MPRI Algorithms Lab: The Steiner Tree Problem

Xiang Wan

October 3, 2025

1 Problem Modeling

1.1 Problem Definition: Minimum Steiner Tree

The following is the standard formal definition of the Minimum Steiner Tree problem, as found in the compendium.

Instance: A complete graph $G = (V, E)$, a metric given by edge weights $w : E \to \mathbb{Z}^+$, and a subset of required vertices $S \subset V$.

Solution: A Steiner tree, i.e., a subtree of G that includes all the vertices in S.

Measure: The sum of the weights of the edges in the subtree, which is to be minimized.

Good News: The problem is approximable within a ratio of $1 + \ln(3)/2 \approx 1.55$. In your project file, this approximation ratio is simplified to 2.

Bad News: The problem is APX-complete.

Garey and Johnson: ND12.

1.2 NP-completeness Proof

To prove that the Minimum Steiner Tree problem (as stated above: *complete graph* with *metric* integer weights) is NP-complete we do two things:

1. Show membership in NP.

2. Provide a polynomial-time reduction from the NP-complete **Exact Cover** problem to our problem statement. The reduction is carried out in two steps: (i) build a general graph with positive integer weights (no zero weights) via a standard construction from Exact Cover; (ii) take the *metric closure* (shortest-path completion) of that graph to obtain a complete graph whose edge weights are positive integers and satisfy the triangle inequality. We prove these transformations preserve the existence of a Steiner tree within the chosen budget.

1.2.1 Membership in NP

Given an instance $(G = (V, E), w, R, B)$ of the Steiner Tree problem (here G is a complete metric graph and $w : E \to \mathbb{Z}^+$), a certificate is a subtree T of G (for example given as a list of edges). We can verify in polynomial time:

- that the listed edges form a connected acyclic subgraph (a tree) on the claimed vertex set,
- that every terminal in R appears in the vertex set of T,
- that the sum of the weights of the edges of T is at most B.

Thus the problem is in NP.

1.2.2 Reduction from Exact Cover (two-step: graph construction + metric closure)

We start from an instance of Exact Cover.

Exact Cover (decision) instance. Let the instance be a universe $U = \{u_1, \ldots, u_n\}$ and a collection of subsets $\mathcal{S} = \{S_1, \ldots, S_k\}$. The question is whether there exists a subcollection $\mathcal{S}' \subseteq \mathcal{S}$ such that the sets in $\mathcal{S}'$ are pairwise disjoint and their union equals U.

Step 1: Construct a (sparse) weighted graph $G' = (V', E')$ with positive integer weights. We construct G' as follows (this construction is polynomial in the size of the Exact Cover instance):

- Vertices:
$$V' = \{a_0\} \ \cup \ \{s_1, \ldots, s_k\} \ \cup \ \{u_1, \ldots, u_n\}.$$
(Here a_0 is a special root vertex, each s_i corresponds to subset S_i, and each u_j corresponds to element $u_j \in U$.)
- Edges and weights (all weights are positive integers):
  - For each $i \in \{1, \ldots, k\}$ add edge (a_0, s_i) with weight
  $$w'(a_0, s_i) = |S_i|.$$
  - For each i and each $u_j \in S_i$ add edge (s_i, u_j) with weight
  $$w'(s_i, u_j) = 1.$$
  - No other edges are added in G'.
- Terminals and budget:
$$R = \{a_0, u_1, \ldots, u_n\}, \qquad B = 2n.$$

All weights in G' are positive integers and the graph size is polynomial in $n + k$.

Correctness of Step 1 (equivalence between Exact Cover and Steiner tree in G' with budget $B = 2n$). We show:

Exact cover exists $\quad \Longleftrightarrow \quad$ there exists a Steiner tree in G' that connects R with total weight ≤ 2

(⇒) Suppose there is an exact cover $\mathcal{S}' \subseteq \mathcal{S}$. Then the sets in $\mathcal{S}'$ are disjoint and their union is U, so $\sum_{S_i \in \mathcal{S}'} |S_i| = n$. Construct a subtree T' of G' as follows: include edges (a_0, s_i) for each $S_i \in \mathcal{S}'$, and for each element $u_j \in U$ include the unique edge (s_i, u_j) where $S_i \in \mathcal{S}'$ is the unique set containing u_j. The resulting subgraph connects a_0 to every u_j and is acyclic (it is a forest that can be pruned to a tree if necessary); its total weight equals

$$\sum_{S_i \in \mathcal{S}'} w'(a_0, s_i) \;+\; \sum_{j=1}^{n} w'(s_{i(j)}, u_j) \;=\; \sum_{S_i \in \mathcal{S}'} |S_i| \;+\; n \;=\; n + n \;=\; 2n.$$

Thus there is a Steiner tree of weight $2n$ (hence $\leq 2n$).

(⇐) Conversely, suppose there exists a Steiner tree T' in G' that connects all terminals R and has total weight $W(T') \leq 2n$. Observe:

- Every terminal u_j has neighbors only among the vertices $\{s_i : u_j \in S_i\}$ in G'. Therefore in any connected subgraph that contains a_0 and u_j, the path from a_0 to u_j must use at least one edge of the form (s_i, u_j) (for some i with $u_j \in S_i$). Hence the tree T' contains at least one (s_i, u_j) edge for each element u_j, so the total contribution to $W(T')$ from edges of type (s_i, u_j) is at least n (each has weight ≥ 1 and there are at least n of them).

- Let $\mathcal{S}''$ be the collection of sets corresponding to those s_i for which T' contains the edge (a_0, s_i); these are exactly the subset-vertices used to connect a_0 into the rest of the tree. Because every u_j must be connected to a_0 in T', every u_j must belong to at least one set in $\mathcal{S}''$, so $\mathcal{S}''$ covers U. Therefore the contribution to $W(T')$ from edges of type (a_0, s_i) is

$$\sum_{S_i \in \mathcal{S}''} w'(a_0, s_i) \;=\; \sum_{S_i \in \mathcal{S}''} |S_i| \;\geq\; n.$$

Combining the two contributions yields $W(T') \geq n + n = 2n$. Since we assumed $W(T') \leq 2n$, we must have equality throughout. Equality implies (i) the number of (s_i, u_j) edges used is exactly n (so for each u_j exactly one such edge is used in T'), and (ii) $\sum_{S_i \in \mathcal{S}''} |S_i| = n$. The latter together with the fact that $\mathcal{S}''$ covers U forces the sets in $\mathcal{S}''$ to be pairwise disjoint and to partition U; hence $\mathcal{S}''$ is an exact cover. This proves the equivalence.

Thus Exact Cover reduces in polynomial time to the Steiner Tree problem on the general graph G' with positive integer weights and budget $2n$.

Step 2: From the general graph G' to a complete metric graph G^* (metric closure). We now transform G' into a complete graph $G^* = (V', E^*)$ on the same vertex set V' by assigning to every unordered pair $\{x, y\}$ the weight equal to the shortest-path distance in G' (with respect to w'):

$$w^*(x, y) := \mathrm{dist}_{G'}(x, y) = \min\{\text{sum of } w'\text{-weights along a path from } x \text{ to } y\}.$$

Since all w' are positive integers, all distances $w^*(x, y)$ are positive integers. By construction w^* satisfies the triangle inequality, so (V', w^*) is a metric and G^* is a complete metric graph matching the format required by the problem statement in Section **??**.

We must show the existence of a Steiner tree for terminals R of weight at most $2n$ in G' is equivalent to the existence of a Steiner tree for R of weight at most $2n$ in G^*. The standard argument is:

- If T is any Steiner tree in G', then viewing T as a subgraph of G^* (each edge of T also corresponds to a pair of vertices in G^*) we have $w^*(x, y) \leq w'(x, y)$ for every edge $(x, y) \in T$. Hence the total weight of the same edge set measured under w^* is at most its weight under w'. Therefore a feasible Steiner tree in G' with weight $\leq 2n$ yields a feasible Steiner tree in G^* with weight $\leq 2n$.

- Conversely, if T^* is a Steiner tree in G^* with weight $W^* \leq 2n$, replace each edge (x, y) of T^* by a shortest path between x and y in G'. The union H of these shortest paths is a connected subgraph of G' whose total w'-weight equals W^*. From H we can extract a spanning tree (on the vertices involved) by removing cycles; removing edges cannot increase total weight. The result is a tree in G' connecting all terminals R with total weight $\leq W^* \leq 2n$. Hence a feasible solution in G^* corresponds to a feasible solution in G'.

Therefore the two-step reduction (Exact Cover $\to G'$ with positive integer weights $\to$ metric closure G^*) produces in polynomial time a complete metric graph instance $(G^*, w^*, R, B = 2n)$ that has a Steiner tree of weight at most $2n$ iff the original Exact Cover instance has a solution.

1.2.3 Conclusion

Exact Cover is NP-complete; we have given a polynomial-time reduction from Exact Cover to the Minimum Steiner Tree problem as stated in Section **??** (complete graph, metric weights in $\mathbb{Z}^+$). Combined with the observation that Steiner Tree belongs to NP, it follows that the Minimum Steiner Tree problem (in the formulation used in this document) is NP-complete.

References

[1] Richard M. Karp. Reducibility Among Combinatorial Problems. In M. Jünger et al. (eds.), *50 Years of Integer Programming 1958-2008*, pages 219–241. Springer-Verlag Berlin Heidelberg, 2010. (Reprint of the original 1972 paper).