DE GRUYTER
OPEN

Przedsiębiorczość i Zarządzanie
Entrepreneurship and Management
University of Social Sciences Publishing House
ISSN 1733–2486
Volume XV, Issue 2, pp. 157–169
DOI 10.2478/eam-2014-0024

Radosław Jadczyk

University of Lodz

Traveling salesman problem: approach to optimality

Abstract: Traveling Salesman Problem (TSP) is a basic and one of the most important transportation problems in operational logistics. It is also known in the literature as a Chinese postman problem or single vehicle routing problem. TSP can be shortly described as follows. Vehicle starting from the selected city must visit a set of another cities exactly once and return to the starting city in such a way that the total distance of the route is minimized.

In this paper first mathematical formulation of decision problem is presented. Then solution strategies of TSP are shown with selected algorithms as examples. In the last part of article, a computational results of selected methods are presented.

Key-words: traveling salesman problem, heuristics, Lin-Kernighan algorithm.

Introduction

Traveling Salesman Problem (TSP) is a basic and one of the most important transportation problems in operational logistics. It is also known in the literature as a Chinese postman problem or single vehicle routing problem. TSP can be described as follows [Całczyński 2000, Ignasiak 1996]. Vehicle starting from the fixed point P_o (base) must visit N points: $P_1, P_2, \dots, P_N$ (cities, customers) and return to starting point P_o . All points must be visited and exactly once. The objective is to find the shortest such trip.

The traveling salesman problem was presented first time in 1859 by W.R. Hamilton, while H. Whitney first time formulated TSP as a combinatorial problem in 1934 [Całczyński 1992].

The main purpose of this article is to present and compare a quality of the simple and fast heuristics used to solve TSP problems in practice by the decision makers who do not have advanced software tools.

In this paper first mathematical formulation of decision problem is presented. Then solution strategies of TSP are shown with selected algorithms as examples. In the last part of article, a computational results of selected methods are presented for a number of test problems.

Problem formulation

Traveling Salesman Problem may be presented as a discrete optimization model [Szymanowski 1977, Szymanowski 1979]. First, it is needed to define decision variables:

$$x_{ij} = \begin{cases} 0 \\ 1 \end{cases} \quad \text{dla } i, j = 0, \dots, N \quad (1)$$

Variable x_{ij} indicates a connection between points P_i and P_j . It takes the value 1, if in the route of vehicle is a connection between points P_i and P_j and 0 in the other case.

Since, each point must be visited by the vehicle only once, constraints in the model are as follows:

$$\sum_{i=0}^N x_{ij} = 1 \quad \text{dla } j = 0, \dots, N \quad (2)$$

$$\sum_{j=0}^N x_{ij} = 1 \quad \text{dla } i = 0, \dots, N \quad (3)$$

$$z_i - z_j + (N+1)x_{ij} \leq (N+1) - 1 \quad \text{dla } i, j = 1, \dots, N; i \neq j; z_i, z_j \in R \quad (4)$$

The vehicle may enter only once given point, and to leave it what is shown by equations (2) and (3).

Another very important condition is that all points in the route must be visited. It is provided by set of inequalities (4). The number of these inequalities is very large and is $N(N-1)$. They contain other variables z_i and z_j , appropriate value of which make condition fulfilled. In other words, inequalities (4) provide that with the element of the set $\{P_i, P_j / x_{ij}=1\}$ it is possible to build the sequence: $(P_0, P_1), (P_1, P_2), (P_2, P_3), \dots, (P_{N-1}, P_N), (P_N, P_0)$ [Solich 1974]. No restrictions described by inequalities (4) transforms TSP to the assignment problem, which one of the solutions is shown on the figure 2.

The goal of the traveling salesman problem is to minimize following function (5):

$$f(x) = \sum_{i=0}^N \sum_{j=0}^N d_{ij} x_{ij} \rightarrow \min \quad \text{dla } i, j = 0, \dots, N \quad (5)$$

Function (5) presents vehicle total length. Factor d_{ij} indicates length between points P_i and P_j . It means, that d_{ij} are the elements of the length matrix D sizes $(N+1) \times (N+1)$.

In the traveling salesman model (1)–(5) the number of decision variables is $(N+1)^2$ and the number of all constraints is equal $(N+1) + (N+1) + N^2$.

Another definition of the traveling salesman problem may be formulated in graph theory [Sysło 1984]. TSP-problem can be presented in the form of undirected, complete and weighted graph $G(V, E, d)$, such that points are the graph's vertices (nodes). The number of vertices is equal $N+1$ (including starting point). Edges connecting each pair of vertices i and j represent paths between visited points. Weight d_{ij} is the distance (length) between nodes i and j . One selected node is the starting (and ending) representing point P_0 , where salesman starting and ending his route. The TSP route is the set of edges, where:

- first edge of the route begins in starting node;
- last edge of the route ends in starting node
- each node of the graph connects two and only two edges.

The set of the graph's edges is the sequence of connections between all graph's vertices and it is called Hamiltonian cycle in graph theory. Therefore, we can say that TSP route is feasible, if it is a Hamiltonian cycle (see Fig. 1). On the figure below there are shown all possible graph's edges for 8 nodes (including starting point P_0) – solid and dotted lines.

Figure 1. Example of Hamiltonian cycle in TSP problem

Source: own elaboration.

On the next figure (see Fig. 2) there is shown infeasible TSP route – solid line. Vehicle starting route from point P_0 is visiting points: P_1, P_5, P_6, P_7 (or: P_7, P_6, P_5, P_1) and returns to point: P_0 . Nodes P_2, P_3, P_4 have been omitted. It means, that these points would have to be visited after restarting the route from

point P_0 . This case cannot be acceptable, because one of the main assumptions (and constraints) says, that vehicle leave the starting point only once.

Figure 2. Example of infeasible TSP route

Source: own elaboration.

Optimal solution of travelling salesman problem in this formulation is the route, where total sum of edge's weights d_{ij} belonging to Hamiltonian cycle is the smallest.

The travelling salesman problem is the one of the most common problems of many logistic companies. It is also in the interest of mathematicians involved in combinatorial optimization. Formulation of the TSP problem is very easy, but to find optimal solution or even near optimal solution is not so clear. The number of feasible solutions (Hamiltonian cycles) is equal to number of all permutations of visiting points. Assuming that the route may be covered in the both directions, only for three points: P_1 , P_2 i P_3 , the number of possible routes is 6 (see Fig. 3).

Figure 3. Possible variants of the vehicle route for three points

Source: own elaboration.

If the number of points will increase the number of feasible solutions will grow at an exponential. For the TSP problem, where vehicle must visit N points and distance matrix $\mathbf{D}$ is not symmetric ($d_{ij} \neq d_{ji}$), the number of all possible routes will be $N!$. Table below (see Tab. 1) shows correlation between the number of visiting points and the number of possible routes.

Table 1. The number of possible routes for asymmetric ($d_{ij} \neq d_{ji}$) and symmetric TSP ($d_{ij} = d_{ji}$)

Number of points	Number of routes (asymmetric)	Number of routes (symmetric)
1	---	---
2	2	1
3	6	3
4	24	12
5	120	60
6	720	360
7	5 040	2570
8	40 320	20 160
9	362 880	181 440
10	3 628 800	1 814 400
15	1 307 674 368 000	653 837 184 000
25	$1,55 \cdot 10^{25}$	$0,77 \cdot 10^{25}$
50	$3,04 \cdot 10^{64}$	$1,52 \cdot 10^{64}$
75	$2,48 \cdot 10^{109}$	$1,24 \cdot 10^{109}$
100	$9,33 \cdot 10^{157}$	$4,66 \cdot 10^{157}$

Source: own elaboration.

Solutions strategies for TSP

The simplest way to find optimal solution for travelling salesman problem is to search all solution space (brute force strategy). List all possible permutations of the points and choose one with the shortest tour length. Unfortunately solution space is too large, so this procedure is not efficient. Time to find the optimal solution for large cases is not acceptable.

Due to the fact that there are not optimal solution (or even near optimal solution) in many parts of the solution space, search strategy should to skip these parts. This strategy is represented in branch and bound method (for example Little algorithm). Solution space is divided for two sets: containing or not connection between to point in the route. Next sets are further subdivided

for other sets and so on until complete route is obtained (set of marked earlier connections).

The Little algorithm is an exact algorithm [Little et al. 1963]. It means, that obtained solution is optimal. In the worst case, for N serving points, the total number of examined sets of the solution space may equal $N!$. It means examining all possible solutions, in other words all possible permutations of serving points (as in the case of brute force strategy). This procedure for large TSP problems is also inefficient.

An alternative for exact methods are heuristic methods. Unfortunately they don't guarantee finding optimal solutions. However their main advantages are:

- simplicity of implementation;
- no need to have special mathematical knowledge about solving problem;
- short searching time of solution compared to exact methods;
- the quality of found solution is very often satisfactory for decision maker (first of all in terms of its search time).

Good heuristic algorithms can find near optimal solution, while these more sophisticated even optimal.

Among the heuristic methods for solving traveling salesman problem two main strategies should be distinguished: construction strategy and improvement strategy (local optimization).

In the first one, TSP route is building step by step by adding new serving point to the partially constructed route (see Fig. 4).

Figure 4. Construction strategy

Source: own elaboration.

The second strategy is the systematic, according to a specific procedure improvement of already built route (see Fig. 5).

Figure 5. Improvement strategy

Source: own elaboration.

Representative of the construction strategy and one of the most popular method is the nearest neighbor search algorithm. The procedure of this algorithm can be stated as follows [Krawczyk 2001].

Step 1

Build the route $T = \{P_0\}$

Step 2 Select point P_j which is the nearest to the last point (last added) in the partially constructed route.

Step 3

Add selected point P_j on the end of the route.

Step 4

Repeat step 2 and 3 until all serving points will be added to the route.

As it is shown above, the procedure is very simple. Vehicle starting from point P_0 , goes to the closest unvisited point. Then starting from this point, vehicle search next point which is located closest to the last visited. This procedure is repeated until TSP route will be completed.

Another, more sophisticated procedure, which realize construction strategy is nearest insertion algorithm [Sysło 1984, Krawczyk 2001].

Step 1

Build the route $T = \{P_0\}$

Step 2

Select point P_j which is the closest to the partially constructed route.

Step 3

Insert selected point P_j in the best place of the partially constructed route.

Step 4

Repeat step 2 and 3 until all serving points will be added to the route.

This algorithm requires an explanation in step 2 and 3. Choosing the point closest to the route, there should be defined the distance of unvisited point

form partially constructed route. This distance can be defined as the length between unvisited point and closest located point in the partially constructed route (cycle). Form all unvisited points, there should be selected one, which distance from cycle is the smallest. Then selected point P_j should be inserted in the best place. In the cycle may be many places, where this point could be inserted. In step 3, for each pair of points P_m and P_n belonging to the partially constructed route, there should be calculated cost of insertion: $\Delta_{m,n} = d_{m,j} + d_{j,n} - d_{m,n}$. Insert point P_j between pair of point P_m and P_n where cost of insertion $\Delta_{m,n}$ is the smallest.

Representative of the improvement strategy is the 2-opt algorithm [Lin 1973]. This is also the representative of wide class of local search heuristic algorithms.

Step 1

Build the route T (Hamiltonian cycle)

Step 2

Perform the following for all nodes until improvement of tour length is obtained:

- select a node i
- examine all 2-opt moves involving the edge between node i and its successor in the cycle; if it is possible do decrease the cycle length this way and construct new route T .

Step 3

Return Step 1 until no improvement of current route T is obtained.

A 2-opt move is deleting two existing edges and reconnecting the two pairs of nodes to obtain a new route. This operation is shown below (see Fig. 6).

Figure 6. 2-opt move

Source: own elaboration.

2-opt algorithm is an iterative procedure, which quality depends on the selected (constructed) starting solution (route). Starting solution may be generated randomly. It may be also the final outcome of another algorithm.

Computational results

The performance of three presented simple heuristic algorithms was tested with four euclidean problems. Each of them contains a set of points represented by coordinates x and y , belonging to the interval $(1,100)$. Starting point has always coordinates $[1,1]$. (see Fig.7)

Figure 7. TSP problem with 25 points (TSP_25)

Source: own elaboration.

Table 2. Coordinates for TSP_25

	x	y
P ₀	1	1
P ₁	13	71
P ₂	71	36
P ₃	1	33
P ₄	90	70
P ₅	42	11
P ₆	85	95
P ₇	91	13
P ₈	77	25
P ₁₃	5	72
P ₁₄	6	95
P ₁₅	40	84
P ₁₆	64	65
P ₁₇	50	81
P ₁₈	36	48
P ₁₉	90	58
P ₂₀	28	15
P ₂₁	99	46

P_9	99	36
P_{10}	22	9
P_{11}	84	8
P_{12}	34	42

P_{22}	22	78
P_{23}	7	82
P_{24}	45	26
P_{25}	98	66

Source: own elaboration.

All TSP problems were solved by the nearest neighbor algorithm (NNA), nearest insertion algorithm (NIA) and 2-opt algorithm (2OPT). In the figures below there are shown routes obtained by each of three algorithms (see Fig. 8, Fig. 9 and Fig. 10).

Figure 8. NNA solution for TSP_25. Tour length: 557,05

Source: own elaboration.

Figure 9. NIA solution for TSP_25. Tour length: 489,09

Source: own elaboration.

Figure 10 2OPT solution for TSP_25. Tour length: 470,29

Source: own elaboration.

The best solution was generated by 2-opt algorithm. The tour length is equal 470,29 and it is about 4% less than NIA's tour length and 15,5% less than NNA's tour length.

As has been told before, 2-opt algorithm may generate different quality solutions according to starting solution. Route shown in Fig. 10 is the best obtained solution for 12 runs 2-opt algorithm. 10 runs had randomly generated starting solution. One run 2OPT started from the solution generated by NNA and one from the route constructed by NIA. The results are shown in the table below (see Tab. 3).

Table 3. Results for 12 runs of 2OPT for TSP_25

Run	2OPT tour length	Number of improvements
1	470,29	50
2	471,97	59
3	471,97	37
4	504,61	56
5	481,72	55
6	526,17	58
7	470,29	59
8	470,29	62
9	470,29	67
10	481,72	60
NNA starting solution	512,75	7
NIA starting solution	470,29	2

Source: own elaboration.

The best tour length 470,29 was obtained 4 times. Average tour length for 10 runs is equal 481,93 with standard deviation 18,89, what is better result than NNA's and NIA's solution. The 2-opt algorithm needs average 56 improvements with standard deviation 8, but only a few improvements if it starts from the route constructed by the one of other heuristics.

In table below there are shown computational results for all tested TSP problems (see Tab. 4).

Table 4. Performance of the selected algorithms for tested problems

TSP problem	Number of points	NNA	NIA	2OPT ^{BEST}	2OPT ^{AVG}
TSP_25	25	557,05	489,09	470,29	481,93±18,89
TSP_50	50	864,80	631,20	625,15	638,87±10,99
TSP_75	75	945,41	771,30	754,67	769,57±16,29
TSP_100	100	890,30	838,70	779,84	812,71±23,03

Source: own elaboration.

Analyzing the results shown in the table above it can be concluded, that solution quality of nearest neighbor algorithm is very poor compared to 2-opt results. The difference between nearest insertion algorithm and 2-opt is not so evident, but it should be noted, that 2-opt procedure is less complicated than NIA. From the other side, NNA and NIA needs less time to generate solution than 2-opt.

Conclusions

In this paper the traveling salesman problem and its solution strategies was described. Construction and improvement strategies are two main approaches to solve TSP. It cannot be clearly concluded which strategy is better. Each has its advantages and disadvantages.

Considering the time of the algorithm execution, presented construction strategy methods run faster than improvement procedures. Improvement algorithms need more time to generate solutions and every time this it may be different. The reason is first of all quality of the starting solution.

Moreover, improvement algorithms not guarantee the same quality of achieved solution in contrast to the construction methods, what can be considered as a significant disadvantage. However, average quality of the solutions seems to be a little better than those obtained by construction algorithms.

It should also be noted, that simplicity and speed of both types of algorithms has another undeniable advantage. There is possibility to combine the advantages of both methods by finding good solution by NIA algorithm and then improving using 2-opt algorithm. Use of not one but several simple algo-

rithms may be a some kind of good proposal for solving hard decision problems which is TSP. It can be a good alternative for analysts or decision makers for which equally important is: simplicity of the algorithms, time required to generate good solution and quality of obtained solution.

References

- Całczyński A. (1992), *Metody optymalizacji w obsłudze transportowej rynku*, PWE, Warszawa.
- Całczyński A., Kędzierska-Stróż D., Orzechowska D., Śleszyński Z. (2000), *Elementy badań operacyjnych w zarządzaniu*, Wydawnictwo Politechniki Radomskiej, Radom.
- Ignasiak E. (1996), *Badania operacyjne*, PWE, Warszawa.
- Junhert M., Reinelt G., Rinaldi G. (1995), *The travelling salesman problem* in: M.O. Ball et al. (Eds), *Handbook in OR&MS*, Vol. 7, Elsevier.
- Krawczyk S. (2001), *Metody ilościowe w logistyce (przedsiębiorstwa)*, Academia Oeconomica, Warszawa.
- Lin S., Kernighan B.W. (1973), *An effective heuristic for the traveling salesman problem*, "Operations Research", Vol. 21, pp. 498–516.
- Little J.D.C., Murty K.G., Sweeney D.W., Karel C. (1963), *An algorithm for the traveling salesman problem*, "Operations Research", Vol. 11.
- Solich R. (1974), *Pewne uogólnienia zadania komiwojażera*, Prace Centrum Obliczeniowego PAN, Nr 14.
- Sysło M.M. (1984), *Optymalizacja kombinatoryczna* [w:] T. Kasprzak (red), *Optymalizacja dyskretna. Zastosowania ekonomiczne*, PWE, Warszawa.
- Szymanowski W. (red.) (1977), *Optymalizacja dostaw do sieci punktów sprzedaży z uwzględnieniem warunków ich działalności operatywnej*, Uniwersytet Warszawski, Wydział Nauk Ekonomicznych.
- Szymanowski W. (red.) (1979), *Optymalizacja dostaw do miejskiej sieci detalicznej*, Praca zbiorowa, Instytut Handlu Wewnętrznego i Rynku, Warszawa.