Global Journal of Business, Economics and Management

Volume 06, Issue 2, (2016) 70-77

www.wjbem.eu

New formulations for the team orienteering problem

Tusan Derya[*]**,** Faculty of Engineering, Department of Industrial Engineering, Baskent University, 06790 Ankara, Turkey.

Imdat Kara, Faculty of Engineering, Department of Industrial Engineering, Baskent University, 06790 Ankara, Turkey.

Papatya Sevgin Bicakci, Faculty of Economics and Administrative Sciences, Department of Management, Baskent University, 06790 Ankara, Turkey.

Baris Kececi, Faculty of Engineering, Department of Industrial Engineering, Baskent University, 06790 Ankara, Turkey.

Suggested Citation:

Derya, T., Kara, I., Bicakci, P., S. & Kececi, B. (2016). New formulations for the team orienteering problem. *Global Journal of Business, Economics and Management: Current Issues*. *6*(2), 70-77.

Received July 08, 2016 ; revised August 12, 2016; accepted November 09, 2016;

Selection and peer review under responsibility of Prof. Dr. Andreea Iluzia IACOB, Bucharest Academy of Economic Studies, Romania.

Abstract

Routing problems have many practical applications in distribution and logistics management. The Traveling Salesman Problem (TSP) and its variants lie at the heart of routing problems. The Orienteering Problem (OP) is a subset selection version of well-known TSP which comes from an outdoor sport played on mountains. In the OP, the traveller must finish its journey within a predetermined time (cost, distance), and gets a gain (profit, reward) from the visited nodes. The objective is to maximize the total gain that the traveller collects during the predetermined time. The OP is also named as the selective TSP since not all cities have to be visited. The Team Orienteering Problem (TOP) is the extension of OP by multiple-traveller. As far as we know, there exist a few formulations for the TOP. In this paper we present two new integer linear programming formulations (ILPFs) for the TOP with $O(n^2)$ binary variables and $O(n^2)$ constraints, where n is the number of nodes on the underlying graph. The proposed formulations can be directly used for the OP when we take the number of traveller as one. We demonstrate that, additional restrictions and/or side conditions can be easily imported for both of the formulations. The performance of our formulations is tested on the benchmark instances from the literature. The benchmark instances are solved via CPLEX 12.6 by using the proposed and existing formulations. The computational experiments demonstrate that both of the new formulations outperform the existing one. The new formulations are capable of solving optimally most of the benchmark instances, which have solved by using special heuristics so far. As a result, the proposed formulations can be used to find the optimal solution of small- and moderate-size real life OP and TOP by using an optimizer.

Keywords: traveling salesman problem, orienteering problem, modeling.

* ADDRESS FOR CORRESPONDENCE: **Tusan Derya,** Faculty of Engineering, Department of Industrial Engineering, Baskent University, Ankara, 06530, Turkey. *E-mail address*: tderya@baskent.edu.tr / Tel.: +90-312-246-6666

1. Introduction

Orienteering is the name of outdoor sport played on mountains and forests. In orienteering sport, the checkpoints with a particular score value are visited by the player within a time constraint that is not too enough to visit all checkpoints. The player scores a point from each visited checkpoint. The aim of this sport activity is to collect the highest score by visiting as more checkpoints as possible in the time allowed. Orienteering can be performed by individuals and teams as well. Each player in the team tries to maximize the total score points by visiting the checkpoints within a specific time limit. Players move from starting point and must reach the destination point on time. Checkpoints should not be visited by more than one player since each checkpoint contributes score once. To do so the route of each player must be determined.

Tsiligirides (1984) defines the Orienteering Problem (OP) inspired by the Orienteering sport. The OP is to find the route which maximizes the total score point by visiting as more nodes (customers) as possible under the cost (time or distance) restrictions. Each node must be visited at most once. The origin and destination points may or may not be same. Because there is no obligation to visit all the nodes in the OP, it is also called as Selective Traveling Salesman Problem (Laporte & Martello, 1990). The OP constitutes node selection problem and the shortest Hamiltonian path problem among the selected nodes. To determine the shortest Hamiltonian path among the selected nodes may be beneficial in terms of visiting more nodes within a time limit. Therefore, OP can be thought of the composition of Knapsack Problem and Travelling Salesman Problem (TSP).

The most important variant of the OP, where more than one traveller exists, is called Team Orienteering Problem (TOP). In some aspects the TOP resembles the Vehicle Routing Problem (VRP) which is commonly known in the operations research literature. VRP is defined as finding the optimal routes with minimum cost while satisfying all the customer demands from a central depot. Unlike the VRP, in the TOP (*i*) it is not possible to visit all the customers because of the cost (time and distance) restrictions; (*ii*) there is no capacity constraint for the vehicles; (*iii*) it is preferred to maximize the return such as profit and income rather than to minimize the cost. The real life applications for the TOP are; excursion in the tourism sector, commodity and service transportation systems, school bus routing.

Golden, Levy & Vohra (1987) proved that the OP is NP-Hard. TOP is an OP where there is one traveller. Hence the TOP is also NP-Hard. Because of this reason usually heuristic algorithms are preferred to solve the TOPs. Interested readers may look at the Vansteenwegen, Souffriau & Oudheusden (2011)'s paper in detail for the solution approaches proposed for the TOP.

The TOP is first studied by Butt and Cavalier (1994) with the name of Multiple Tour Maximum Collection Problem. Later on it is named as TOP by Chao, Golden & Vasil (1996). Spectacular studies are published on the exact solution approaches to find the optimal solution of TOP. The first exact algorithm for the TOP is proposed by Butt and Ryan (1999). Butt and Ryan (1999) use the column generation approach together with the branch-and-bound algorithm. Additionally, Boussier, Feillet & Gendreau (2007) propose a branch-and-price algorithm. The proposed algorithm can find optimal solutions up to 100 nodes. Poggi, Viana and Uchoa (2010) propose a branch-and-cut algorithm and branch-and-price algorithm for the TOP.

Butt and Cavalier (1994), Tang and Miller-Hooks (2005), Ke, Archetti & Feng (2008), Vansteenwegen, Souffriau & Oudheusden (2011), Dang, Guibadj & Moukrim (2013) propose mixed integer programming formulations for the TOP. The subtour elimination constraints of mathematical formulations in the literature, except Vansteenwegen et. al. (2011), increases exponentially with respect to the number of nodes in the underlying graph. Therefore these formulations cannot be used directly by any optimization software. Vansteenwegen et al. (2011) present a general TOP formulation whose sub-tour elimination constraints increase polynomially. As far as we are aware, this is the only formulation that can be used directly by an optimizer.

The main motivation of this paper is to present new linear programming formulations for solving small- and moderate-size TOPs directly by using any optimizer. Our contributions are two folds: (1) We present new ILPFs for the TOP with $O(n^2)$ constraints and $O(n^2)$ binary decision variables. Both formulations are useable directly by any optimizer. (2) We found the optimal solutions of the benchmark instances and observe that most of the best known solutions obtained by heuristics are far away from the optimal values.

The remainder of the paper is organized as follows: In section 2, problem TOP is defined more precisely and general formulation of the TOP is given, while two new ILPFs for the TOP are presented in section 3. The performances of the proposed formulations are analysed in section 4. The conclusion and further remarks appeared in section 5.

2. Problem Definition and General Formulation

The Team Orienteering Problem (TOP) can be defined as follows: Let $G = (V, A)$ be a complete graph, where $V=\{1, 2, \ldots, n\}$ is the set of nodes (vertices), $\{1\}$ is the departure node (depot, origin), $\{n\}$ is the arrival node (and might actually correspond to a same physical location) and the remaining nodes are customer nodes. The set $A=\{(i, j): i, j\in V, i\neq j\}$ is the arc (or edge) set. Each node i is associated with a profit (reward) p_i ($p_1 = 0$ and $p_n = 0$). The traveller gains a profit p_i if the i^{th} node is visited. A travel cost (time, distance) t_{ij} is associated with each arc $(i, j)\in A$. The travel costs are assumed to satisfy the triangular inequality. There are m travellers located at the depot. The journey starts from the departure node and ends at the arrival node, and must complete within a predetermined time (or cost or distance) T_{max}. Some nodes may not be visited because of T_{max} restriction. The TOP consists of determining a set of m paths, each goes from node 1 to node n and keeps to the time limitation; such that each customer is visited at most once and the total profit collected is maximized. No capacity constraints are considered.

The TOP can be formulated as an integer linear programming model (ILPF) by using the polynomial number of decision variables: $x_{ij}=1$ if the traveller goes from node i to node j and 0 otherwise. Then, a general two indexed ILPF for the TOP may be given as follows:

$$\text{Maximize} \sum_{i=1}^{n}\sum_{j=1}^{n-1} p_i x_{ij} \tag{1}$$

Subject to

$$\sum_{i=2}^{n} x_{1i} = m \tag{2}$$

$$\sum_{i=1}^{n-1} x_{in} = m \tag{3}$$

$$\sum_{i=1}^{n-1} x_{ij} \leq 1 \quad , j = 2, 3, \ldots, n-1 \tag{4}$$

$$\sum_{j=1}^{n} x_{ij} \leq 1 \quad , i = 2, 3, \ldots, n-1 \tag{5}$$

$$\sum_{i=1}^{n-1} x_{ij} = \sum_{i=2}^{n-1} x_{ji} \quad , j = 2, 3, \ldots, n-1 \tag{6}$$

+ Sub-tour elimination constraints (7)

+ Time constraints (T_{max}) (8)

$$x_{ij} \in \{0,1\} \quad , \quad \forall (i,j) \in A \tag{9}$$

In this formulation, the objective function (1) maximizes the total profit (reward) collected. Constraints (2) and (3) guarantee that there are *m* paths each starts from node 1 and ends at node *n*; whereas (4) and (5) are the degree constraints ensuring that each node (except node 1 and *n*) is visited at most once. Constraint (6) named as the conservation of flow constraints; guarantees the connectivity of each path. Constraint (7) is the implicit form of sub-tour elimination constraints. These constraints are necessary to prevent the sub-tours. Constraint (8) ensures the time limit (T_{max}) is not exceeded for each path. Constraint (9) indicates that the decision variables are binary.

Several ILPFs with three-index variables for the TOP have been proposed in the literature (see Butt and Cavalier, 1994; Tang and Miller-Hooks, 2005; Ke et. al., 2008, Vansteenwegen et. al., 2011; Dang et. al., 2013). Existing ILPFs of the TOP differ from each other with respect to the sub-tour elimination and time constraints proposed instead of the ones given in (7) and (8) above. These formulations except Vansteenwegen et. al. (2011), are based on Dantzig-Fulkerson-Johnson's sub-tour elimination constraints of the TSP (Dantzig, Fulkerson & Johnson, 1954). The number of inequalities in these types of constraints grows exponentially with the number of nodes, i.e., they are exponential size formulations. Thus these formulations become significantly larger as the problem size increases. Therefore, they cannot be used directly to solve instances of small or moderate sizes by any optimizer. To the best of our knowledge, there is only one polynomial size formulation with three-index variables, proposed by Vansteenwegen et. al. (2011) for the TOP, which is based on Miller-Tucker-Zemlin subtour elimination constraints (Miller, Tucker & Zemlin, 1960).

3. New Formulation

In this paper we focus only on the polynomial size formulations. In order to prevent illegal tours, i.e., eliminate sub-tours, auxiliary decision variables are defined and then sub-tour elimination constraints of the formulation are developed. Therefore, the proposed formulations may be divided into two groups according to the new auxiliary variables. A formulation is named as node-based if the auxiliary decision variables defined on the nodes of graph and as arc-based if the auxiliary decision variables are defined on the arcs of graph. In the subsequent sections, we present two new ILPFs, one of them is a node based and the other is an arc based formulation, for the TOP.

3.1. Node Based Formulation for the TOP

Let us define the auxiliary variable as in the following.

v_i: is the time passed up to the node *i* if a traveler visits on its journey and zero otherwise.

Proposition 1: The following inequalities together with the constraints (2) - (6) are valid initializing and bounding constraints of the auxiliary variables v_i's for the TOP. Constraints (10) and (11) are the time bounding constraints ensuring that $t_{1i} \le v_i \le T_{\max} - t_{in}$. Constraint (12) ensures that the journey must end before the predetermined value T_{max}, and also v_i is zero if the i^{th} node does not visited.

$$v_i \ge t_{1i}\, x_{1i} + \sum_{i-1} (t_{1j} + t_{ji}) x_{ji} \quad , \; i = 2,\dots, n \tag{10}$$

$$v_i \le T_{\max} - t_{in} - (T_{\max} - t_{1i} - t_{in}) x_{1i} \quad , \quad i = 2,\dots,n\text{-}1 \tag{11}$$

$$v_i \le T_{\max} \sum_{i-1} x_{ij} \quad , \quad i = 2,\dots,n\text{-}1 \tag{12}$$

Proposition 2: The following inequality together with the constraints (2) - (6) and (10) - (12) are valid sub-tour elimination constraints for the TOP. Those constraints guarantee that the solution contains no illegal sub-tours.

$$v_i - v_j + (T_{\max} + t_{ij})x_{ij} + (T_{\max} - t_{ji})x_{ji} \le T_{\max} \ , \ i \neq j = 2,\ldots, n-1 \qquad (13)$$

Constraint (13) is the route continuity and sub-tour elimination constraint ensuring that the auxiliary variable v_i of each tour works as a step function in accordance with the time between the consecutive nodes of tour, hence no illegal sub-tour can be formed.

We propose the following ILPF for the TOP with respect to the proposition 1 and 2. The formulation is called as *NBF* since the auxiliary decision variables defined on the nodes of graph.

NBF:

Maximize (1)

Subject to (2)-(6) and (9)-(13) and

$$v_i \ge 0 \ , \ i = 2,\ldots, n \qquad (14)$$

where $x_{ij} = 0$ whenever $t_{1i} + t_{jn} + t_{ij} > T_{max}$ and the distance matrix is Euclidean. The *NBF* given above has $O(n^2)$ binary decision variables and $O(n^2)$ constraints.

3.2. Arc Based Formulation for the TOP

Let us define the flow variable y_{ij} as in the following.

y_{ij}: is the total time passed from origin to node *j* if the arc (*i*, *j*) is on the path and zero otherwise.

Proposition 3: The following inequalities together with the constraints (2) - (6) are valid initializing and bounding constraints of the auxiliary variables y_{ij}'s for the TOP. Note that, if the arc (*i*, *j*) is not on the path then y_{ij} must be equal to zero. We do not need non-negativity constraints for y_{ij}'s since the constraint (16) will force this restriction.

$$y_{1i} = t_{1i}x_{1i} \ , \ i = 2, \ldots, n \qquad (15)$$

$$y_{ij} \ge (t_{1i} + t_{ij})x_{ij} \ , \ i = 2, \ldots, n-1 \ , \ j = 2, \ldots, n \qquad (16)$$

$$y_{ij} \le T_{\max}x_{ij} \ , \ i = 1, \ldots, n-1 \ , \ j = 2, \ldots, n \qquad (17)$$

Proposition 4: The following inequality together with the constraints (2) - (6) and (15) - (17) are valid sub-tour elimination constraints for the TOP. Those constraints guarantee that the trip ends within the predetermined time and the solution contains no illegal sub-tours.

$$\sum_{\substack{j=1 \\ j\neq i}} y_{ij} - \sum_{\substack{j=1 \\ j\neq i}} y_{ji} - \sum_{\substack{j=1 \\ j\neq i}} t_{ij}x_{ij} = 0 \ , \ i = 2, \ldots, n-1 \qquad (18)$$

In accordance with the constraints given in (15) – (18), the auxiliary variable yij of each arc on the path forms a step function with respect to the time between previous nodes. Thus, those constraints guarantee that, an auxiliary variable yij shows the time passed just after leaving the node j.

We propose the following ILPF for the TOP with respect to the proposition 3 and 4. The formulation is called as ABF since the auxiliary decision variables defined on the arcs of graph.

ABF:

Maximize (1)

Subject to (2)-(6) and (8) and (15)-(18)

where $x_{ij} = 0$ whenever $t_{1i} + t_{jn} + t_{ij} > T_{max}$ and the distance matrix is Euclidean. The *ABF* given above has $O(n^2)$ binary decision variables and $O(n^2)$ constraints.

4. Computational Analysis

In order to investigate the performance of proposed formulations, the results of models are compared on a set of test instances. Seven problem sets of Chao et al. (1996), including 387 benchmark instances in total, for the TOP are used. The test instances in each set have 32, 21, 33, 100, 66, 64 and 102 numbers of nodes, respectively. The coordinate and profit of each node is identical in all instances of the same set. In each set, there are three groups which have different numbers of vehicles. An instance in each group is characterized by a different value of T_{max}. All computational experiments are performed on a notebook PC with Intel Core Duo CPU @1.66 GHz processor and 2 GB RAM.

Instances are solved by two new formulations, *NBF* and *ABF*, presented in this paper and the formulation (*VF*) proposed by Vansteenwegen et al. (2011), as well. CPLEX 12.6 solver engine is used to solve the formulations. The solution time for each instance is limited with 7200 seconds. For each instance, the results found by solving our proposed formulations are compared with the results of existing *VF* formulation. The comparison of results is shown in Tables 1 and 2. Table 1 gives the comparison of formulations on the small-sized test instances up to 33 nodes. In Table 1 the first column shows the problem set, the second column shows the number of nodes (*n*), the third column shows the number of travellers (*m*) and the fourth column shows the number of test instances in the problem set. The next three columns indicate the number of optimally solved problems in each set by the model *VF*, *NBF* and *ABF*, respectively.

Table 1. The comparison of formulations on the small-sized test instances

				# of optimal solutions found		
Problem	*n*	*m*	#	*VF*	*NBF*	*ABF*
	32	2	18	16	9	18
Set 1	32	3	18	12	10	18
	32	4	18	16	15	18
	21	2	11	11	10	11
Set 2	21	3	11	11	11	11
	21	4	11	11	11	11
	33	2	20	18	4	20
Set 3	33	3	20	13	10	20
	33	4	20	14	13	20

According to the results in Table 1, ABF is able to find the optimal solutions of all problems in the first three problem sets within 7200 seconds. The VF and NBF can find the optimal solutions of 122 and 93 of 147 test problems, respectively. The results in Table 1 show that the ABF is outperforming to the other formulations. Hence, the ABF formulation is used to solve the medium-sized test problems. In Table 2 the first four columns are explained as in the Table 1. The next two columns indicate the number of optimally solved problems in the test instances and the average CPU time in seconds, respectively.

Table 2. The results of *ABF* formulation on the medium-sized test instances

Problem	*n*	*m*	#	# of optimal solutions	Average CPU (sec.)
	100	2	20	16	3081
Set 4	100	3	20	8	4640
	100	4	20	7	4790
	66	2	26	26	499
Set 5	66	3	26	24	4342
	66	4	26	24	4029
	64	2	14	14	1118
Set 6	64	3	14	14	1077
	64	4	14	14	716
	102	2	20	20	2049
Set 7	102	3	20	16	5050
	102	4	20	13	4755

Table 2 indicates that the formulation *ABF* is able to optimally solve the 196 of 240 instances within the 7200 seconds. The average CPU time to obtain the optimal solution by the *ABF* formulation is around 3000 seconds. And the average solution time differs between 500 and 5000 seconds approximately.

5. Computational Analysis

In this study, we address the team orienteering problem (TOP), where the objective function is to maximize the total profit that travelers collect from the visited nodes. For this case, two integer linear programming formulations (each has polynomial number of binary variables and constraints) are developed to solve the TOP instances optimally by a standard optimization solver. These two formulations are compared with a formulation already exists in the literature.

To investigate the performance of formulations over 650 computational experiments are done. According to the results obtained on the small-sized benchmark problems *ABF* is superior to the other formulations. And with the experimentation on the medium-sized problems, *ABF* is able to optimally solve the medium-sized instances with 7200 seconds time limit. As a consequence of these computational analyses, we conclude that *ABF* may be used for solving TOP directly by any optimizers.

References

Boussier, S. Feillet., D., & Gendreau, M. (2007). An exact algorithm for team orienteering problems. *4OR, 5*, 211 – 230.

Butt, S., & Cavalier, T. (1994). A heuristic for the multiple tour maximum collection problem. *Computers and Operations Research, 21*, 101 – 111.

Butt, S., & Ryan, D. (1999). An optimal solution procedure for the multiple tour maximum collection problem using column generations. *Computers and Operations Research, 26*, 427 - 441.

Chao, I., Golden, B., & Wasil, E. (1996). The team orienteering problem. *European Journal of Operational Research, 88*, 464 – 474.

Dang, D. C., Guibadj, R. N., & Moukrim, A. (2013). An effective pso-inspired algorithm for the team orienteering problem. *European Journal of Operational Research, 229*(2), 332 – 344.

Dantzing, G., Fulkerson, R., & Johnson, S. (1954). Solution of a large-scale traveling salesman problem. *Operations Research, 2*, 393 - 410.

Golden, B., Levy, L., & Vohra, R. (1987). The Orienteering Problem. *Naval Research Logistics, 34*, 307 - 318.

Ke, L., Archetti, C., & Feng, Z. (2008). Ants can solve the team orienteering problem. *Computers and Industrial Engineering, 54*, 648 – 665.

Laporte, G., & Martello, S. (1990). The Selective Travelling Salesman Problem. *Discrete Applied Mathematics, 26*, 193 - 207.

Miller, C. E., Tucker, A. W., & Zemlin, R.A. (1960). Integer Programming Formulations and Traveling Salesman Problems. *Journal of the Association for Computing Machinery, 7*, 326 - 329.

Poggi, M., Viana, H., & Uchoa, E. (2010). The Team Orienteering Problem: Formulations and Branch-Cut and Price. In *10th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS '10)*, 142 – 155.

Tang, H., & Miller-Hooks, E. (2005). A tabu search heuristic for the team orienteering problem. *Computers and Operations Research, 32*, 1379 - 1407.

Tsiligrides, T. (1984). Heuristic methods applied to orienteering. *Journal of the Operational Research Society, 35*, 797 - 809.

Vansteenwegen, P., Souffriau, W., & Oudheusden, D. V. (2011). The orienteering problem: A survey. *European Journal of Operational Research, 209*, 1 - 10.