive the search process. Simple and easy to implement, the stochastic procedure allows for the fast computation of a near-optimal solution, a suitable property to address time-constrained reasoning constraints. A computational experiment shows that even though the proposed hybrid genetic algorithm does not yet outperform some recent tabu search techniques, computed solutions remain very competitive and nearly match some of the best known solutions. Future work will explore the introduction of new features to the proposed hybrid genetic algorithm. Accordingly, a combination of various routing heuristics aimed at diversifying problem-solving strategies to search the solution space will be investigated. The contribution of alternate knowledge will also be examined to enrich the recombination phase of the algorithm. Additional research directions include the impact of partial constraint relaxation during problem-solving as well as knowledge associated with statistical distribution of key domain attributes over problem instances.

5. References

1. Desrosiers, J. and al.: Time Constrained Routing and Scheduling. In: Handbooks in Operations Research and Management Science, Vol. 8. Network Routing, M.O. Ball, T.L. Magnanti, C.L. Monma, G.L. Nemhauser Eds, North-Holland, Amsterdam, (1995) 35-139
2. Solomon, M.M. and Desrochers, J.: Time Window Constrained Routing and Scheduling Problems. Transportation Science 22 (1988) 1-13
3. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, New York, (1989)
4. Reeves, C.R.: Modern Heuristics Techniques for Combinatorial Problems. Halsted Press, New York, (1993)
5. Glover, F.: Tabu Search - Part I. ORSA Journal on Computing 1 (1989) 190-206
6. Glover, F.: Tabu Search - Part II. ORSA Journal on Computing 2 (1990) 4-32
7. Potvin, J-Y. and Bengio, S.: The Vehicle Routing Problem with Time Windows Part II: Genetic Search. INFORMS Journal on Computing 8(2) Spring (1996)
8. Thangiah, S.R. and al.: Vehicle Routing with Time Deadlines using Genetic and Local Algorithms. Proceedings of the 5th International Conference on Genetic Algorithms, University of Illinois at Urbana-Champaign, (1993) 506-513
9. Thangiah, S.R.: An Adaptive Clustering Method using a Geometric Shape for Vehicle Routing Problems with Time Windows. Proceedings of the 6th International Conference on Genetic Algorithms, University of Pittsburg, (1995) 536-543
10. Blanton, J.L. and Wainwright, R.L.: Multiple Vehicle Routing with Time and Capacity Constraints using Genetic Algorithms. Proceedings of the 5th International Conference on Genetic Algorithms, Champaign, IL, (1993), 452-459
11. Rochat, Y. and Taillard, E.D.: Probabilistic diversification and intensification in local search for vehicle routing. Journal of Heuristics 1(1) (1995) 147-167
12. Kirkpatrick, S., Gelatt, C.D. and Vecchi, P.M.: Optimization by Simulated Annealing. Science 220 (1983) 671–680
13. Solomon, M.M.: Algorithms for the Vehicle Routing and Scheduling Problems with Time Window Constraints. Operations Research 35(2) (1987) 254-265
14. Bodin, L. and al.: The State of the Art in the Routing and Scheduling of Vehicles and Crews. Computers and Operations Research 10(2) (1983).
15. Wall, M.: GAlib - A C++ Genetic Algorithms Library, version 2.4. (http://lancet.mit.edu/galib-2.4/), MIT, Mass., (1995)
16. Desrochers M., and al.: A New Optimization Algorithm for the Vehicle Routing Problem with Time Window. Operations Research 40 (1992) 342-354
17. Taillard, E.D. and al.: A Tabu Search Heuristic for the Vehicle Routing Problem with Soft Time Windows. Transportation Science 31 (1997) 170-186
18. Badeau, P. and al.: A Parallel Tabu Search Heuristic for the Vehicle Routing Problem with Time Windows. Transportation Research-C 5 (1997) 109-122

Distribution list

INTERNAL DISTRIBUTION

1 - Director General
3 - Document Library
1 - Dr. É. Bossé
1 - Mr. J. Berger (author)
1 - Dr. A. Guitouni
1 - Dr. A. Benaskeur
1 - Mr. A. Boukhtouta
1 - Mrs. M. Bélanger
1 - Dr I. Abi-Zeid
1 - Mr. M. Blanchette
1 - Mr. D. Demers
1 - Dr R. Breton
1 - Dr A.-C. Bourry-Brisset
1 - M. J. Roy
1 - LCdr D. Morrissey
1 - Maj. B. Deschênes
1 - Mr D. Gouin
1 - Mr R. Charpentier
1 - Mr G. Thibault
1 - Mr Y. Van Chestein

EXTERNAL DISTRIBUTION

1 - DRDKIM (unbound copy)
1 - DRDC
1 – Director – Science and Technology (Air) 3
    Bob Elliott
1 – Director – Aerospace Requirements 4
1 – Director – Avionics, Simulators and Photography 4
1 – Director – Land Requirements 4
1 – Director – Naval Requirements 4
1 - DRDC Toronto
    Att: Ross Pigeau
8 - Air Command Headquarters
    Westwin, MB R3J 0T0
    Attn: Charles Hunter
        Deputy Chief of Staff Ops Research
    Attn: Dale Reding
        Centre for Operational Research and
    Analysis (CORA) 2
    Attn: Dr Bao
        CORA
    Attn: LCol Hunter
        A3 Plans
    Attn: Major P.P.J. Lessard
        G3 Plans
    Attn: LCol Shawn E. Burtenshaw
        2 Wing Operations Officer
    Attn: Major C. Lynn Doucette
        A3 Exercise Coordination 3
    Attn: Major W.F. Seymour
        A3 Exercice Coordination 2
1 - 3D Thrust Leader: Chris McMillan, DRDC Ottawa
NDHQ
- Major L. Lefebvre, AFCCIS Project Manager

DOCUMENT CONTROL DATA

1. ORIGINATOR (name and address) Jean Berger, Martin Salois RDDC Valcartier 2459 Pie-XI Blvd. North, Val-Belair, Qc G3J 1X5, CANADA		2. SECURITY CLASSIFICATION (Including special warning terms if applicable) Unclassified
3. TITLE (Its classification should be indicated by the appropriate abbreviation (S, C, R or U) A Hybrid Genetic Algorithm for the Vehicle Routing Problem with Time Windows (U)		
4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.) J. Berger and M. Salois		
5. DATE OF PUBLICATION (month and year) June 2004	6a. NO. OF PAGES 25	6b. NO. OF REFERENCES 17
7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. Give the inclusive dates when a specific reporting period is covered.) Tech Report		
8. SPONSORING ACTIVITY (name and address)		
9a. PROJECT OR GRANT NO. (Please specify whether project or grant) Project 3dm		9b. CONTRACT NO.
10a. ORIGINATOR'S DOCUMENT NUMBER		10b. OTHER DOCUMENT NOS N/A

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

- [x] Unlimited distribution
- [ ] Contractors in approved countries (specify)
- [ ] Canadian contractors (with need-to-know)
- [ ] Government (with need-to-know)
- [ ] Defense departments
- [ ] Other (please specify)

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

A Variety of hybrid genetic algorithms has been recently proposed to address the vehicle routing problem with time windows (VRPTW), a problem known to be NP-hard. However, very few genetic-based approaches exploit implicit knowledge provided by the structure of the intermediate solutions computed during the evolutionary process to explore the solution space. This report presents a new hybrid genetic algorithm for the VRPTW. It investigates the impact of using explicitly domain knowledge and prior knowledge/characteristics about typical solutions expected from the recombination and mutation phases of the algorithm. Basic principles borrow from recent hybrid and standard genetic algorithms, and features of well-known heuristics to drive the search process. Designed to support time-constrained reasoning tasks, the procedure is entended to be conceptually simple, easy to implement and allow fast computation of near-optimal solution. A computational experiment has been conducted to compare the performance of the proposed algorithm with similar and standard techniques.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

genetic algorithms, vehicle routing problem with time windows, VRPTW, optimization, heuristics

dcd03e