Distance Metrics and Fitness–Distance Analysis for the Capacitated Vehicle Routing Problem

Marek Kubiak*

*Institute of Computing Science, Poznan University of Technology
Piotrowo 3A, 60-965 Poznan, Poland
Marek.Kubiak@cs.put.poznan.pl

1 Introduction

Metaheuristics, usually inspired by nature, define schemes of algorithms which have to be further adapted to a given problem in order to be practically useful (e.g. crossover operators have to be designed for an evolutionary algorithm).

One of the conclusions which may be derived from the No Free Lunch theorem [12] states that there is no algorithm which performs better than any other on all possible optimisation problems. Thus, an algorithm may be efficient only on a certain subclass of all problems, while in other cases it will perform worse than others, even a random search or a complete enumeration of solutions. This conclusion applies to metaheuristics as well. In this context the process of adaptation of a metaheuristic may be viewed as a means of moving the ‘efficient’ (for this algorithm) subclass of problems in the space of all possible problems.

This process of adaptation, the design of components of a metaheuristic, may profoundly influence the efficiency of an algorithm. For a class of metaheuristics, namely evolutionary algorithms, this phenomenon was even observed experimentally [8]. Thus, the adaptation of a metaheuristic to the problem is a crucial operation which should be well justified and carefully performed.

In many cases, however, this adaptation is done by a designer based on intuition and experience. It may result, of course, in an efficient algorithm for the problem, which is quite often the case, but the algorithm may just as well be a poorly performing one, which happens not so rarely. In addition, this way of adaptation provides neither justification for the choices made by the designer, nor gives knowledge which might be useful for others in the future.

A different way of adaptation of a metaheuristic algorithm to a given problem has received some attention recently. It is based on statistical analyses of the search space (fitness landscape) of instances of the problem. Such analyses, e.g. fitness–distance analysis [2, 3, 5, 7], random–walk correlation measurement [7], estimation of the size of attractor of local optima [7], provide objective information about certain properties of the fitness landscape and justify the design or choice of components for a metaheuristic algorithm [3, 5, 7].

Vienna, Austria, August 22–26, 2005
Therefore, the first goal of this paper is to present results of statistical analysis, namely the fitness–distance analysis, of some instances of the Capacitated Vehicle Routing Problem (CVRP) taken from literature. Such analysis may be further used to justify certain designs of metaheuristics to the problem and to provide explanation for efficiency of existing algorithms.

The second goal of this work is to present distance metrics for solutions of the CVRP. In the author’s opinion many publications about algorithms for combinatorial optimisation problems describe fitness landscapes using the notion of distance without actually providing its definition; moves in the landscape are told to be near of far jumps, search strategies are said to intensify the search in a promising region of the search space or to diversify the search by identifying new regions with good solutions. However, accurate, objective, quantitative information about what is near or what is a region is rarely provided, with few exceptions [9, 11]. Thus, distance metrics for the CVRP, presented in this paper, demonstrate that distance may be defined strictly and become a useful tool for verification of research intuition.

2 The Capacitated Vehicle Routing Problem

The CVRP [1, 5, 10] is a very basic formulation of a problem which a transportation company might face in its everyday operations. The goal is to find the shortest–possible set of routes for the company’s vehicles in order to satisfy demands of customers for certain goods. Each of identical vehicles starts and finishes its route at the company’s depot, and must not carry more goods than its capacity specifies. All customers have to be serviced, each exactly once by one vehicle. Distances between the depot and customers are given.

The version of CVRP considered here does not fix the number of vehicles (it is a decision variable): also the distance to be travelled by a vehicle is not constrained. Names of instances used in this study are provided in table 1, in section 4.1.

In order to properly define distance metrics in the next section, some basic definitions related to solutions of the CVRP have to be given. A solution \( s \) of the CVRP is a set of \( T(s) \) routes:

\[
 s = \{ t_1, t_2, \ldots, t_{T(s)} \}.
\]

A route is a sequence of customers (nodes) starting with the depot, \( v_0 \), and has the form:

\[
 t_i = (v_0, v_{i,1}, v_{i,2}, \ldots, v_{i,n(t_i)}) \quad \text{for } i = 1, \ldots, T(s),
\]

where \( n(t_i) \) is the number of customers on route \( t_i \).

Some additional specific constraints should also be satisfied by a feasible solution, but they are not provided here. Refer to [1, 5, 10] for more information about the CVRP.

3 Distance metrics for solutions of the CVRP

The distance metrics presented in this section correspond to certain properties of solutions of the CVRP which are deemed to be important for their quality: existence of certain edges (or even paths) or specific way of partition of the set of customers into routes (clusters).
3.1 Distance in terms of edges: \(d_e\)

The idea of this metric is based on a very similar concept formulated for the travelling salesman problem (TSP): the number of common edges in TSP tours [2]. In the research cited it was found that good solutions of the TSP have many common edges and are, thus, very similar (close) to each other. Due to similarity between solutions of the TSP (one tour) and the CVRP (a set of disjoint tours/routes) the idea of common edges may be easily adapted to the latter.

In order to properly define the distance metric some definitions are required:

\[
E(t_i) = \{\{v_0, v_{i,1}\}\} \cup \left( \bigcup_{j=1}^{n(t_i)-1} \{\{v_{i,j}, v_{i,j+1}\}\} \right) \cup \{\{v_{i,n(t_i)}, v_0\}\} \\
E(s) = \bigcup_{t_i \in s} E(t_i)
\]

\(E(t_i)\) is a multiset of undirected edges appearing in route \(t_i\). \(E(s)\) is a multiset of edges in solution \(s\). The notion of a multiset is required here, because routes in some solutions of the CVRP may include certain edges twice.

Using the general concept of distance between subsets of the same universal set, as defined by Marczewski and Steinhaus [6] (cited after [4]), the distance \(d_e\) between two solutions \(s_1, s_2\) of the same CVRP instance may be defined as:

\[
d_e(s_1, s_2) = \frac{|E(s_1) \cup E(s_2)| - |E(s_1) \cap E(s_2)|}{|E(s_1) \cup E(s_2)|}
\]

Due to the fact that \(d_e\) is only a special case of the Marczewski–Steinhaus distance, it inherits all its properties of a metric; its values are also normalized to the interval [0,1].

This distance metric perceives solutions of the CVRP as multisets of edges: solutions close to each other will have many common edges; distant solutions will have few common ones. However, closer investigation of the metric reveals that it is not intuitively ‘linear’ (although it is ‘monotonic’), e.g. \(d_e = 0.5\) does not mean that exactly half of each \(E(s_i)\) is common; 50% of common edges implies \(d_e \approx 2/3\).

3.2 Distance in terms of partitions of customers: \(d_{pc}\)

The concept behind the second distance metric is based on the ‘cluster first – route second’ heuristic approach to the CVRP [1]: first find a good partition of customers into clusters and then try to find routes (solve TSPs) within these clusters, separately. According to this idea the distance metric should identify dissimilarities between solutions perceived as partitions of the set of customers.

An example of a distance metric for partitions of a given set may be found in [4] (it is even more generally defined there, for hypergraphs or binary trees). This example is easily adaptable to solutions of the CVRP. Let us define:

\[
C(s) = \{c_1(s), c_2(s), \ldots, c_T(s)\}
\]
\[ c_i(s) = \{v_{i,1}, v_{i,2}, \ldots, v_{i,n(t_i)}\} \]

\[ \sigma(c_i(s_1), c_j(s_2)) = \frac{|c_i(s_1) \cup c_j(s_2)| - |c_i(s_1) \cap c_j(s_2)|}{|c_i(s_1) \cup c_j(s_2)|} \]

\( C(s) \) is a partition of the set of customers into clusters; one cluster, \( c_i(s) \), holds customers from route \( t_i \) of \( s \); \( \sigma(\cdot) \) is the distance between two clusters.

According to [4], the distance between solutions may be defined as:

\[
d_{pc}(s_1, s_2) = 1/2 \left\{ \max_{i=1}^{T(s_1)} \T(s_2) \min_{j=1}^{T(s_2)} \sigma(c_i(s_1), c_j(s_2)) + \max_{i=1}^{T(s_2)} \T(s_1) \min_{j=1}^{T(s_1)} \sigma(c_i(s_1), c_j(s_2)) \right\}
\]

This function is a distance metric for partitions; it is also normalized. It is not exactly a metric for solutions of the CVRP, because \( d_{pc}(s_1, s_2) = 0 \) does not imply \( s_1 = s_2 \) (the number of solutions which are not discriminated by \( d_{pc} \) may be exponentially large).

Formula 1 has the following sense: firstly, the best–possible assignment of clusters from \( C(s_1) \) to clusters from \( C(s_2) \) is made (the one which minimizes \( \sigma(\cdot) \)), and vice versa; that is the idea behind internal min operators. Secondly, two worst assignments are chosen among those pairs (the max operators), and distance in those assignments is averaged to form the overall distance between partitions. Thus, it may be concluded that \( d_{pc} \) is somehow 'pessimistic' in the choice of 'optimistic' matches of clusters.

This mixture of max and min operators in \( d_{pc} \) makes interpretation of its values difficult. Certainly, values near to 0 indicate great similarity of solutions. However, larger values do not necessarily indicate very dissimilar partitions; it is sufficient that there are 'outliers' in partitions, which can hardly be well assigned to clusters in the other solution, and the max operator will result in large values, implying distant solutions.

### 3.3 Distance in terms of pairs of nodes: \( d_{pn} \)

The third distance metric, \( d_{pn} \), is based on the same idea as \( d_{pc} \): distance between solutions viewed as partitions of the set of customers. However, this idea has a different, more straight-forward, mathematical formulation in \( d_{pn} \). Here, the Marczewski–Steinhaus [6] concept of distance is applied to sets of pairs of nodes (customers).

Let's define:

\[
PN(t_i) = \bigcup_{j=1}^{n(t_i) - 1} \bigcup_{k=j+1}^{n(t_i)} \{v_{i,j}, v_{i,k}\}
\]

\[
PN(s) = \bigcup_{t_i \in s} PN(t_i)
\]

\( PN(t_i) \) is the set of undirected pairs of nodes (customers) which are assigned to the same route \( t_i \) (it is similar to a clique). \( PN(s) \) is the set of all such pairs in solution \( s \).

The distance \( d_{pn} \) between solutions is defined as:

\[
d_{pn}(s_1, s_2) = \frac{|PN(s_1) \cup PN(s_2)| - |PN(s_1) \cap PN(s_2)|}{|PN(s_1) \cup PN(s_2)|}
\]

Vienna, Austria, August 22–26, 2005
Similarly to $d_{pc}$, this function is not exactly a metric for solutions of the CVRP, but for partitions implied by those solutions.

Formula 2 has a more straightforward sense than 1; here, the value of distance roughly indicates how large are parts of sets of pairs which are not shared by two compared solutions. If $d_{pn} = 0$ then two solutions imply identical partitions; $d_{pn} = 1$ implies completely different partitions (not even one pair of nodes is assigned to a route in the same way in $s_1$ and $s_2$).

4 Fitness–distance analysis of the CVRP

4.1 Random solutions vs. local optima

The first stage of the fitness–distance analysis focuses on possible differences between sets of local optima and random solutions of a given instance in terms of distance in these sets.

In order to check these differences, large random samples of 2000 different solutions of each type were generated:

- a random solution: a number of routes was randomly drawn and then these routes were randomly filled with customers (uniform probability whenever possible); if a route could not be continued due to the capacity constraint, the vehicle returned to the depot and a new route was started;
- a local optimum: a random solution was subject to a greedy local search with a neighbourhood of 3 joined operators: 2–opt, exchange of 2 customers, joining of 2 routes [5].

Finally, statistics on values of distance in these samples were computed, as shown in table 1. It may be seen in the table that average values of each type of distance in samples of local optima are usually much lower than values of distance for the best random solutions. It is $d_e$ which usually results in the highest gains.

It might be concluded that local optima are clustered together in some parts of the fitness landscapes rather than scattered all over it. This also means that they usually share many common properties: edges or assignments of customers to same routes.

4.2 Trends in sets of local optima

The second stage of the fitness–distance analysis is an attempt to find trends in the sets of local optima themselves: if when solutions become better they tend to be more similar (close) to each other [2, 3, 5, 7, 9].

This hypothesis was tested in the following manner: first, each type of distance was computed for each of 1000 pairs of local optima from the previously generated samples. Then, pairs of solutions with roughly the same quality (max. 5% difference) were arranged on scatter plots of fitness versus distance. Finally, lines of regression were added and correlations computed.

Vienna, Austria, August 22–26, 2005
Table 1: Values of distance for the best random solutions and average distance for local optima (differences with the values for random solutions). Also the squares of fitness–distance correlation (FDC) values are shown.

<table>
<thead>
<tr>
<th>name</th>
<th>best_rnd</th>
<th>avg_lopt</th>
<th>(\rho^2_{f,d} )</th>
<th>(d_e)</th>
<th>(d_{pc})</th>
<th>(d_{pn})</th>
<th>(\rho^2_{f,d_e} )</th>
<th>(\rho^2_{f,d_{pc}} )</th>
<th>(\rho^2_{f,d_{pn}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>c100</td>
<td>0.977</td>
<td>0.908</td>
<td>0.950</td>
<td>-0.304</td>
<td>-0.142</td>
<td>-0.219</td>
<td>0.197</td>
<td>0.062</td>
<td>0.126</td>
</tr>
<tr>
<td>c100b</td>
<td>0.959</td>
<td>0.904</td>
<td>0.963</td>
<td>-0.419</td>
<td>-0.201</td>
<td>-0.374</td>
<td>0.347</td>
<td>0.211</td>
<td>0.468</td>
</tr>
<tr>
<td>c120</td>
<td>0.976</td>
<td>0.902</td>
<td>0.937</td>
<td>-0.262</td>
<td>-0.108</td>
<td>-0.191</td>
<td>0.020</td>
<td>0.091</td>
<td>0.152</td>
</tr>
<tr>
<td>c150</td>
<td>0.972</td>
<td>0.917</td>
<td>0.958</td>
<td>-0.293</td>
<td>-0.107</td>
<td>-0.199</td>
<td>0.296</td>
<td>0.015</td>
<td>0.232</td>
</tr>
<tr>
<td>c199</td>
<td>0.972</td>
<td>0.929</td>
<td>0.970</td>
<td>-0.271</td>
<td>-0.103</td>
<td>-0.179</td>
<td>0.237</td>
<td>0.014</td>
<td>0.242</td>
</tr>
<tr>
<td>c50</td>
<td>0.946</td>
<td>0.894</td>
<td>0.935</td>
<td>-0.341</td>
<td>-0.151</td>
<td>-0.276</td>
<td>0.167</td>
<td>0.069</td>
<td>0.173</td>
</tr>
<tr>
<td>c75</td>
<td>0.909</td>
<td>0.882</td>
<td>0.948</td>
<td>-0.221</td>
<td>-0.083</td>
<td>-0.196</td>
<td>0.117</td>
<td>0.012</td>
<td>0.113</td>
</tr>
<tr>
<td>f134</td>
<td>0.972</td>
<td>0.929</td>
<td>0.952</td>
<td>-0.339</td>
<td>-0.147</td>
<td>-0.243</td>
<td>0.160</td>
<td>0.099</td>
<td>0.045</td>
</tr>
<tr>
<td>f71</td>
<td>0.962</td>
<td>0.908</td>
<td>0.926</td>
<td>-0.485</td>
<td>-0.262</td>
<td>-0.588</td>
<td>0.237</td>
<td>0.046</td>
<td>0.378</td>
</tr>
<tr>
<td>tai100a</td>
<td>0.949</td>
<td>0.911</td>
<td>0.951</td>
<td>-0.354</td>
<td>-0.126</td>
<td>-0.261</td>
<td>0.083</td>
<td>0.036</td>
<td>0.113</td>
</tr>
<tr>
<td>tai100b</td>
<td>0.907</td>
<td>0.900</td>
<td>0.978</td>
<td>-0.297</td>
<td>-0.094</td>
<td>-0.280</td>
<td>0.238</td>
<td>0.032</td>
<td>0.233</td>
</tr>
<tr>
<td>tai100c</td>
<td>0.978</td>
<td>0.902</td>
<td>0.963</td>
<td>-0.352</td>
<td>-0.093</td>
<td>-0.256</td>
<td>0.246</td>
<td>0.036</td>
<td>0.380</td>
</tr>
<tr>
<td>tai100d</td>
<td>0.944</td>
<td>0.894</td>
<td>0.933</td>
<td>-0.335</td>
<td>-0.070</td>
<td>-0.256</td>
<td>0.178</td>
<td>0.082</td>
<td>0.272</td>
</tr>
<tr>
<td>tai150a</td>
<td>0.932</td>
<td>0.947</td>
<td>0.992</td>
<td>-0.305</td>
<td>-0.128</td>
<td>-0.261</td>
<td>0.009</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>tai150b</td>
<td>0.961</td>
<td>0.921</td>
<td>0.981</td>
<td>-0.298</td>
<td>-0.095</td>
<td>-0.250</td>
<td>0.072</td>
<td>0.035</td>
<td>0.223</td>
</tr>
<tr>
<td>tai150c</td>
<td>0.958</td>
<td>0.928</td>
<td>0.975</td>
<td>-0.329</td>
<td>-0.116</td>
<td>-0.239</td>
<td>0.190</td>
<td>0.046</td>
<td>0.217</td>
</tr>
<tr>
<td>tai150d</td>
<td>0.979</td>
<td>0.901</td>
<td>0.958</td>
<td>-0.343</td>
<td>-0.072</td>
<td>-0.208</td>
<td>0.042</td>
<td>0.011</td>
<td>0.057</td>
</tr>
<tr>
<td>tai385</td>
<td>0.950</td>
<td>0.938</td>
<td>0.989</td>
<td>-0.205</td>
<td>-0.060</td>
<td>-0.128</td>
<td>0.122</td>
<td>0.047</td>
<td>0.158</td>
</tr>
<tr>
<td>tai75a</td>
<td>0.907</td>
<td>0.887</td>
<td>0.968</td>
<td>-0.286</td>
<td>-0.092</td>
<td>-0.253</td>
<td>0.165</td>
<td>0.021</td>
<td>0.156</td>
</tr>
<tr>
<td>tai75b</td>
<td>0.902</td>
<td>0.891</td>
<td>0.956</td>
<td>-0.284</td>
<td>-0.083</td>
<td>-0.224</td>
<td>0.025</td>
<td>0.026</td>
<td>0.075</td>
</tr>
<tr>
<td>tai75c</td>
<td>0.945</td>
<td>0.889</td>
<td>0.918</td>
<td>-0.364</td>
<td>-0.101</td>
<td>-0.236</td>
<td>0.187</td>
<td>0.034</td>
<td>0.258</td>
</tr>
<tr>
<td>tai75d</td>
<td>0.944</td>
<td>0.875</td>
<td>0.950</td>
<td>-0.425</td>
<td>-0.112</td>
<td>-0.401</td>
<td>0.244</td>
<td>0.059</td>
<td>0.277</td>
</tr>
</tbody>
</table>

Samples of such plots, which are representative to the instances listed in table 1, are presented in figures 1 and 2. In these plots one might notice some (not very strong; see table 1) correlation between fitness and distance, which indicates that better local optima tend to be even closely clustered than worse ones. The strongest tendencies are observed generally for \(d_e\) and \(d_{pn}\); \(d_{pc}\) is worse perhaps due to its ‘pessimistic’ nature mentioned in section 3.2.

5 Conclusions

The study presented in the paper revealed a certain structure existing in the fitness landscapes of instances of the CVRP. This structure might explain why efficient algorithms for this problem have to include local search components and also why the intensification technique presented in [10], which has been extremely conservative in changing ‘good’ edges or routes, has been so successful.

Vienna, Austria, August 22–26, 2005
The paper also confirms that fitness–distance analysis should become a tool widely used by designers of metaheuristic algorithms, because once the notion of distance is clearly and unambiguously defined, the analysis may provide objective information about properties of solutions which are important to the overall fitness. A research intuition might be, therefore, confirmed or rejected by this kind of analysis [3, 5, 7].

Acknowledgements

This research has been supported by the Polish State Committee for Scientific Research (KBN) through grant no. 3T11F00426.

References


Figure 2: Fitness–distance plots with local optima for instance c199 and 3 distance types: $d_e$, $d_{pc}$, $d_{pn}$, together with lines of regression.


Vienna, Austria, August 22–26, 2005