The relation of Connected Set Cover and Group Steiner Tree

Slobodan Jelić\(^1\) \quad Domagoj Matijević\(^1\)

\(^1\)Josip Juraj Strossmayer University of Osijek
Department of Mathematics

**ApplMath11**
7\(^{th}\) Conference on Applied Mathematics and Scientific Computing
June 13-17, 2011
Trogir, Croatia
Introduction

- NP-hard combinatorial optimizations problem
- efficient algorithm not known
- polynomial approximation

\[ OPT \leq ALG \leq \rho \times OPT \]

- \( \rho \) - approximation algorithm
- \( \rho = \rho(\cdot) \) - function of input size = approximation ratio
Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$
  - $\bigcup_{S \in S} S = U$
- Problem: Find $\mathcal{R} \subseteq S$
  - each $u \in U$ is contained in at least one set from $\mathcal{R}$
  - among all such subfamilies, $\mathcal{R}$ has minimal cardinality
Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$
  - $\bigcup_{S \in S} S = U$
- Problem: Find $\mathcal{R} \subseteq S$
  - each $u \in U$ is contained in at least one set from $\mathcal{R}$
  - among all such subfamilies, $\mathcal{R}$ has minimal cardinality
Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$
  - $\bigcup_{S \in S} S = U$
- **Problem:** Find $\mathcal{R} \subseteq S$
  - each $u \in U$ is contained in at least one set from $\mathcal{R}$
  - among all such subfamilies, $\mathcal{R}$ has minimal cardinality
Set Cover

\[ S_1 \quad S_2 \]

\[ S_3 \]

\[ S_4 \]
Set Cover

The relation of Connected Set Cover and Group Steiner Tree
Set Cover

The relation of Connected Set Cover and Group Steiner Tree

optimal

$S_3$

$S_4$
Connected Set Cover

- \( U \) - set of elements (universe)
- \( S \) family of subsets of \( U \)
- \( \bigcup_{S \in S} S = U \)
- \( G = (S, E) \) connected graph
Connected Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$

\[ \bigcup_{S \in S} S = U \]

- $G = (S, E)$ connected graph
Connected Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$
  \[ \bigcup_{S \in S} S = U \]
- $G = (S, E)$ connected graph
**Connected Set Cover**

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$

$$\bigcup_{S \in S} S = U$$

- $G = (S, E)$ connected graph
Connected Set Cover

- $U$ - set of elements (universe)
- $S$ family of subsets of $U$

\[ \bigcup_{S \in S} S = U \]

- $G = (S, E)$ connected graph
Connected Set Cover

**Problem:** Find $\mathcal{R} \subseteq \mathcal{S}$

- each $u \in U$ is contained in at least one set from $\mathcal{R}$
- $G[\mathcal{R}]$ is connected
- among all such subfamilies, $\mathcal{R}$ has minimal cardinality

Node weighted case:

- each set $S$ has nonnegative weight $w(S)$
- $\sum_{S \in \mathcal{R}} w(S)$ minimized
- NWCSC - Node Weighted Connected Set Cover
Connected Set Cover

**Problem:** Find $\mathcal{R} \subseteq S$
- each $u \in U$ is contained in at least one set from $\mathcal{R}$
- $G[\mathcal{R}]$ is connected
- among all such subfamilies, $\mathcal{R}$ has minimal cardinality

**Node weighted case:**
- each set $S$ has nonnegative weight $w(S)$
- $\sum_{S \in \mathcal{R}} w(S)$ minimized
- NWCSC - Node Weighted Connected Set Cover
Connected Set Cover
Connected Set Cover

The relation of Connected Set Cover and Group Steiner Tree
Connected Set Cover

The relation of Connected Set Cover and Group Steiner Tree
Connected Set Cover

Slobodan Jelić, Domagoj Matijević

The relation of Connected Set Cover and Group Steiner Tree
The relation of Connected Set Cover and Group Steiner Tree

Slobodan Jelić, Domagoj Matijević
Connected Set Cover

The relation of Connected Set Cover and Group Steiner Tree
Applications of CSC problems

- constructing optimal biodiversity reserve system
  - set of species live on protected areas - reserves
  - choose subset of reserves which represents all species
  - overall number (cost) of reserves has minimized


"...movement and species richness are positively affected by corridors and connectivity..."

- problem can be modeled as CSC
Applications of CSC problems

- constructing optimal biodiversity reserve system
  - set of species live on protected areas - reserves
  - choose subset of reserves which represents all species
  - overall number (cost) of reserves has minimized

  - "...movement and species richness are positively affected by corridors and connectivity..."

- problem can be modeled as CSC
Applications of CSC problems

- constructing optimal biodiversity reserve system
  - set of species live on protected areas - reserves
  - choose subset of reserves which represents all species
  - overall number (cost) of reserves has minimized

  - "...movement and species richness are positively affected by corridors and connectivity..."

- problem can be modeled as CSC
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: $1 + D_c(G)H(\gamma - 1)$
    - $D_c(G)$ - length of the longest path in graph G between two non-disjoint sets
    - $\gamma = \max \{|S| : S \in S\}$
- NWCSC presented as an open problem
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: \(1 + D_c(G)H(\gamma - 1)\)
    - \(D_c(G)\) - length of the longest path in graph \(G\) between two non-disjoint sets
    - \(\gamma = \max \{|S| : S \in S\}\)
- NWCSC presented as an open problem
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: $1 + \frac{D_c(G)H(\gamma - 1)}{D_c(G)}$
    - $D_c(G)$ - length of the longest path in graph $G$ between two non-disjoint sets
    - $\gamma = \max \{|S| : S \in \mathcal{S}\}$
  - NWCSC presented as an open problem
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: $1 + D_c(G)H(\gamma - 1)$
    - $D_c(G)$ - length of the longest path in graph $G$ between two non-disjoint sets
    - $\gamma = \max \{|S| : S \in S\}$
  - NWCSC presented as an open problem
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: $1 + D_c(G)H(\gamma - 1)$
    - $D_c(G)$ - length of the longest path in graph $G$ between two non-disjoint sets
    - $\gamma = \max \{|S| : S \in S\}$
- NWCSC presented as an open problem
Previous work on CSC

  - valid inequalities for CSC polytope
  - connected set cover on line and spider graphs
  - first greedy approximation algorithm for CSC on general graph
  - approximation ratio: $1 + D_c(G)H(\gamma - 1)$
    - $D_c(G)$ - length of the longest path in graph $G$ between two non-disjoint sets
    - $\gamma = \max \{|S| : S \in S\}$
- NWCSC presented as an open problem
Our improvements on CSC and NWCSC

- $D_c(G) \in O(m)$
- Improvement on CSC:
  - Approximation ratio $O(\log^2 m \log \log m \log n)$
- Results on NWCSC:
  - We are not aware of any previous algorithm for NWCSC
  - Our algorithm: $O(\sqrt{m} \log m)$
Our improvements on CSC and NWCSC

- $D_c(G) \in O(m)$
- improvement on CSC:
  - approximation ratio $O(\log^2 m \log \log m \log n)$
- results on NWCSC:
  - we are not aware of any previous algorithm for NWCSC
  - our algorithm: $O(\sqrt{m} \log m)$
Our improvements on CSC and NWCSC

- $D_c(G) \in O(m)$
- Improvement on CSC:
  - Approximation ratio $O(\log^2 m \log \log m \log n)$
- Results on NWCSC:
  - We are not aware of any previous algorithm for NWCSC
  - Our algorithm: $O(\sqrt{m} \log m)$
Bad example for greedy strategy

- $n \in \mathbb{N}$ is even, $n \geq 6$

  \[ U = \{1, 2, \ldots, n\} \]

- $S = \{S_1, \ldots, S_n\}$
  - $S_1 = \{1, 2, \ldots, n/2\}$
  - $S_n = \{n/2 + 1, \ldots, n\}$
  - $S_i = \{i - 1, i\}$ for $2 \leq i \leq n - 1$
Bad example for greedy strategy

- $n \in \mathbb{N}$ is even, $n \geq 6$

  $$U = \{1, 2, \ldots, n\}$$

- $S = \{S_1, \ldots, S_n\}$
  - $S_1 = \{1, 2, \ldots, n/2\}$
  - $S_n = \{n/2 + 1, \ldots, n\}$
  - $S_i = \{i - 1, i\}$ for $2 \leq i \leq n - 1$
greedy: $S$

optimal: $\{S_1, S_n\}$
Group Steiner Tree (GST)

- introduced by Reich and Widmayer (1990).
- motivation: wire routing with multiport terminals in physical VLSI design
- definition:
  - graph $G$
  - edge weight function $w : E(G) \rightarrow \mathbb{R}^+$
  - $G = \{g_1, g_2, \ldots, g_k\}$, $g_i \subset V$
  - objective: find subtree $T$
    - $V(T) \cap g_i \neq \emptyset$ for all $i \in \{1, \ldots, k\}$
    - minimize $\sum_{e \in E(T)} w(e)$
- node weighted case:
  - each vertex $v \in V(G)$ has nonnegative weight $w_N(v)$
  - minimize $\sum_{v \in V(T)} w_N(v)$
- NWGST - Node Weighted Group Steiner Tree
Group Steiner Tree (GST)

- introduced by Reich and Widmayer (1990).
- motivation: wire routing with multiport terminals in physical VLSI design
- definition:
  - graph $G$
  - edge weight function $w : E(G) \rightarrow \mathbb{R}^+$
  - $G = \{g_1, g_2, \ldots, g_k\}$, $g_i \subset V$
  - objective: find subtree $T$
    - $V(T) \cap g_i \neq \emptyset$ for all $i \in \{1, \ldots, k\}$
    - minimize $\sum_{e \in E(T)} w(e)$
- node weighted case:
  - each vertex $v \in V(G)$ has nonnegative weight $w_N(v)$
  - minimize $\sum_{v \in V(T)} w_N(v)$
- NWGST - Node Weighted Group Steiner Tree
introduced by Reich and Widmayer (1990).

motivation: wire routing with multiport terminals in physical VLSI design

definition:

- graph $G$
- edge weight function $w : E(G) \rightarrow \mathbb{R}^+$
- $G = \{g_1, g_2, \ldots, g_k\}$, $g_i \subset V$
- objective: find subtree $T$
  - $V(T) \cap g_i \neq \emptyset$ for all $i \in \{1, \ldots, k\}$
  - minimize $\sum_{e \in E(T)} w(e)$

node weighted case:

- each vertex $v \in V(G)$ has nonnegative weight $w_N(v)$
- minimize $\sum_{v \in V(T)} w_N(v)$

NWGST - Node Weighted Group Steiner Tree
introduced by Reich and Widmayer (1990).

motivation: wire routing with multiport terminals in physical VLSI design

definition:

- graph $G$
- edge weight function $w : E(G) \rightarrow \mathbb{R}^+$
- $G = \{g_1, g_2, \ldots, g_k\}$, $g_i \subset V$
- **objective**: find subtree $T$
  - $V(T) \cap g_i \neq \emptyset$ for all $i \in \{1, \ldots, k\}$
  - minimize $\sum_{e \in E(T)} w(e)$

node weighted case:

- each vertex $v \in V(G)$ has nonnegative weight $w_N(v)$
- minimize $\sum_{v \in V(T)} w_N(v)$

NWGST - Node Weighted Group Steiner Tree
Group Steiner Tree (GST)

- introduced by Reich and Widmayer (1990).
- motivation: wire routing with multiport terminals in physical VLSI design
- definition:
  - graph $G$
  - edge weight function $w : E(G) \rightarrow \mathbb{R}^+$
  - $G = \{g_1, g_2, \ldots, g_k\}$, $g_i \subset V$
  - objective: find subtree $T$
    - $V(T) \cap g_i \neq \emptyset$ for all $i \in \{1, \ldots, k\}$
    - minimize $\sum_{e \in E(T)} w(e)$
  - node weighted case:
    - each vertex $v \in V(G)$ has nonnegative weight $w_N(v)$
    - minimize $\sum_{v \in V(T)} w_N(v)$
- NWGST - Node Weighted Group Steiner Tree

Slobodan Jelić, Domagoj Matijević

The relation of Connected Set Cover and Group Steiner Tree
Group Steiner Tree

4 groups:
total weight = 5
optimal weight = 4
The relation of Connected Set Cover and Group Steiner Tree

4 groups:

- Total weight = 5
- Total weight = 4
- Optimal
Group Steiner Tree

4 groups:

optimal

total weight = 5

The relation of Connected Set Cover and Group Steiner Tree

Slobodan Jelić, Domagoj Matijević
total weight = 5
Group Steiner Tree

total weight = 4

optimal
Previous work on GST

  - first polylogarithmic approximation algorithm for GST
  - solving problem on trees in approximation ratio $O(\log k \log N)$
  - technique: LP relaxation + randomized rounding
  - generalization: stretch $= O(\log n \log \log n)$
  - Bartal probabilistic approximation of metric spaces

- result on general graph: $O(\log N \log n \log \log n \log k)$
Previous work on GST

  - first polylogarithmic approximation algorithm for GST
  - solving problem on trees in approximation ratio $O(\log k \log N)$
  - technique: LP relaxation + randomized rounding
  - generalization: stretch $= O(\log n \log \log n \log \log k)$
  - Bartal probabilistic approximation of metric spaces

- result on general graph: $O(\log N \log n \log \log n \log k)$
Previous work on NWGST

  - first approximation of NWGST
  - $O(\sqrt{n} \log n)$ - approximation algorithm
Previous work on NWGST

  - first approximation of NWGST
  - $O(\sqrt{n} \log n)$ - approximation algorithm
Previous work on NWGST

  - first approximation of NWGST
  - $O(\sqrt{n} \log n)$ - approximation algorithm
Main Result: Equivalence of NWCSC and NWGST

- NWCSC and NWGST are equivalent
  - NWCSC reducible to NWGST
  - NWGST reducible to NWCSC
- NWCSC algorithm:
  - reduce NWCSC to NWGST
  - solve NWGST by Khandekar et al. algorithm
- CSC algorithm:
  - reduce CSC to GST with edge weights equal to 1
  - solve GST by Garg et al. algorithm
Main Result: Equivalence of NWCSC and NWGST

• NWCSC and NWGST are equivalent
  • NWCSC reducible to NWGST
  • NWGST reducible to NWCSC

• NWCSC algorithm:
  • reduce NWCSC to NWGST
  • solve NWGST by Khandekar et al. algorithm

• CSC algorithm:
  • reduce CSC to GST with edge weights equal to 1
  • solve GST by Garg et al. algorithm
Main Result: Equivalence of NWCSC and NWGST

- NWCSC and NWGST are equivalent
  - NWCSC reducible to NWGST
  - NWGST reducible to NWCSC

- NWCSC algorithm:
  - reduce NWCSC to NWGST
  - solve NWGST by Khandekar et al. algorithm

- CSC algorithm:
  - reduce CSC to GST with edge weights equal to 1
  - solve GST by Garg et al. algorithm
Main Result: Equivalence of NWCSC and NWGST

- NWCSC and NWGST are equivalent
  - NWCSC reducible to NWGST
  - NWGST reducible to NWCSC
- NWCSC algorithm:
  - reduce NWCSC to NWGST
  - solve NWGST by Khandekar et al. algorithm
- CSC algorithm:
  - reduce CSC to GST with edge weights equal to 1
  - solve GST by Garg et al. algorithm
Main Result: Equivalence of NWCSC and NWGST

- NWCSC and NWGST are equivalent
  - NWCSC reducible to NWGST
  - NWGST reducible to NWCSC
- NWCSC algorithm:
  - reduce NWCSC to NWGST
  - solve NWGST by Khandekar et al. algorithm
- CSC algorithm:
  - reduce CSC to GST with edge weights equal to 1
  - solve GST by Garg et al. algorithm
CSC with requirements

  - solution subgraph has to be $k$-connected
  - each element covered by at least $m$ sets

- $(k,m)$-CSC
- generalization of $(1,m)$-CSC
  - each element $u$ has requirement $r_u \in \mathbb{N}$
  - element $u$ has to be covered by at least $r_u$ sets

- Connected Set Cover with requirements (CSC-R)
CSC with requirements

  - solution subgraph has to be $k$-connected
  - each element covered by at least $m$ sets

- $(k, m)$-CSC
  - generalization of $(1,m)$-CSC
    - each element $u$ has requirement $r_u \in \mathbb{N}$
    - element $u$ has to be covered by at least $r_u$ sets

- Connected Set Cover with requirements (CSC-R)
CSC with requirements

  - solution subgraph has to be $k$-connected
  - each element covered by at least $m$ sets

- $(k, m)$ - CSC
- generalization of $(1, m)$-CSC
  - each element $u$ has requirement $r_u \in \mathbb{N}$
  - element $u$ has to be covered by at least $r_u$ sets

- Connected Set Cover with requirements (CSC-R)
CSC with requirements

  - solution subgraph has to be \(k\)-connected
  - each element covered by at least \(m\) sets
- \((k, m)\) - CSC
- generalization of \((1, m)\)-CSC
  - each element \(u\) has requirement \(r_u \in \mathbb{N}\)
  - element \(u\) has to be covered by at least \(r_u\) sets
- Connected Set Cover with requirements (CSC-R)
Covering Steiner Tree problem

- generalization of GST
  - each group $g \in \mathcal{G}$ has to be covered at least $k_g$ times, $k_g \in \mathbb{N}$
  - polylog approximation for CST
  - approximation ratio: $O(\log n \log \log n \log N \log(K \cdot k))$
Covering Steiner Tree problem

- generalization of GST
  - each group \( g \in \mathcal{G} \) has to be covered at least \( k_g \) times, \( k_g \in \mathbb{N} \)
  - polylog approximation for CST
  - approximation ratio: \( O(\log n \log \log n \log N \log(K \cdot k)) \)
Our algorithm for CSC-R

- CSC-R equivalent to CST with mutually equal edge weights \((w(e) = 1 \text{ for all } e \in E(G))\)
- Algorithm for CSC-R:
  - reduce CSC-R to CST
  - Approximation ratio for CSC-R: \(O(\log^2 m \log \log m \log(R \cdot n))\)
  - \(R = \max_{u \in U} r_u\)
Our algorithm for CSC-R

- CSC-R equivalent to CST with mutually equal edge weights ($w(e) = 1$ for all $e \in E(G)$)
- algorithm for CSC-R:
  - reduce CSC-R to CST
  - approximation ratio for CSC-R: $O(\log^2 m \log \log m \log(R \cdot n))$
  - $R = \max_{u \in U} r_u$
NWCSC to NWGST reduction

- Reduction from NWCSC to NWGST
  - Fixed element \( u \in U \)
  \[
g_u = \{ S \in S : u \in S \}
\]
- Graph \( G \), node weights \( w_N(\cdot) \)
- NWGST instance
- Let \( T \) is solution of NWGST
- Take \( R = V(T) \) as solution of NWCSC
  - \( G[R] \) is connected since \( T \) is connected
  - \( R \) is set cover since \( T \) touches each group

- CSC to GST reduction: \( w(e) = 1 \) for all \( e \in E(G) \)
- CSC-R to CST reduction: \( k_{g_u} = r_u \)
NWCSC to NWGST reduction

- reduction from NWCSC to NWGST
- fixed element $u \in U$
  \[ g_u = \{ S \in S : u \in S \} \]
- graph $G$, node weights $w_N(\cdot)$
- NWGST instance
- let $T$ is solution of NWGST
- take $\mathcal{R} = V(T)$ as solution of NWCSC
  - $G[\mathcal{R}]$ is connected since $T$ is connected
  - $\mathcal{R}$ is set cover since $T$ touches each group

- CSC to GST reduction: $w(e) = 1$ for all $e \in E(G)$
- CSC-R to CST reduction: $k_{gu} = r_u$
NWCSC to NWGST reduction

- reduction from NWCSC to NWGST
  - fixed element $u \in U$

$$g_u = \{ S \in S : u \in S \}$$

- graph $G$, node weights $w_N(\cdot)$
- NWGST instance
- let $T$ is solution of NWGST
- take $\mathcal{R} = V(T)$ as solution of NWCSC
  - $G[\mathcal{R}]$ is connected since $T$ is connected
  - $\mathcal{R}$ is set cover since $T$ touches each group

CSC to GST reduction: $w(e) = 1$ for all $e \in E(G)$
CSC-R to CST reduction: $k_{g_u} = r_u$

- approximation ratio for GST is $\Omega(\log^2 n)$
- even for trees

NWCSC and NWGST are equivalent

NWGST on trees is reducible to GST

NWCSC is $\Omega(\log^2 n)$ inapproximable

Open problems:

- Polylog approximation of NWCSC
- Log approximation of CSC (or proof of nonexistence)

- approximation ratio for GST is $\Omega(\log^2 n)$
- even for trees

Our conclusion:

- NWGST on trees is reducible to GST
- NWCSC and NWGST are equivalent
- NWCSC is $\Omega(\log^2 n)$ inapproximable
Inapproximability results on NWCSC and open problems

  - approximation ratio for GST is $\Omega(\log^2 n)$
  - even for trees
- our conclusion:
  - NWGST on trees is reducible to GST
  - NWCSC and NWGST are equivalent
  - NWCSC is $\Omega(\log^2 n)$ inapproximable
- open problems:
  - polylog approximation of NWCSC
  - log approximation of CSC (or proof of nonexistence)
Thank you for attention!!!