Computing and Informatics, Vol. 29, 2010, 1349–1357

APPROXIMABILITY OF THE MINIMUM STEINER CYCLE PROBLEM

Monika Steinová

ETH Zürich
Informationstechnologie und Ausbildung
CAB F 13.1
Universitätsstrasse 6
8092 Zürich, Switzerland
e-mail: monika.steinova@inf.ethz.ch

Revised manuscript received 11 May 2010

Abstract. In this paper, we consider variants of a new problem that we call minimum Steiner cycle problem (SCP). The problem is defined as follows. Given is a weighted complete graph and a set of terminal vertices. In the SCP problem, we are looking for a minimum-cost cycle that passes through every terminal exactly once and through every other vertex of the graph at most once. We show that, if $P \neq NP$, there is no approximation algorithm for SCP on directed graphs with an approximation ratio polynomial in the input size. Moreover, this result holds even in the case when the number of terminals is 4. On the contrary, we show that SCP on undirected graphs with constant number of terminals and edge costs satisfying the β-relaxed triangle inequality is approximable with the ratio $\beta^2+\beta$. When the number of terminals k is a part of the input, the problem is obviously a generalization of TSP. For the metric case, we present a $\frac{3}{2}$- and a $\frac{2}{3}\log_2 k$-approximation algorithm for undirected and directed graphs $G=(V,E)$, respectively. For the case with the β-relaxed triangle inequality, we present a $(\beta^2+\beta)$-approximation algorithm.

Keywords: Approximation, TSP, Steiner tree, terminals, cycle

Mathematics Subject Classification 2000: 68Q25, 68W25, 68W40

1 INTRODUCTION

Many practical problems in operations research can be formulated as variants of the famous *traveling salesman problem* (TSP). For a complete graph and an edge cost function, the TSP asks to build a minimum-cost cycle that contains all the vertices, each exactly once. Another prominent problem is the well-known *minimum Steiner tree* (MST) problem, where, given a weighted complete graph and a subset of vertices, called terminals, the goal is to find a minimum-cost connected subgraph containing all terminals.

In our paper, we consider a new modification of TSP to an MST-like problem that we refer to as the *minimum Steiner cycle problem* (SCP). As the input for the problem, a complete graph, an edge cost function and a subset of vertices, called terminal vertices, are given. The objective is to minimize the cost of a cycle passing through each terminal vertex exactly once while possibly using some non-terminal vertices, each at most once. The paper deals with a first characterization of SCP and gives some hardness and approximation results.

The paper [12] gives an alternate, more general definition of the SCP. The definition in [12] includes penalties for unvisited vertices. It can easily be shown that, in the general case, the version with vertex penalties can be reduced to a version without them – however, this reduction may break other properties such as the non-negativity of edge costs and triangle inequalities between edge costs. Hence, in order to investigate the problem in more detail, we need to distinguish between these two definitions. As our definition matches the original Steiner tree problem more closely, we suggest to use the name "SCP with vertex penalties" for the general definition from [12].

In [12], the author considers the SCP with vertex penalties as a 0-1 integer linear problem, examines the polytope defined by the set of SCP solutions and develops two lifting procedures to extend facet-defining inequalities from the traveling salesman polytope. These results are not related to our line of research.

The TSP in general is intractable, unless $P = NP$. On the other hand, when the edge cost function satisfies a certain property, the problem becomes much easier. When the cost function satisfies the triangle inequality, the problem is known to be APX-complete [13], but approximable with ratio $\frac{3}{2}$ due to the Christofides algorithm [7]. When the inequality is generalized to a β-relaxed triangle inequality (for all vertices u, v, w, the edge cost function satisfies $c(uv) \leq \beta(c(uw) + c(wv)$ for some $\beta > 1$), the problem still stays constantly approximable with ratio $\min\{\frac{3}{2}\beta^2, \beta^2 + \beta, 4\beta\}$ [6, 1, 4]. When instances of TSP are points in a fixed-dimension Euclidean space and the edge cost function is their Euclidean distance, the problem admits a PTAS [3]. The situation is more difficult in the directed TSP also known as asymmetric TSP. The best currently known approximation algorithm for the directed version of TSP with triangle inequality is $\frac{2}{3}\log_2 n$ (n is the number of vertices in the graph) due to [10]. We are not aware of any approximation algorithm for directed TSP with β-relaxed triangle inequality.

The MST problem is in some sense easier than TSP. In general, it is APX-hard unless $P = NP$ [5, 2], but approximable with a factor $1 + \frac{\ln 3}{2} \approx 1.55$ independent on the edge cost function [16]. When the number of terminals is constant, the problem is known to be in P due to the Dreyfus-Wagner algorithm [8, 14].

In the case where the number of terminals in SCP is constant, we discuss two results here: The main result is that, unless $P = NP$, there is no approximation algorithm for SCP on directed graphs with a constant number of terminals with an approximation ratio polynomial in the input size. This even holds for only four terminals. For undirected graphs with β-relaxed triangle inequality, we give a $(\beta^2 + \beta)$-approximation algorithm.

When the number of terminals k is a part of the input, the problem is obviously a generalization of TSP. For the metric case, we present a $\frac{3}{2}$- and a $\frac{2}{3}\log_2 k$-approximation algorithm for undirected and directed graphs, respectively. For the undirected case with the β-relaxed triangle inequality, we present a $(\beta^2 + \beta)$-approximation algorithm.

The paper is organized as follows. In Section 2, we give the formal definitions of the problems and fix our notation. Section 3 is devoted to the inapproximability results. The approximation algorithms are discussed in Section 4 and the paper is concluded in Section 5.

2 PRELIMINARIES

Unless declared otherwise, we define all the notions in this section for both directed and undirected graphs.

We define a simple graph as a pair $G = (V, E)$, where V is the set of vertices and E is the set of edges. In a directed/undirected graph, an edge is an ordered/unordered pair of two vertices (u, v) that we denote as $\overrightarrow{uv}/uv$.

A (simple) *path* $v_0 v_1 v_2 \dots v_{k-1}$ in a graph G is a non-empty sequence of vertices such that $v_i \neq v_j$ for all $0 \leq i < j \leq k-1$ and $(v_i, v_{i+1}) \in E$ for all $0 \leq i < k-1$. When $v_0 = v_{k-1}$, the path forms a *cycle* of length $k-1$. By $v_0 \rightsquigarrow v_{k-1}$ we denote a path starting in vertex v_0 and ending in v_{k-1}. A *Hamiltonian cycle* of a graph $G = (V, E)$ is a cycle of length $|V|$. A *tree* $T = (V, E)$ is a connected graph without cycles.

A complete graph $G = (V, E)$ with an edge cost function $c : E \to \mathbb{Q}^+$ satisfies the *triangle inequality* if $c(u, v) \leq \beta(c(u, w) + c(w, v))$ for all $u, v, w \in V$ and $\beta = 1$. When $\beta > 1$, the inequality is called *β-relaxed triangle inequality.* The cost of a subgraph of a graph is the sum of the costs of its edges. For simplicity, we write $c(X)$ when we mean the cost of subgraph X.

For a graph $G = (V, E)$, a subset $K \subseteq V$ and a cost function $c : E \to \mathbb{Q}^+$, we refer to the subgraph of G on the vertices from K by $G(K, c)$. The *complete distance network* G_D of a graph G we define as $G(K, c_D)$, where the cost function c_D between any two vertices is defined as the cost of the shortest path connecting them in graph G.

The *minimum Steiner tree* (MST) problem is defined in any undirected graph with edge cost function. The objective is, given a subset of vertices, called *terminal vertices* or *terminals*, to find a minimum-cost subgraph that spans all terminals.

The problem of finding a minimum-cost Hamiltonian cycle in a complete graph satisfying the triangle inequality is called *metric traveling salesman problem* or metric TSP. When the cost function satisfies the β-relaxed triangle inequality, then the problem is called *traveling salesman problem with β-relaxed triangle inequality* and is denoted as β-TSP.

Given a graph $G = (V, E)$ together with a set $K \subseteq V$ of terminals, a *Steiner cycle* in G is a cycle in G containing every terminal from K. The *minimum Steiner cycle problem* (SCP) is the problem of finding a minimum-cost Steiner cycle in a complete graph. We are characterizing several variations of this problem. The SCP problem with a constant number k of terminals is denoted as k-SCP. When the cost function of the graph satisfies the triangle inequality, we denote the problem variants as k-ΔSCP and ΔSCP, respectively. When the cost function satisfies the β-relaxed triangle inequality, we use the notation k-Δ_βSCP and Δ_βSCP, respectively.

3 HARDNESS OF *K*-SCP ON DIRECTED GRAPHS

In this section, we prove that 4-SCP on directed graphs is not approximable with a ratio polynomial in the input size. We show this by reducing the following *NP*-complete problem to the 4-SCP.

Definition 1. The *directed 2-vertex-disjoint path problem* (D-2VDP) is the following decision problem. Given a directed graph $G = (V, E)$ and four different vertices $s_0, s_1, t_0, t_1 \in V$, decide whether G contains a path from s_0 to t_0 and a path from s_1 to t_1 that are vertex-disjoint.

Lemma 1 ([11])**.** The directed 2-vertex-disjoint path problem is *NP*-complete.

Theorem 1. Unless $P = NP$, there is no approximation algorithm for k-SCP ($k > 3$) on directed graphs with an approximation ratio polynomial in the input size.

Proof. Assume that the directed 4-SCP is approximable with ratio $p(n) \geq 1$, where p is a polynomial and n is the number of vertices in the graph. We show how to decide D-2VDP by using the directed 4-SCP approximation algorithm with ratio $p(n)$. Since D-2VDP is *NP*-complete, and under the assumption that $P \neq NP$, this leads to a contradiction; hence, such approximation algorithm cannot exist.

Let $G = (V, E)$ and s_0, s_1, t_0, t_1 be an instance of D-2VDP and let $W := p(|V|) \cdot |V| + 1$. We build a complete directed graph $G' = (V, V \times V \setminus \{\overrightarrow{vv} \mid v \in V\})$ where the cost function c is defined as follows (see Figure 1).

- $c(t_i, s_{1-i}) = 1$, $i \in \{0, 1\}$,
- $c(t_i, x) = W$, $x \in V \setminus \{s_{1-i}\}$, $i \in \{0, 1\}$,

- $c(x, s_i) = W$, $x \in V \setminus \{t_{1-i}\}$, $i \in \{0, 1\}$,
- the costs of all other edges from E are set to 1,
- the costs of all remaining edges that are not in E are set to W.

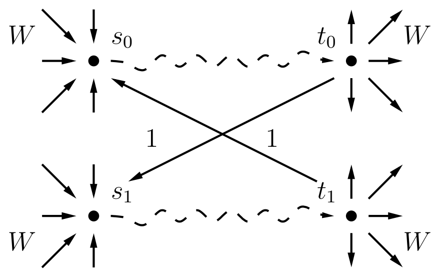

Fig. 1. The edge cost function for graph G'

Let the Steiner cycle S be the outcome of the $p(|V|)$-approximation algorithm for 4-SCP on graph G' and terminals s_0, s_1, t_0, t_1. Then, the graph G contains two vertex-disjoint paths $s_0 \rightsquigarrow t_0$ and $s_1 \rightsquigarrow t_1$ if and only if $c(S) < W$.

Assume that G contains two vertex-disjoint paths $s_0 \rightsquigarrow t_0$ and $s_1 \rightsquigarrow t_1$. Then the sequence $s_0 \rightsquigarrow t_0 s_1 \rightsquigarrow t_1 s_0$ forms a Steiner cycle of cost at most $|V| < W$ proving the $\Rightarrow$ implication.

We prove the $\Leftarrow$ direction by contradiction: Assume that G does not contain vertex-disjoint paths $s_0 \rightsquigarrow t_0$ and $s_1 \rightsquigarrow t_1$ and $c(S) < W$, since $c(S) < W$, S contains edges of cost 1 only. Therefore, except for the edges $\overrightarrow{s_1 t_0}$ and $\overrightarrow{s_0 t_1}$, all other edges of S are present also in the graph G. We show that the only order in which the terminals can be present in S is s_0, t_0, s_1, t_1. The paths between these terminals in the Steiner cycle S identify the paths $s_0 \rightsquigarrow t_0$ and $s_1 \rightsquigarrow t_1$ which will lead to a contradiction.

First start from terminal t_1. Since all other of its outgoing edges are expensive, the following vertex on S is the terminal s_0. Then the following terminal on the cycle S has to be t_0. Otherwise, terminal s_1 would be connected in the cycle S after terminal s_0 by some too expensive edge. Then, since all other outgoing edges from terminal t_0 are too expensive, terminal t_0 is connected in S directly to terminal s_1 by the directed edge $\overrightarrow{t_0 s_1}$. Since all vertices in S occur exactly once and the directed edges present in S, except the two connecting edges $\overrightarrow{t_0 s_1}$ and $\overrightarrow{t_1 s_0}$, are present in graph G, we found the two vertex disjoint paths $s_0 \rightsquigarrow t_0$ and $s_1 \rightsquigarrow t_1$. This is a contradiction. □

Corollary 1. The problem of finding a minimum Steiner cycle on $k > 3$ terminals in a directed graph is *NP*-complete.

The undirected version of k-SCP seems to be different, similarly as there are differences between directed and undirected TSP. The undirected version of the decision problem of finding the k vertex-disjoint paths in a graph (k is constant) is in P, see Robertson and Seymour [15]. The problem of deciding whether an undirected graph contains a shortest edge-/vertex-disjoint path between two given pairs of vertices is known to be in P [9], too. On the other hand, for a given graph and two distinct pairs of vertices, the problem of deciding whether there exist vertex-disjoint paths between them such that one is the shortest path is *NP*-complete for undirected graphs with unit edge-lengths due to [9].

To conclude this discussion, the hardness of k-SCP on undirected graphs is open. We expect that the problem is *NP*-hard for some high constant k, but the previous results mentioned above might be an evidence that the problem is still in P for a small number of terminals.

4 APPROXIMATING THE MINIMUM STEINER CYCLE PROBLEM

In this section, we discuss several approximation algorithms for various settings of SCP, in particular, undirected versions of k-Δ_βSCP and Δ_βSCP and both directed and undirected versions of ΔSCP.

We use the following notation. Let $G = (V, E)$ be a directed/undirected complete graph, let $K \subseteq V$ be a set of terminals and let c be a cost function of G satisfying the (β-relaxed) triangle inequality. Let OPT be an optimal solution of the SCP in the setting that is particularly used.

4.1 Approximation of Directed and Undirected ΔSCP

In both directed and undirected cases, we observe that, due to the triangle inequality, the cost of the direct edge between two vertices is not more expensive than any bypass. Therefore, in this setting, the Hamiltonian cycle on $G(K, c)$ is the minimum Steiner cycle of G with terminals K. In other words, the application of an α-approximation algorithm for metric TSP on the graph $G(K, c)$ gives us an α-approximation of directed/undirected ΔSCP. On the other hand, ΔSCP is a generalization of the metric TSP which allows us to approximate metric TSP as good as ΔSCP.

Theorem 2. ΔSCP is approximable with factor $\frac{3}{2}$ on undirected graphs and with factor $\frac{2}{3}\log_2 k$ on directed graphs (V, E) with k terminals.

Proof. The result for undirected graphs follows from the Christofides algorithm [7]. The result for directed graph follows from algorithm [10]. □

4.2 Approximation of Undirected Δ_βSCP and k-Δ_βSCP

The presented approximation algorithm first builds a tree T that spans all the terminal vertices and whose cost is not higher than the cost of the optimal solution. Then we apply the approximation algorithm for building a Hamiltonian cycle from [1] to obtain a Steiner cycle whose cost is a constant factor away from the cost of the optimal solution.

The tree T is constructed as follows. We build $G_D(K) = (K, c_D)$ and compute a minimum-cost spanning tree T_D for it. Then we transform T_D back to the graph G by replacing the edges by corresponding shortest paths. In this graph, we might have some extra edges that form cycles. Therefore, we build a minimum spanning tree on this graph and we denote it as T. Assume now that we remove the most expensive edge from the optimal solution OPT to obtain a path P. This path sequentially connects the terminals from K. We can replace each such path with the shortest path between the two terminals obtaining some new graph $G' = (V', E')$, $V' \subseteq V$, $E' \subseteq E$. Our tree T is the cheapest out of all such graphs and therefore $c(T) \leq c(G') \leq c(P) \leq c(OPT)$.

The work of Andreae [1] presents one of the approximation algorithms for dealing with β-TSP. The algorithm is very general: After plugging in a tree, it recursively builds a Hamiltonian cycle that contains all the vertices of the given tree and has bounded cost. In our case, we use the algorithm to solve k-Δ_βSCP and Δ_βSCP. According to Theorem 1 of [1], we plug in our tree T and we obtain a Hamiltonian cycle C that contains all the vertices of the tree T. Since T contains all the terminals, C is a Steiner cycle of graph G. The theorem gives the bound $c(C) \leq (\beta^2 + \beta)c(T)$ on the cost of C. In both k-Δ_βSCP and Δ_βSCP, we have $c(T) \leq c(OPT)$ giving us the bound $c(C) \leq (\beta^2 + \beta)c(T) \leq (\beta^2 + \beta)c(OPT)$.

Theorem 3. The undirected versions of k-Δ_βSCP and Δ_βSCP are approximable with factor $(\beta^2 + \beta)$ using the Refined T^3 algorithm from [1].

5 CONCLUSION

In this paper, we have presented hardness results and approximation algorithms for various settings of the minimum Steiner cycle problem. The results are summarized in Table 1. The results for the cases with non-constant number of terminals are presented in the right table. The SCP in this setting is obviously at least as hard as TSP and therefore improving any approximation ratio would directly imply the improvement of TSP approximation. The directed case seems to be harder. It is still open how well the directed TSP with β-relaxed triangle inequality can be approximated. This carries over to our Δ_βSCP. The results for the cases where the number of terminals is constant is shown in the left table. Here the most interesting result is the inapproximability of directed SCP with 4 terminals within a polynomial factor of the input size. Since the directed ΔSCP is in P, two questions arise: What is the setting where directed k-SCP becomes hard? How hard is the undirected version of SCP when the cost function is not bounded?

	constant # of terminals	
c function	directed G	undirected G
Δ	P	P
Δ_β, $\beta > 1$		$\beta^2 + \beta$
general	hard	

	non-constant # of terminals	
c function	directed G	undirected G
Δ	$\frac{2}{3}\log_2 k$	$\frac{3}{2}$
Δ_β, $\beta > 1$		$\beta^2 + \beta$
general	hard	hard

Table 1. An overview of the approximation ratios of the algorithms analyzed in this paper. The table on the left contains results for the setting with a constant number of terminals, the right table contains the results for the non-constant terminal setting. By Δ, Δ_β and "general" we denote the different restrictions on the cost function, k denotes the number of terminals in the graph and "hard" expresses that the problem is not approximable with a polynomial that is dependent on the size of the input.

Acknowledgement

I would like to thank Juraj Hromkovič for introducing the problem, Hans-Joachim Böckenhauer for valuable comments in preparation of this paper and Davide Bilò for helpful discussions about the topic.

This work was partially supported by SNF grant 200021/132510/1.

REFERENCES

[1] Andreae, T.: On the Traveling Salesman Problem Restricted to Inputs Satisfying a Relaxed Triangle Inequality. Networks, Vol. 38, 2001, No. 2, pp. 59–67.

[2] Arora, S.—Lund, C.—Motwani, R.—Sudan, M.—Szegedy, M.: Proof Verification and Hardness of Approximation Problems. In Proc. of the 33rd Annual Symposium on Foundations of Computer Science (FOCS 1992), pp. 14–23, IEEE 1992.

[3] Arora, S.: Polynomial Time Approximation Schemes for Euclidean TSP and Other Geometric Problems. In Proc. of the 37th Annual Symposium on Foundations of Computer Science (FOCS 1996), pp. 2–11, 1996.

[4] Bender, M.—Chekuri, C.: Performance Guarantees for TSP with a Parametrized Triangle Inequality. Information Processing Letters, Vol. 73, 2000, pp. 17–21.

[5] Bern, M. W.—Plassmann, P. E.: The Steiner Problem with Edge Lengths 1 and 2. Information Processing Letters, Vol. 32, 1989, No. 4, pp. 171–176.

[6] Böckenhauer, H.-J.—Hromkovič, J.—Klasing, R.—Seibert, S.—Unger, W.: Towards the Notion of Stability of Approximation for Hard Optimization Tasks and the Traveling Salesman Problem. In Proc. of the 4th Italian Conference on Algorithms and Complexity (CIAC 2000), Vol. 1767 of Lecture Notes in Computer Science, pp. 72–86, Springer-Verlag 2000.

[7] Christofides, N.: Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem. Technical Report 388, Graduate School of Industrial Administration, Carnegie-Mellon University 1976.

[8] Dreyfus, S. E.—Wagner, R. A.: The Steiner Problem in Graphs. Networks, Vol. 1, 1971/72, pp. 195–207.

[9] Eilam-Tzoreff, T.: The Disjoint Shortest Paths Problem. Discrete Applied Mathematics, Vol. 85, 1998, No. 2, pp. 113–138.

[10] Feige, U.—Singh, M.: Improved Approximation Ratios for Traveling Salesperson Tours and Paths in Directed Graphs. In Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 10^{th} International WorkApproximability of the Minimum Steiner Cycle Problem 1009 shop, APPROX 2007, and 11^{th} International Workshop, RANDOM 2007, pp. 104–118.

[11] Fortune, S.—Hopcroft, J. E.—Wyllie, J.: The Directed Subgraph Homeomorphism Problem. Theoretical Computer Science, Vol. 10, 1980, pp. 111–121.

[12] Salazar González, J. J.: The Steiner Cycle Polytope. European Journal of Operational Research, Vol. 147, 2003, No. 3, pp. 671–679.

[13] Papadimitriou, C. H.—Yannakakis, M.: The Traveling Salesman Problem With Distances One and Two. Mathematics of Operations Research, Vol. 18, 1993, No. 1, pp. 1–11.

[14] Prömel, H. J.—Steger, A.: The Steiner Tree Problem. Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn 2002.

[15] Robertson, N.—Seymour, P. D.: Graph Minors. XIII. The Disjoint Paths Problem. Journal of Combinatorial Theory, Series B, Vol. 63, 1995, No. 1, pp. 65–110.

[16] Robins, G.—Zelikovsky, A. Z.: Improved Steiner Tree Approximation in Graphs. In Proc. of the 11^{th} Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2000), pp. 770–779, ACM/SIAM 2000.

Monika Steinová is currently a Ph. D. student at ETH Zürich in Switzerland. She received a master's degree in computer science from Comenius University, Slovakia in 2007. She was awarded the Google Europe Anita Borg Scholarship in 2007. Her research interests include complexity theory, approximation algorithms, parametrized complexity and distributed computing.