

Dominika Crnjac Milić¹

Ljiljanka Kvesić²

SOME NEW RESULTS ON THE TRAVELLING SALESMAN PROBLEM

Abstract:

The travelling salesman problem (or The sales representative problem) has been insufficiently explored so far. One of the first results on this issue was provided by Euler in 1759 (The problem of moving a knight on the chess board), Knight's Tour Problem. Papers on this subject were written by A.T. Vandermonde (1771), T. P. Kirkman (1856) and many others. The sales representative problem is a major challenge due to the application in solving theoretical and practical problems such as the quality of algorithms and of optimization methods. This well-known optimization problem has been extensively studied from several aspects since 1930. In general form the study was started by Karl Menger, seeking the shortest route through all points of a finite set with known distances between every two

points. Since then, there have been many formulations of the problem.

In this paper we shall provide an analysis of the nature of the commercial representative problem, and highlight its complexity and some ways of its solution. We shall use graph theory, and pay particular attention to the search of Hamiltonian cycle of minimum weight in the weighted graph.

During the paper development we were led by the following question: "How to minimize the total distance travelled by a sales representative in order to visit n given locations exactly once and return to the starting point?"

Keywords:

graph, top, edge, trail, algorithm, cycle

Author's data:

¹ Dominika Crnjac Milić, Associate Professor, Faculty of Electrical Engineering, J. J Strossmayer University of Osijek, Kneza Trpimira 2B, 31000 Osijek, Croatia,
E-mail: dominika.crnjac@etfos.hr;

² Ljiljanka Kvesić, Assistant Professor, Faculty of Science and Education, University of Mostar, 88000 Mostar, Bosnia and Herzegovina,
E-mail: ljkkvesic@gmail.com

Introduction

Although the problem of sales representative is a well-known optimization problem and so far it has been studied from several aspects [6], in this research paper, using graph theory, we are going to give our attention to the search of Hamiltonian cycle of minimum weight in the weighted graph. Looking at the problem from an economic aspect, the target of the companies in selling business is to animate potential customers within as short as possible time, on a daily basis and with the least possible costs, to buy their products. For planning daily activities of their sales representatives, it is important to minimize the total distance that a representative should pass in order to visit n given locations exactly once and return to the starting point.

The number of locations to visit in one period of time is known, the distance between the locations visited is given (constant).

Costs required for tours, and time lapse of each tour are inclined to a constant.

Suppose that a sales representative can pass only once through one place.

The problem can be mathematically formulated as follows:

$$x_{ij} = \begin{cases} 1, & \text{using path } i, j \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where x_{ij} are decision variables (the j -th job is performed or not performed in the i -th position), and x_{ij} are path costs (or time) spent on the i -th site of the client tour (i, j) .

Target Function is

$$\min z = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (2)$$

with restrictions

$$\sum_{i=1}^n x_{ij} = 1, \quad \forall j = 1, 2, \dots, n \quad (3)$$

$$\sum_{j=1}^n x_{ij} = 1, \quad \forall i = 1, 2, \dots, n \quad (4)$$

$$x_{ij} \geq 0,$$

$$x_{j_1 j_2} + x_{j_2 j_3} + \dots + x_{j_m j_1} \leq m - 1, \quad m = 2, \dots, n - 1 \quad (5)$$

Omitting the last restriction, we obtain an assignment problem [5].

Definitions and Basic Concepts

Definition

A graph is an ordered triplet $G = (V, E, \varphi)$, where $V = V(G)$ is a non-empty set whose elements are called vertices, $E = E(G)$ is a set whose elements are called edges and it is disjointive with V , and φ is a function which connects (u, v) to each edge e from E , where $u, v \in V$.

Furthermore, the vertices u, v are said to be adjacent if there is an edge e whose ends are u and v . The edge e is incident to vertices u and v with labeled $e = (u, v)$ or $e = u \cdot v$. The degree of vertex v in the graph is the number of graph G edges incidental with v . A walk is the sequence $W \equiv v_0 e_1 v_1 e_2 \dots v_k e_k$ whose members are alternately vertices v_i and edges e_i in the way that the ends of e_i are vertices v_{i-1} i v_i , $i = 1, 2, \dots, k$.

The number k is said to be a walk length of sequence W , wherein v_0 is the beginning and v_k the end of the walk W .

A walk is closed when $v_0 = v_k$.

If all edges are mutually different, a walk is called a path. If all the vertices are mutually different, the path is called a route.

A cycle is closed route whose vertices, except the end vertices, are different from each other.

Note that paths and cycles containing all the vertices of the graph are interesting for our problem in the study.

Definition 2

For a path (cycle) that contains all the graph vertices is said to be Hamiltonian path (cycle) in the graph.

One of the most difficult algorithmic problems is to answer the question whether the graph has got Hamiltonian cycle or not, especially Hamiltonian cycle of minimum weight [10].

Theoretically, in the final graph it is possible to find a Hamiltonian cycle in a finite number of steps. We are interested in quick and efficient algorithms.

Therefore, the interval of algorithm execution is important.

Time algorithm essentiality is observed using the function $f: \mathbb{N} \rightarrow \mathbb{R}^+$ (\mathbb{N} is the set of natural numbers), where $f(n), n \in \mathbb{N}$ is the number of elementary operations that are necessary in algorithm.

The time course of the algorithm is measured by the total number of operations, such as arithmetic operations, comparison, etc.

Furthermore, if $f: \mathbb{N} \rightarrow \mathbb{R}^+$ is the function, it is said that $f(n) = O(q(n))$ if there are $c, n_0 > 0$ such, that for each $n \geq n_0$, $f(n) \leq cq(n)$, $q: \mathbb{N} \rightarrow \mathbb{R}^+$ and for $q(n)$ it is said to be upper limit for $f(n)$.

In the event that $f(n) = O(q(n))$ than algorithm has time complexity $O(q(n))$. Consequent to q , the most frequently used algorithms are polynomial algorithms and exponential algorithms.

An algorithm is polynomial if there is a polynomial solution for the algorithm. Time for

solving algorithm is increased very slowly compared to the input data. It is uncertain that there is polynomial algorithm for the solution of many problems.

Definition 3

It is said that the ordered pair (G, w) is weighted, wherein $G = (V, E)$ is a graph and $w: E \rightarrow \mathbb{R}^+$ is a function, \mathbb{R}^+ (non-negative real numbers).

The number $w(e)$ is called the edge weight e .

Note that the weight of a given edge could mean any measure which characterizes an edge in economic terms, such as cost, profit, route length, etc.

Some mathematical models

a) *The transport mode* has a very important role in the management of supply chains.

The task is to minimize the total transportation costs and improve service.

Suppose there are m storage areas of a_i capacities at i locations for goods to be transported to j locations that have a demand for the goods b_j .

The task is to minimize transport costs between i place and j place if the unit cost of transport between the two destinations c_{ij} are known.

The quantity to be transported from place i to place j is x_{ij} .

Mathematical formulation can be written in the form

$$\min T = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}, \quad (6)$$

wherein T is the function representing the total cost, x_{ij} the quantity to be transported, and c_{ij} the unit cost of transportation.

It is necessary to find the minimum of a T function, subject to certain restrictions:

$$\sum_{j=1}^n x_{ij} = a_i, \quad \forall i = 1, 2, 3, \dots, m \quad (7)$$

$$\sum_{i=1}^m x_{ij} = b_j, \quad \forall j = 1, 2, 3, \dots, n \quad (8)$$

$$x_{ij} \geq 0, \quad \text{za } i = 1, 2, \dots, m, j = 1, 2, \dots, n \quad (9)$$

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j \quad (10)$$

b) *Assignment model* is a special case of transport model.

The core of the problem is in the distribution or in the assignment of n tasks and duties to n locations, people, etc., subject to correspondence. In other words, one job is assigned to only one employee, etc.

Each position can perform some or all of n possible tasks for a certain time (with certain costs). It is necessary allocate tasks such that every position performs only one job and the total time (or the total costs) required for the performance of all operations is minimal.

Mathematically, it is an injective projection [4]. The goal is to find the optimum using a measure of individual success. Of course, this can be applied to many management issues in economic practice. The aforementioned can be mathematically formulated as follows.

A square matrix of $A_{n \times n}$ type is given with elements $a_{ij} \geq 0$, for $i, j = 1, 2, \dots, n$, ($n \geq 3$).

It is necessary to determine the square matrix $X_{n \times n}$ with elements x_{ij} subject to

$$\sum_{i=1}^n x_{ij} = \sum_{j=1}^n a_{ij} = 1 \quad (11)$$

$$\min T = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_{ij} \quad (12)$$

Note that it is useful to define binary variables.

$$x_{ij} = \begin{cases} 1, & \text{means that } i \text{ applicant should be assigned to } j \text{ job} \\ 0, & i \text{ applicant is not assigned to } j \text{ job} \end{cases} \quad (13)$$

There are several different methods for solving this method, out of which we emphasise the method of transforming the above model in the corresponding network model whose solution is reduced to the problem of determining the shortest path in the network [5].

Three subtypes of the sales representative problem usually appear in practice:

1) Symmetric sales representative problem, which was previously described in this paper, and where there is an undirected weighted graph.

2) Asymmetric sales representative problem.

If there is at least one pair of places $(u, v) \in E(G)$ the path length (edge weight) has unequal values, depending on the direction of the tour; then this is an asymmetric sales representative problem. In this case, a directed graph is used as a model. Directed graph is usually called digraph and recorded as ordered triplet $D = (V, A, \psi)$, wherein V is non-empty set of vertices, namely $V \neq \emptyset$.

Cross section of A and V sets is empty, namely A and V are disjunctive. The elements of A set are called arcs, wherein a function ψ joins an ordered pair of vertices (u, v) , $u, v \in V$ to each arc $a \in A$ [7].

We can generalize the aforementioned to more sales representatives.

Suppose that m number of sales representatives departs from the same starting place, visit a set of places and return to where they started. It is necessary to determine the tours for all sales representatives, so that each place is visited only once with minimal total cost.

Note that the cost can be caused by the distance of places, by the time spent travelling, and by the

cost of transportation. For the application purpose we will list some variations of multiple travelling salesman problem, but they won't be analyzed in this paper because of their extensiveness.

1. Let a certain number of representatives start off from each of few starting places. After the tour the sales representatives return, either each of them in their starting place or in any of the starting places on the condition that in the end the number of sales representatives in every starting place is equal as in the beginning.
2. Suppose that the number of sales representatives is not fixed. Out of the available m sales representatives, we desire to make a selection of sales representatives to participate in the tour. Clearly, each sales representative has her/his own fixed costs that are taken into consideration when deciding how many sales representatives to be activated in order to minimize the total cost.
3. Suppose that there is a demand for a sales representative to visit some places in a given time intervals. This problem occurs in the organization of air transport, maritime transport, transport by roads and the like.
4. It is possible to introduce various restrictions, such as a limited number of places that a sales representative should visit, minimum distance, maximum distance etc.

Some methods of solving the travelling salesman problem

1. Exact methods

These are the methods that give exact algorithms. Their disadvantage is a prolonged performance time.

In this place, we might add the method of branching and restrictions, where you can estimate all possible solutions and reject the adverse ones on the basis of pre-set certain criteria.

2. Approximate methods

Algorithms that provide approximate solutions are used here. These are lower time complexity algorithms.

- a) Nearest neighbour method

This method is highly developed and most simple approximate method. It is about visiting the next nearest neighbouring places which were not visited.

When places are visited, it is necessary to return to the starting place.

- b) Greedy algorithms

The route is always built by adding the shortest possible edge. During this process, neither a cycle whose length is less than the number of places, nor a vertex with degree higher than 2 may occur, ie. a place may not be visited more than once, and the same edge may not be added multiple times.

In addressing the above concerns, the question is how to find all the possible tours, compute their lengths, and choose the best of them.

In order to simplify the solution of travelling salesman problem, many approximate methods are developed today and continue to develop, which will provide acceptable solutions. Effective methods have been developed so far, which enable solving the problem for a couple of million places within quite reasonable amount of time. [2]

- a) Method of inserting

We start with the shortest tour of n given places.

Most usually it is a triangle.

A follower of the previous place is an added place.

It is closest to any of the preceding visits and is added to the optimum position in the tour.

The process is repeated until all the places are added.

- b) Optimization method by ant colonies
Scientists often solve complex problems by watching and imitating natural processes. By studying and imitating the movement of ant colonies, we find solutions to the problems of a small number of places.

Namely, during the search for new areas ants leave a trail of pheromones that leads the other ants to new places (places of food). Let's observe a group of ants in different places. They do not return to the places where they were, and they visit all the remaining places.

The ant that uses the shortest route leaves the most intense pheromone trail, inversely proportional to the length of path. Naturally, the rest of ants will follow the peak intensity of the pheromone and follow its trail. The procedure is repeated until the shortest tour [10].

To improve previously studied approximate algorithms, and to find optimal algorithms, the following ideas are useful:

2- optimal algorithm: We arbitrary choose two edges in the cycle, which we remove and join two newly-created paths. Connecting is done so that visiting conditions are maintained. The tour is still used if it is shorter than the transitional. The process is continued until the impossibility of improvement.

3-optimal algorithm: We arbitrarily choose 3 edges and remove them. Two ways of reconnecting occur by

removing those three edges. We choose a path that has a shorter tour. Note that the 3-optimal tour is the two-optimal one.

k -optimal algorithm: This algorithm improves as the previous two do, but the work is done with k edges, $k > 3$.

Of course, the greater k requires more computational time.

One result of the travelling salesman problem optimization

To find the shortest path you need to find the path of least weight that connects two given vertices u_0 and v_0 . For the purpose of simplicity, instead of path weight p let's introduce the term path length p , wherein $p = \sum_{e \in p} w(e)$, and the least path weight (u, v) is a distance from w to v , which is written $d(u, v)$.

In order to find the shortest path, we provide the following Algorithm which finds the shortest path (u_0, v_0) ; what is more, all the shortest paths from u_0 to all other vertices in G .

Suppose that $S \subseteq V$, so that $w_0 \in S$, $\bar{S} = V \setminus S$. If $p = w_0 \dots \bar{w}v$ is the shortest path from u_0 to \bar{S} , then $\bar{u} \in S$ and a (u_0, \bar{u}) part of p has to be the shortest (u_0, \bar{u}) path.

Out of this

$$d(u_0, \bar{v}) = d(u_0, \bar{u}) + w(\bar{u}\bar{v}), \quad (14)$$

distance from u_0 to \bar{S} is

$$d(u_0, \bar{S}) = \min_{u \in S, v \in \bar{S}} \{d(u_0, u) + w(uv)\}. \quad (15)$$

Let's start from the set $S_0 = \{u_0\}$ and construct an increasing range of subsets out of

$V, S_0, S_1, \dots, S_{v-1}$, so that in the end of i -th step, the shortest path from u_0 to all vertices out of S_i are known. The first step is finding a vertex which is closest to vertex u_0 , which is obtained by computing $d(u_0, \bar{S}_0)$ and selecting vertex $u_1 \in \bar{S}_0$ so that

$$d(u_0, u_1) = d(u_0, \bar{S}_0), \quad (16)$$

as follow out of (14) results that

$$d(u_0, \bar{S}_0) = \min_{u \in \bar{S}_0, v \in S_0} \{d(u_0, u) + w(uv)\} = \min_{v \in S_0} \{w(u_0v)\} \quad (17)$$

In addition, suppose $S_1 = \{u_0, u_1\}$, p_1 - path u_0v_1 , so it is the shortest (u_0, u_1) - path.

For the purpose of generalization, suppose that $S_k = \{u_0, u_1, \dots, u_k\}$ and suppose that the shortest paths $(u_0, u_k), p_1, \dots, p_k$ have already been determined, then with (15) we can compute $d(u_0, \bar{S}_k)$ and select vertex $u_{k+1} \in \bar{S}_k$ such that

$$d(u_0, u_{k+1}) = d(u_0, \bar{S}_k). \quad (18)$$

Notice that according to (15)

$$d(u_0, u_{k+1}) = d(u_0, u_j) + w(u_j u_{k+1}), \quad (19)$$

for some $j \leq k$, and so we obtain the shortest (u_0, u_{k+1}) - path by adding edge u_j, u_{k+1} to path p_j .

Note that in each step these shortest paths together form a connected graph without cycles. These graphs are called trees (wood) and the previous algorithm is called the process of tree growth [3].

The idea for the previous algorithm originates from Edsger Dijkstra Wybe (1959) who, by

means of his algorithm, managed to determine the distance of individual points (destinations) to other points considered to be important, but he did not manage to determine the shortest distance in this way. Knowing that the task of a sales representative is to visit some business destinations and return after the job, provided that each destination was visited exactly once, the question is how to make the itinerary, and to travel as short as possible?

According to the aforementioned (the shortest path problem) Hamiltonian cycle of minimum weight, which is called the optimal cycle should be found in the complete weighted graph. Here we will provide an "approximate" approach, consisting of finding a Hamiltonian cycle, and then search for another cycle of less weight that is slightly modified.

If $c = v_1v_2 \dots v_{v-1}v_1$, then, for all $i, j, 1 < i+1 < j$ we can find a new Hamiltonian cycle

$$c_{ij} = v_1v_2 \dots v_i v_j v_{j-1} \dots v_{i+1} v_{j+1} v_{j+2} \dots v_{v-1} v_1 \quad (20)$$

where we removed edges $v_i v_{i+1}, v_j v_{j+1}$, and added edges $v_i v_j$ and $v_{i+1} v_{j+1}$.

Furthermore, if for some i, j , we can apply

$$w(v_i v_j) + w(v_{i+1} v_{j+1}) < w(v_i v_{i+1}) + w(v_j v_{j+1}), \quad (21)$$

then the cycle c_{ij} is an improvement of c .

By continuing likewise we shall reach the cycle that cannot be improved anymore by this method. It is clear that the final cycle is not optimal. In order to achieve greater accuracy, the procedure can be repeated several times, starting with different cycles.

Conclusion

This paper is a result of analytical research of the theoretical basis of the travelling salesman problem.

Numerous formulations of the travelling salesman problem are known in the literature, and most of them remained unresolved to date. Here, as confirmation of prior thought we mention an open problem of graph theory: Finding necessary and sufficient condition for a graph to have a Hamiltonian cycle. In this research paper, we pointed out the variations of the travelling salesman problem. Some solving methods were pointed out, and certain improvements were provided.

The paper gives an approximate result of travelling salesman problem optimization by using a Hamiltonian cycle with the aim of contributing to solving the travelling salesman problem, which, to this day and with great effort of scientists, has not yet been resolved.

References

- [1] D. L. Applegate (2006). *The Traveling Salesman Problem: A Computational Study*, Princeton University Press, ISBN 978-0-691-12993-8, Princeton, New Jersey, pp.593
- [2] T.Bektas (2005). The multiple traveling salesman problem: an overview of formulations and solution procedures, *The International Journal of Management Science Omega*, Volume 34, Issue 3, June 2006, pp. 209-219, ISSN 0305-0483, The University of Michigan
- [3] D.Crnjac Milić (2009). Stablo (drvo) logičkih mogućnosti i njegova primjena u logistici, *IX međunarodni znanstveni skup Business logistics in modern management*, Segetlija Z., Karić M. (Ed.), pp. 265-270, ISBN 978-953-253-072-8, Ekonomski fakultet u Osijeku, 11.2009., Sveučilište Josipa Jurja Strossmayera u Osijeku, Osijek
- [4] M.Crnjac, D.Jukić, R.Scitovski (1994). *Matematika*, Ekonomski fakultet u Osijeku, ISBN:953-6073-03-X, Osijek
- [5] L. Čaklović (2010). *Geometrija linearnog programiranja*, Element, ISBN:978-953-197-566-7, Zagreb
- [6] Rajesh Matai, Surya Singh and Murari Lal Mittal (2010). *Traveling Salesman Problem: an Overview of Applications, Formulations, and Solution Approaches, Traveling Salesman Problem, Theory and Applications*, ISBN: 978-953-307-426-9, InTech
- [7] G. Sierksma (2002). *Linear and Integer Programming. Theory and Practice*, 2nd Edition, Marcel Dekker, New York,
- [8] M. Singer (1997). *Introduction to the Theory of Computation*, PWS Pub.co., Boston
- [9] M.Sipser (2012). *Introduction to the Theory of Computation*, Thomson Course Technology, (Ed.: 3rd), ISBN 978-1133187790, Boston, Massachusetts, United States of America
- [10] D. Veljan (2001). *Kombinatorna i diskretna matematika*, Algoritam, ISBN 86-03-99051-4, Zagreb
- [11] H.S.Wilf (1986). *Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs
- [12] S.G. Williamson (1985). *Combinatorics for Computer Science*, Computer Science Press, ISBN 978-0486420769, Rockville