manuscript No.
(will be inserted by the editor)

Approximating the asymmetric profitable tour

Viet Hung Nguyen[1], **Thi Thu Thuy Nguyen**[2]

[1] LIP6, Université Pierre et Marie Curie Paris 6, 4 place Jussieu, 75252 Paris Cedex 05, France
[2] Ecole Internationale des Sciences de Traitement de l'Information (EISTI), Avenue du Parc 95011, Cergy Cedex, France

Received: date / Revised version: date

Abstract We study the version of the asymmetric prize collecting traveling salesman problem, where the objective is to find a directed tour that visits a subset of vertices such that the length of the tour plus the sum of penalties associated with vertices not in the tour is as small as possible. In [3], the authors defined it as the *Profitable Tour Problem* (PTP). We present an $(1 + \log(n))$-approximation algorithm for the asymmetric PTP with n is the vertex number. The algorithm that is based on Frieze et al.'s heuristic for the asymmetric traveling salesman problem as well as a method to round fractional solutions of a linear programming relaxation to integers (feasible solution for the original problem), represents a directed version of the Bienstock et al.'s [2] algorithm for the symmetric PTP.

1 Introduction

Let $G = (V, A)$ be a complete directed graph with vertex set $V = \{1, 2, \dots, n\}$ and arc set A. We associate with each arc $e = (i, j)$ a cost c_e and with each vertex $i \in V$ a nonnegative penalty π_i. The arc costs are assumed to satisfy the triangle inequality, that is, $c_{(i,j)} \le c_{(i,k)} + c_{(k,j)}$ for all $i, j, k \in V$. In this paper, we consider the *Asymmetric Profitable Tour Problem* (APTP) which is a simplified version of the *Asymmetric Prize Collecting Traveling Salesman Problem* (APCTSP), namely, to find a tour that visits a subset of the vertices such that the length of the tour plus the sum of penalties of all vertices not in the tour is as small as possible.

In the general version of APCTSP, introduced by Balas [1], the arc costs are no assumed to satisfy the triangle inequality. Further, associated with each vertex there is a certain reward or prize, and in the optimization problem one must choose a subset of vertice to be visited so that the total reward is at least a given a parameter W_0.

The Profitable Tour Problem was defined formally in [3] with two versions: SPTP for *Symmetric Profitable Tour Problem* (i.e. when $c_{(i,j)} = c(j, i)$ for all $i, j \in V$) and APTP for the asymmetric one. In [3], the authors also note that two approximation algorithms have been developped for SPTP. The first one given by Bienstock et al. [2] achieves a factor $\frac{5}{2}$. This algorithm is based on the solution of a linear programming problem. The second approximation algorithm, developped by Goemans and Willamson [5], is purely combinatorial. They presented a general approximation technique for constrained forest problems, that can be extended to SPTP with $2 - \frac{1}{n-1}$ as approximation factor.

In [3], the authors also noted that no approximation algorithm had been designed for APTP and to our knowledge, no one has been developped from that time either.

In this paper, we propose a first approximation algorithm for APTP. The algorithm uses the same framework as the Bienstock et al.'s one which is based on the followings:

Result 1: The value of a solution given by Christofides heuristic for STSP is at most $\frac{3}{2}$ times the value of an optimal solution of the linear programming relaxation of the Held-Karp integer formulation (i.e. subtour elimination integer formulation) for STSP. This result is due to Shmoys and Williamson [9].

Result 2: The parsimonious property of the linear programming relaxation of the Held-Karp integer formulation for the traveling salesman on a subset $S \subseteq V$. This result is based on the works of Lovasz [7], Goemans and Bertsimas [5].

We show in this paper that one can have the equivalent results to Result 1 and Result 2 for APTP. Precisely,

- The value of a solution given by Frieze et al. heuristic [4] for ATSP is at most $\log(n)$ times the value of an optimal solution of the linear programming relaxation of the Held-Karp integer formulation (i.e. subtour elimination integer formulation) for STSP. This result is due to Williamson [10].
- We prove the parsimonious property of the linear programming relaxation of the Held-Karp integer formulation for the asymmetric traveling salesman on a subset $S \subseteq V$. Our proof is based on the work of Jackson [6].

For solving APTP, we can then apply the same framework as Bienstock et al.'s algorithm and get a $(1 + \log(n))$-approximation algorithm.
The paper is organized as follows. In Section 2, we present a sketch of our algorithm for APTP. We state the parsimonious property for the linear programming relaxation of the Held-Karp integer formulation for the asymmetric traveling salesman on a subset $S \subseteq V$ in Section 3. In Section 4, we show that algorithm guarantee a factor of $(1 + \log(n))$. At last, in the appendix, we sketch a proof for the parsimonious property stated in Section 3.

2 Overview of the algorithm

We use the same framework as in the Bienstock et al.'s algorithm. The algorithm can be summarized as follows. For $j = 1, \ldots, n$, let APTP(j) be the subproblem of APTP which imposes that the vertex j must be in the tour. Let Z^* and $Z^*(j)$ be respectively the optimal solutions to APTP and APTP(j). It then follows that $Z^* = \min\{\sum_{i \in V} \pi_i, \min_{j \in V} Z^*(j)\}$. Based on the fact that APTP(j) can be formulated as an integer program, the algorithm is divided into three steps:

1. Solve the linear programming relaxation of APTP(j) by using the ellipsoid method.
2. Transform the fractional solution into a feasible solution to APTP(j) by first rounding into integers and then by applying the Frieze et al.'s algorithm [4].
3. Out of the $n + 1$ solutions generated (including the empty solution) simply choose the best.

In what follows we describe the integer program for APTP(j), whose optimal solution is denoted by $Z^*(j)$. Let y_i be one if vertex $i \in V$ is in the tour and zero otherwise. Let x_e be one if arc e is in the tour and zero otherwise. For every subset S, let $\delta^+(S)$ be the set of arcs with tail in S and head in $V \setminus S$ and $\delta^-(S)$ be the set of arcs with head in S and tail in $V \setminus S$. Then, the APTP(j) can be formulated as follows:

$$Z^*(j) = \min \sum_{e \in A} c_e x_e + \sum_{i \in V} \pi_i (1 - y_i)$$

subject to

$$x(\delta^+(i) = x(\delta^-(i) = y_i \ \forall i \in V, \tag{1}$$

$$x(\delta^+(S)) = x(\delta^-(S)) \geq y_i \ \forall i \in V, S \subset V \text{ such that } |S \cap \{i, j\}| = 1, \tag{2}$$

$$y_j = 1, \tag{3}$$

$$0 \leq x_e \leq 1 \text{ and integer}, \tag{4}$$

$$0 \leq y_i \leq 1 \text{ and integer } \forall i \neq j, \tag{5}$$

3 Held-Karp relaxation and the parsimonious property

Consider the asymmetric traveling salesman problem defined on the graph G with vector cost c. A well-known lower bound on the length of the optimal tour is given by Held and Karp (1971) and is the solution to the following LP:

$$Z_{HK} = \min \sum_{e \in A} c_e x_e \tag{6}$$

$$\text{subject to} \tag{7}$$

$$x(\delta^+(i)) = x(\delta^-(i)) = 1 \ \forall i \in V, \tag{8}$$

$$x(\delta^+(S)) = x(\delta^-(S)) \geq 1 \ \forall S \subset V, S \neq \emptyset \tag{9}$$

$$x_e \geq 0 \tag{10}$$

Let $L(V)$ be the cost of the optimal asymmetric traveling tour and $L^F(V)$ be the cost of the tour given by Frieze et al.'s heuristic. We have

Theorem 1 *[10]* $Z_{HK}/L(V) \geq Z_{HK}/L^F(V) \geq \log(n)$.

Proof The Frieze et al's algorithm involves iterating the assignment problem. The assignment problem yields a collection of subtours on the nodes. A representative node from each subtour is selected, and the process is iterated. When all remaining nodes are in one subtour, the subtours represented by the remaining nodes are patched in. Since the number of nodes is at least halved on every iteration, at most $\lceil \log(n) \rceil$ iterations are needed. Let A_1, $A_2, \ldots, A_{\lceil \log(n) \rceil}$ be the cost of the $\lceil \log(n) \rceil$ assignment problems. William proved in Lemma 3.2.4 of [10] that $A_i \leq Z_{HK}$ for every $1 \leq i \leq \lceil \log(n) \rceil$. Since $L^F(V) \leq \sum_{i=1} \lceil \log(n) \rceil A_i$, this implies directly Lemma 1. □

For the next lemma we need to formulate the following LP. Associated with each vertex $i \in V$ is a given number r_i which is either zero or one. Let $V_1 = \{i \in V \mid r_i = 1\}$.
Problem P_1:

$$\min \sum_{e \in A} c_e x_e$$

subject to

$$x(\delta^+(i) = x(\delta^-(i) = r_i \;\forall i \in V, \tag{11}$$

$$x(\delta^+(S)) = x(\delta^-(S)) \geq 1 \;\forall S \subset V \text{ such that } V_1 \cap S \neq \emptyset, V_1 \cap V \setminus S \neq \emptyset \tag{12}$$

$$0 \leq x_e \tag{13}$$

Lemma 1 *The optimal solution value to Problem P_1 is unchanged if we solve it without constraint* (11).

4 Analysis of the algorithm for APTP

The algorithm generates n different solution to the APTP by solving the LP relaxation of APTP(j) for every $j \in V$. The j^{th} solution associated with Problem APTP(j) is generated in the following way.
Let $\bar{x}$ and $\bar{y}$ be the optimal solution to the LP relaxation of APTP(j). Define new vectors $\hat{x}$ and $\hat{y}$ as follows:

$$\hat{x}_e = \frac{1 + \log(n)}{\log(n)} \bar{x}_e \;\forall e \in A, \tag{14}$$

and for any $i \in V$

$$\hat{y}_i = \begin{cases} 1, & \text{if } \bar{y}_i \geq \frac{\log(n)}{1+\log(n)} \\ 0, & \text{otherwise} \end{cases}$$

Observe that by definition of $\hat{y}_i$, we have

$$1 - \hat{y}_i \leq (1 + \log(n))(1 - \bar{y}_i) \;\forall i \in V. \tag{15}$$

Notice that we are not claiming that $\hat{x}$, $\hat{y}$ is a feasible solution to the LP relaxation of APTP(j).
Let $T = \{i \in V \mid \hat{y}_i = 1\}$. Our algorithm constructs a traveling salesman tour through all vertices in T using Frieze et al.'s heuristic and therefore charges penalty costs for all vertices not in T. Define

$$Z^F(j) = L^F(T) + \sum_{i \in V} \pi_i (1 - \hat{y}_i),$$

that is, $Z^F(j)$ is the cost of the solution produced by our algorithm, assuming j is in the tour.
Our algorithm chooses the best solution among all such solutions or the solution in which no vertex is visited, whichever yeilds the minimum cost. Hence,

$$Z^F = \min\{\sum_{i \in V} \pi_i, \min_{j \in V}\{Z^F(j)\}\}.$$

Theorem 2 $Z^F/Z^* \leq (1 + \log(n))$.

Proof It is sufficient to show that $Z^F(j)/Z^*(j) \leq (1+\log(n))$, for every j. First, note that the following LP yields the Held and Karp lower bound on the length of the optimal traveling salesman tour through the subset of vertices T:
Problem P_2:

$$\min \sum_{e\in A} c_e x_e \tag{16}$$
subject to
$$x(\delta^+(i) = x(\delta^-(i) = 1 \; \forall i \in T, \tag{17}$$
$$x(\delta^+(i) = x(\delta^-(i) = 0 \; \forall i \in V \setminus T, \tag{18}$$
$$x(\delta^+(S)) = x(\delta^-(S)) \geq 1 \; \forall S \subset V \text{ such that } T\cap S \neq \emptyset, T\cap (V\setminus S) \neq \emptyset \tag{19}$$
$$0 \leq x_e \tag{20}$$

By Lemma 1, the solution value to Problem P_2 is unchanged when we take out contraints (17) and (18). Let Problem P_3 be (16), (19) and (20), and denote by $\tilde{x}$ its optimal solution. Using Theorem 1, we have

$$L^F(T) \leq \log(n) \sum_{e\in A} c_e \tilde{x}_e \tag{21}$$

We now show that $\hat{x}$ is feasible for Problem P_3. Clearly, $\hat{x}$ satisfies (20). To prove that it also satisfies (19) consider any $S \subset V$ such that $T\cap S \neq \emptyset$ with some vertex $i \in T\cap S$ and $T\setminus S \neq \emptyset$ with with some vertex $j \in T\setminus S$. By feasibility of $\bar{x}$ in LP relaxation of APTP(j) and the definition of T we have, using the constraint (1) and equation (14),

$$\sum_{e\in\delta^+(S)} \bar{x}_e = \sum_{e\in\delta^-(S)} \bar{x}_e \geq \bar{y}_i \geq \frac{\log(n)}{1+\log(n)} \; \forall S \subset V \text{ such that } T\cap S \neq \emptyset, T\cap(V\setminus S) \neq \emptyset$$

Hence, for any $S \subset V$ such that $T\cap S \neq \emptyset$, $T\cap (V\setminus S) \neq \emptyset$ we have

$$\sum_{e\in\delta^+(S)} \hat{x}_e = \sum_{e\in\delta^-(S)} \hat{x}_e = \frac{1+\log(n)}{\log(n)} \sum_{e\in\delta^+(S)} \bar{x}_e \geq 1,$$

and therefore $\hat{x}$ satisfies (19). Consequently, since $\tilde{x}$ is optimal

$$\sum_{e\in A} c_e \hat{x}_e \geq \sum_{e\in A} c_e \tilde{x}_e. \tag{22}$$

Hence,

$$\begin{aligned}
Z^F(j) &= L^F(T) + \sum_{i\in V} \pi_i (1-\hat{y}_i) \\
&\leq \log(n) \sum_{e\in A} c_e \tilde{x}_e + \sum_{i\in V} \pi_i(1-\hat{y}_i) \text{ (from (21))} \\
&\leq \log(n) \sum_{e\in A} c_e \hat{x}_e + \sum_{i\in V} \pi_i(1-\hat{y}_i) \text{ (from (22))} \\
&\leq \log(n) \sum_{e\in A} c_e \frac{1+\log(n)}{\log(n)} \bar{x}_e + (1+\log n) \sum_{i\in V} \pi_i (1-\bar{y}_i) \text{ (from (14), (15))} \\
&= (1+\log(n)) (\sum_{e\in A} c_e \bar{x}_e + \sum_{i\in V} \pi_i (1-\bar{y}_i)) \\
&= (1+\log(n)) Z^*(j).
\end{aligned}$$

□

References

1. Balas, E.: The prize collecting traveling salesman problem. *Networks*, Vol. 19, 621-636, (1989).
2. Bienstock, D., Goemans, M.X., Simchi-Levi, D. and Williamson, D.P.: A note on the prize collecting traveling salesman problem, *Math. Prog.,* Vol. 59, 413-420 (1993).
3. Dell'Amico, M., Maffioli, F. and Väbrand P.: On Prize-collecting Tours and the Asymmetric Travelling Salesman Problem, *Int. Trans. Opl Res.,* Vol. 2, 297-308 (1995).
4. Frieze, A.M., Galbiati, G. and Maffioli, F.: On the worst case performance of some algorithms for the asymmetric traveling salesman problem, Networks, 12, 23-39 (1982)
5. Goemans, M.X., Bertsimas, D.J.: On the parsimonious property of connectivity problems, *Proc. 1st ACM-SIAM Symposium on Disc. Alg.*, 388-396 (1990).
6. Jackson, B.: Some remarks on arc-connectivity, vertex splitting, and orientation in graphs and digraphs, *Journal of Graph Theory*, Vol. 12, No.3, 429-436 (1988).
7. Lovasz, L.: On some connectivity properties of Eulerian graphs, *Acta Mathematica Hungarica*, 28, 129-138 (1976).
8. Nguyen, V.H.,: A $2\log(n)$-approximation algorithm for Directed Tour Cover. *LNCS 5573 Proc. of COCOA 2009*, 208-218 (2009).
9. Shmoys, D. and Williamson, D.P.: Analyzing the Held-Karp TSP bound: A monotonicity property with application, *Information Processing Letters*, Vol. 35, 281-285 (1988).
10. Williamson, D. P.: Analysis of the Held-Karp heuristic for the traveling salesman problem, Master thesis, MIT (1990).

5 Appendix

In this section, we provide a short proof of Lemma 1. Nguyen [8] have obtained a more general result which represents the directed version of a similar result for the undirected case of Goemans and Bertsimas [5]. Note that a weaker version of Goemans and Bertsimas's result had been given in a theorem of Lovasz [7] and a directed version of the theorem of Lovasz has been proved by Jackson (Theorem 3 in [6]). In the proof of Lemma 1, we will not need the stronger result in [8], just a simplified version of Jackson's theorem expressed for our purpose as follows.

Lemma 2 *Let $G = (V, A)$ be an Eulerian directed multigraph, $T \subseteq V$ and $s \in V$, such that G is k-strongly connected between any two vertices of T different from s. Then, for any inneighbor u of s (i.e. $u \in \delta^-(s)$), there exists another outneighbor w of s (i.e. $w \in \delta^+(s)$), such that the multigraph obtained from G by removing the arcs (u, s) and (s, w), and adding a new arc (u, w) is also k-strongly connected between two any vertices of T different from s.*

Proof of Lemma 1: Let $V_0 = V \setminus V_1$, that is, $V_0 = \{i \in V \mid r_i = 0\}$. Let Problem P_1' be the problem P_1 without (11). Finally, let $\tilde{x}$ be a rational vector feasible for Problem P_1', chosen such that

C1: $\tilde{x}$ is optimal for Problem P_1', and
C2: subject to C1, $\sum_{e\in A} \tilde{x}_e$ is minimized.

Let M be a positive integer, large enough so that $\tilde{v} = M\tilde{x}$ is a vector of integers. We may regard $\tilde{v}$ (with a slight abuse of notation) as the incidence vector of the arc-set $\tilde{A}$ of a multigraph $\tilde{G}$ with vertex V. Clearly, $\tilde{G}$ is Eulerian, and by (12), it is M-strongly directed between any two elements of V_1. Now suppose that for some vertex s, $\sum_{e\in\delta^+(s)} \tilde{x}_e = \sum_{e\in\delta^-(s)} \tilde{x}_e > r_s$ (i.e. s has a degree larger than Mr_s in $\tilde{G}$). Let us apply Lemma 2 to s and any inneighbor u of s (where $T = V_1$), and let $\tilde{H}$ be the resulting multigraph, with incidence vector $\tilde{z}$.
Clearly,

$$\sum_{e\in A} c_e \tilde{z}_e \le \sum_{e\in A} c_e \tilde{v}_e,$$

and so

$$\sum_{e\in A} c_e \frac{\tilde{z}_e}{M} \le \sum_{e\in A} c_e \tilde{x}_e.$$

Moreover,

$$\sum_{e\in A} \frac{\tilde{z}_e}{M} = \sum_{e\in A} \tilde{x}_e - \frac{1}{M}.$$

Hence, by the choice of $\tilde{x}$, $z = \tilde{z}/M$ cannot be feasible for Problem P_1'.
If $s \in V_0$, then by Lemma 2, z is feasible for Problem P_1'. Thus, we must have $s \in V_1$, and, in fact $\sum_{e\in\delta(t)} \tilde{x}_e = 0$

for all $t \in V_0$. In other word $\tilde{A}$ spans precisely V_1, $\tilde{G}$ is M-strongly connected, and $\sum_{e\in\delta^+(s)} \tilde{v}_e = \sum_{e\in\delta^-(s)} \tilde{v}_e \geq M + 1$. But we claim now that the multigraph $\tilde{H}$ is M-strongly connected. By Lemma 2, it could only fail to be M-strongly connected between s and some other vertex, but the only possible cut of size less than M is $\delta^+(s)$ and $\delta^-(s)$ (after removing the arcs (u, s) and (s, w) and adding a new arc (u, w), the cardinality of all the cuts stay unchanged except the value of $\delta^+(s)$ and $\delta^-(s)$). Since these cuts has at least $M + 1 - 1 = M$ arcs, the claim is proved as desired. Consequently, again we obtain that z is feasible for Problem P_1', a contradiction.
In other words, $\sum_{e\in\delta^+(i)} \tilde{v}_e = \sum_{e\in\delta^-(i)} \tilde{v}_e = Mr_i$ for all i, that is, (11) holds as required. □