A Note on the Prize Collecting Bottleneck TSP and Related Problems

Fabian Hargesheimer *

Abstract

We design a $\frac{10}{3}$ approximation algorithm for a penalty-based prize collecting version of the bottleneck TSP. In this variant of the TSP we look for a tour through a subset of vertices so that the maximum of (1.) the most expensive edge taken and (2.) the highest penalty for not visiting a node, is minimized. The presented algorithm combines linear programming and a rounding method with a heuristic approach for computing a bottleneck optimal tour.

1 Introduction

The original definition of the prize collecting traveling salesman problem (PC-TSP) was given by Balas in 1989 [1]. Given a graph G with associated weights, penalties and prizes, the goal was to find a tour trough a subgraph of G so that the sum of all edge costs in the resulting tour plus the sum of penalties taken for vertices not visited where minimized. As an additional constraint, the sum of all prizes from visited vertices had to fulfill a certain quota in relation to the sum of all available prizes. Without this constraint, the problem is called penalty traveling salesman problem (P-TSP).

The first result on the penalty TSP was an algorithm with a 2.5 approximation ratio by Bienstock et al. in 1993 [2], using a LP-based approach. A 2-approximate algorithm

*Dept. of Computer Science, University of Bonn. Email: hargeshe@cs.uni-bonn.de

was given in 1995 by Goemans and Williamson [3]. It uses a primal-dual approach and also solves the penalty STP with the same approximation ratio. The best known upper bound for P-TSP today is the 2009 result by Goemans [4] which features an approximation ratio of 1.91457 and is based on an improvement to the original Goemans and Williamson algorithm by Archer et al. [5] from the same year.

The bottleneck traveling salesman problem (BN-TSP) was introduced 1979 by Garey and Johnson [6]. In this variant of the TSP the objective is to minimize the cost of the most expensive edge in the resulting tour. In their paper, Garey and Johnson showed the NP-Hardness of the problem and the NP-Completeness of the associated decision problem. For the metric case, there exists a 2-approximate algorithm, shown by Parker and Rardin in 1984 [7]. It works by first computing a bottleneck-optimal biconnected subgraph for the instance in question, and then constructing the square of the bigraph to find an approximate solution. They also showed that this approximation is optimal, unless P = NP.

In the penalty bottleneck TSP (P-BN-TSP) we are given a weighted graph $G = (V, E, c)$ and a penalty function π that assign to each vertex a penalty π_i. The objective is to find a sub-graph $H \subseteq G$, so that H is a cycle (also called tour in this context) and the maximum value of the most expensive edge in H and the highest penalized node not in H is minimized.

In this paper, we will design an algorithm for P-BN-TSP with an approximation rate of $\frac{10}{3}$, using a similar strategy as used by Bienstock et al. in their "MLP heuristic", while taking advantage of some of the unique properties of bottleneck problems. We call it "modified MLP" (MMLP). It combines a LP-relaxation and a rounding technique with a known heuristic of the non-prize collecting version of the problem to get a constant factor approximation algorithm. In this case, we use the BN-TSP heuristic published by Parker and Rardin, which gives 2-approximations for the BN-TSP on graphs fulfilling the triangle inequality.

Since the Parker-Rardin heuristic has a worse approximation ratio then the Christofides heuristic, it is not surprising that MMLP also has a worse approximation ratio than MLP, at $\frac{10}{3}$ compared to $\frac{5}{2}$. However, we have to consider that the lower bound for metric BN-TSP is proven to be 2, which is consequentially also a lower bound for P-BN-TSP, while the best known upper bound for metric penalty TSP is below 2. Also, MMLP's running time is faster than MLP by a factor of n, since unlike MLP, we do not have to compute a solution for every possible initial node.

2 The Algorithm

The algorithm works in two steps: First, the LP-relaxation to the ILP describing the P-BN-TSP is used to determine which of the nodes will be used in the resulting tour. A constant rounding technique is then applied to the result to achieve integer assignments. In the second step, the Parker-Rardin heuristic is used to compute the tour through these nodes.

We will need the following definitions for MMLP: Let $G = (V, E, c)$ be a complete, metric, undirected, weighted graph with positive edge costs and $\pi : V \to \mathbb{Q}^+$ be its assigned penalty function. Let Z^* be the cost of the optimal solution for the penalty BN-TSP and Z^{MMLP} be the resulting cost for the penalty BN-TSP given by the MMLP heuristic. Let BN^* be the cost of the optimal solution for the BN-TSP and let BN^{PR} be the cost for the solution to the BN TSP given by the Parker-Rardin heuristic.

It can be easily seen that if there exist solutions with a lower bottleneck value then the solution consisting of the empty tour with all penalties taken, any node with a penalty value equal to the highest assigned penalty must be part of the resulting tour in all such solutions. This means we will only have to compare two solutions: The one with an empty tour, incurring the highest penalty, and one tour including one of the vertices assigned with the highest penalty. We will denote this vertex by j.

We proceed with the formulation of the underlying ILP, which we will denote by P_1:

$$
\begin{aligned}
Z^* = \ \mathit{minimize} \ \ & \max\left[\max_{e\in E}(c_e x_e), \max_{i\in V}(\pi_i(1-y_i))\right] && (2.1)\\
\mathit{subject\ to} \ \ & \sum_{e\in\delta(\{i\})} x_e = 2y_i \quad \forall i \in V, && (2.2)\\
& \sum_{e\in\delta(S)} x_e \geq 2y_i \quad \forall i \in V, S \subset V s.t. \ |S \cap \{i,j\}\,| = 1, && (2.3)\\
& 0 \leq x_e \leq 1 \ \mathit{and\ integer}, && (2.4)\\
& 0 \leq y_i \leq 1 \ \mathit{and\ integer} \quad \forall i \neq j, && (2.5)\\
& y_j = 1. && (2.6)
\end{aligned}
$$

$y_i, i \in V$ will be 1 for vertices in the tour and 0 if otherwise, $x_e e \in E$ will be 1 for edges in the tour and 0 for edges not in the tour. δ denotes cut sets. The objective function compares the edge with the highest cost and the highest penalty incurred and takes the greater value. Constraint 2.2 manages penalties: If there are no edges taken to y_i, then y_i will be zero, which means that π_i will be fully added to the total cost. 2.3 ensures that any subset that has edges taken is connected by at least two edges to the tour containing j.

We now solve the LP-relaxation of our program using the ellipsoid method and transform our solution into a feasible solution to the P-BN-TSP: To convert the fractional results of P_1 to integers, we round any y_i of value $\frac{3}{5}$ or greater to 1, and to 0 otherwise.

We then define $T = \{i \in V | \hat{y}_i = 1\}$ as the set of nodes with a rounded value of 1 and compare the highest penalty value of the nodes not in T with the resulting bottleneck value when we use T as input to the Parker-Rardin heuristic to compute the value for the highest weight edge in the tour.

$$
Z^{MMLP^*} = \max\left[BN^{PR}(T), \max_{i\in V}(\pi_i(1-y_i))\right] \qquad (2.7)
$$

As final step, we compare this result with the cost of the solution representing the empty tour, given by the penalty of j:

$$Z^{MMLP} = \min \left[Z^{MMLP^*}, \pi_j \right] \tag{2.8}$$

3 Analysis

In this section, we will prove our main theorem:

Theorem 3.1.
MMLP has a worst-case approximation ratio of $\frac{10}{3}$, i.e.

$$\frac{Z^{MMLP}}{Z^*} \leq \frac{10}{3} \tag{3.1}$$

To show this, it is sufficient to show that 2.7 holds. We will now provide the LP-relaxation of the ILP solving the bottleneck TSP:

$$\begin{aligned}
&\textit{minimize} \ \max_{e \in E}(c_e x_e) && (3.2)\\
&\textit{subject to} \sum_{e \in \delta(S)} x_e \geq 2 \quad \forall S \subset V, S \neq \emptyset, && (3.3)\\
&\sum_{e \in \delta(\{i\})} x_e = 2 \quad \forall i \in V, && (3.4)\\
&0 \leq x_e. && (3.5)
\end{aligned}$$

We also need the following:

Lemma 3.2.

$$\frac{BN^{LP}}{BN^*} \geq \frac{BN^{LP}}{BN^{PR}} \geq \frac{1}{2}$$

Where BN^{LP} is the cost of the LP-relaxation for the bottleneck TSP

Proof.

$$\frac{BN^{LP}}{BN^{*}} \geq \frac{BN^{LP}}{BN^{PR}}$$

This can be easily seen, since by definition, the optimal solution can never be worse then any other feasible solution.

$$\frac{BN^{LP}}{BN^{PR}} \geq \frac{1}{2}$$

This follows from the facts that the Parker-Rardin heuristic has a worst-case approximation rate of 2, and the inability of the LP-based heuristic to be better than the optimum solution. □

Lemma 3.3.
The bottleneck value of the solution given by the LP-relaxation of the bottleneck TSP does not change when we remove constraint 3.4.

Proof.
The constraint 3.3 guarantees at least 2-connectivity between all vertices, while constraint 3.4 limits the resulting network to a Hamiltonian cycle. Since the objective value only depends on the maximum of the weights of all edges taken, it can be seen easily that the resulting value for the bottleneck will not change, even if the solution is no longer Hamiltonian. □

We can now proof theorem 3.1:

Let $\bar{x}$ and $\bar{y}$ be the optimal solutions produced by the LP-relaxation of P_1. We use a constant rounding technique to obtain vectors $\hat{x}$ and $\hat{y}$ (These vectors are no longer necessarily feasible solutions to P_1).

$\forall e \in E$:

$$\hat{x}_e = \frac{5}{3}\bar{x}_e \tag{3.6}$$

$\forall i \in V:$

$$\hat{y}_i = \begin{cases} 1 & \text{if } \bar{y}_i \geq \frac{3}{5} \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

The definition of $\hat{y}_i$ leads us to the following observation:

$$1 - \hat{y}_i \leq \frac{5}{2}(1 - \bar{y}_i) \quad \forall i. \tag{3.8}$$

Now we define a modified version of the LP that, given a subset of vertices $T \subseteq V$, results in a tour visiting only vertices in T.

$$\begin{aligned}
&\textit{minimize} \max_{e \in E}(c_e x_e) && (3.9)\\
&\textit{subject to} \sum_{e \in \delta(S)} x_e \geq 2 \quad \forall S \subset V \textit{ such that } T \cap S \neq \emptyset, T \cap (V \setminus S) \neq \emptyset, && (3.10)\\
&\sum_{e \in \delta(\{i\})} x_e = 2 \quad \forall i \in T, && (3.11)\\
&\sum_{e \in \delta(\{i\})} x_e = 0 \quad \forall i \notin T, && (3.12)\\
&0 \leq x_e. && (3.13)
\end{aligned}$$

Using lemma 3.3, we obtain a shortened version of this LP, which we denote as P_2:

$$\begin{aligned}
&\textit{minimize} \max_{e \in E}(c_e x_e) && (3.14)\\
&\textit{subject to} \sum_{e \in \delta(S)} x_e \geq 2 \quad \forall S \subset V \textit{ such that } T \cap S \neq \emptyset, T \cap (V \setminus S) \neq \emptyset, && (3.15)\\
&0 \leq x_e. && (3.16)
\end{aligned}$$

Let $\acute{x}$ be the optimal solution of P_2. We obtain the following using lemma 3.2:

$$BN^{PR} \leq 2 \max_{e \in E} c_e \acute{x}_e. \tag{3.17}$$

Now we prove that $\hat{x}$ is feasible for P_2: It is easy to see that $\hat{x}$ satisfies 3.16 by definition 3.6. To show that $\hat{x}$ satisfies 3.15 we use the feasibility of $\bar{x}$ in P_1, the definition of T, constraint 2.3 and equation 3.6 to get

$$\sum_{e\in\delta(S)} \bar{x}_e \geq 2\bar{y}_i \geq 2\frac{3}{5} = \frac{6}{5} \quad \forall S \subset V \textit{ such that } T \cap S \neq \emptyset, T \cap (V \setminus S) \neq \emptyset. \tag{3.18}$$

Considering any $S \subset V$ such that $T \cap (V \setminus S) \neq \emptyset$ we get

$$\sum_{e\in\delta(S)} \hat{x}_e = \frac{5}{3} \sum_{e\in\delta(S)} \bar{x}_e \geq \frac{5}{3} \cdot \frac{6}{5} = 2 \tag{3.19}$$

Which shows that $\hat{x}$ satisfies 3.15. From the optimality of $\acute{x}$ we also know

$$\max_{e\in E} c_e \hat{x}_e \geq \max_{e\in E} c_e \acute{x}_e \tag{3.20}$$

Finally we can now start from 2.7 and use these results to prove our main theorem:

$$\begin{aligned}
Z^{MMLP^*} &= \max\left[BN^{PR}(T), \max_{i\in V}(\pi_i(1-\hat{y}_i))\right] && (3.21)\\
&\leq \max\left[2\cdot\max_{e\in E}(c_e\acute{x}_e), \max_{i\in V}(\pi_i(1-\hat{y}_i))\right] \quad from(3.17) && (3.22)\\
&\leq \max\left[2\cdot\max_{e\in E}(c_e\hat{x}_e), \max_{i\in V}(\pi_i(1-\hat{y}_i))\right] \quad from(3.20) && (3.23)\\
&\leq \max\left[2\cdot\max_{e\in E}(c_e\frac{5}{3}\bar{x}_e), \max_{i\in V}(\pi_i\frac{5}{2}(1-\bar{y}_i))\right] \quad from(3.6, 3.8) && (3.24)\\
&\leq \max\left[\frac{10}{3}\cdot\max_{e\in E}(c_e\bar{x}_e), \frac{5}{2}\cdot\max_{i\in V}(\pi_i(1-\bar{y}_i))\right] && (3.25)\\
&\leq \frac{10}{3}\cdot\max\left[\max_{e\in E}(c_e\bar{x}_e), \max_{i\in V}(\pi_i(1-\bar{y}_i))\right] && (3.26)\\
&\leq \frac{10}{3}Z^* && (3.27)
\end{aligned}$$

. □

4 Application to Related Problems

The similar general method used in MMLP can also be applied to a number of related problems, like the penalty-sum bottleneck TSP (PS-BN-TSP), which minimizes the maximum of the sum of all penalties incurred and the highest edge cost paid. Using MMLP for PS-BN-TSP we obtain the same approximation ratio of $\frac{10}{3}$. However, our observation about highest penalty nodes being part of optimal tours does no longer hold in this case, which means we will have to run the algorithm on every possible initial node and then compare the results to get a final solution.

References

[1] E. Balas. "The prize collecting travelling salesman problem," *Networks* 19:621-636. 1989.

[2] D. Bienstock, M. X. Goemans, D. Simchi-Levi, D. Williamson. "A note on the prize collecting travelling salesman problem," *Mathematical Programming*, 59:413-420. North-Holland, 1993.

[3] M. Goemans, D. P. Williamson. "A general approximation technique for constrained forest problems," *SIAM Journal on Computing* 24(2):296-317. 1995.

[4] M. X. Goemans. "Combining Approximation Algorithms for the Prize-Collecting TSP," 2009.

[5] A. Archer, M. Bateni, M. Hajiaghayi, H. Karloff "Improved Approximation Algorithms for Prize-Collecting Steiner Tree and TSP," *Proc. of the 50th Annual Symp. on Foundations of Computer Science* 2009.

[6] M. R. Garey, D. S. Johnson. "Computers and Intractability: A Guide to the Theory of NP-Completeness" A2.3:24:212-215. W.H. Freeman. 1979.

[7] R. G. Parker, R. L. Rardin. "Guaranteed performance heuristics for the bottleneck traveling salesman problem," *Operations Research Letters* 2(6):269-272. 1984.