Approximating Rooted Steiner Networks

Joseph Cheryian\textsuperscript{1}, Bundit Laekhanukit\textsuperscript{2}, Guyslain Naves\textsuperscript{2}, Adrian Vetta\textsuperscript{2}

December 2011, Banff

\textsuperscript{1}Waterloo University

\textsuperscript{2}McGill University
Directed Steiner Tree problem (DST)

A network design problem:

**Input:**
- $G$ a directed graph, with costs $c : E(G) \to \mathbb{N}$,
- $r$ a vertex of $G$ (the root),
- a set $T \subseteq V(G)$ of terminals,

**Output:** A subgraph $G'$ of $G$ such that there is one path from $s$ to $t$ in $G'$, for all $t \in T$

**Goal:** $\min \sum_{e \in E(G')} c(e)$
A network design problem:

**Input:**
- $G$ a directed graph, with costs $c : E(G) \to \mathbb{N}$,
- $r$ a vertex of $G$ (the *root*),
- a set $T \subseteq V(G)$ of *terminals*,
- requirements $k : T \to \mathbb{N}$.

**Output:** A subgraph $G'$ of $G$ such that there are $k_t$ disjoint paths from $s$ to $t$ in $G'$, for all $t \in T$

**Goal:** $\min \sum_{e \in E(G')} c(e)$
Outline

1. $k$-DRC with $O(1)$ terminals.
2. Hardness of $k$-DRC (directed graph).
3. Hardness of $k$-URC (undirected graphs).
4. Integrality gap of $k$-DRC.
Directed Steiner Forest with $O(1)$ terminals

**Theorem (Feldman, Ruhl (2006))**

The Directed Steiner Forest with $O(1)$ terminals is polynomial-time solvable.

**Proof:** Guess nodes of degree $> 2$ and how they are linked, compute shortest paths.

Generalization to Directed Rooted Connectivity?
Proposition

If $G$ is an acyclic digraph and $\sum_{t \in T} k_t = O(1)$, then there is a polynomial-time algorithm.

Proof: Pebbling game (Fortune, Hopcroft, Wyllie).

Open problem: (polynomial or NP-hard?)

$$\sum_{t \in T} k_t = O(1)$$ but $G$ is not acyclic.
Let $\alpha = 2\beta \geq 2$, $k_t^1 = 1$ and $k_t^2 = 2$. 
Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

- Integral solution: $6\beta + 6$
Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$
Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$

Integrality gap: $\frac{6}{5}$
Non-integrality for requirement 3

Let $\alpha = 2\beta \geq 2$, $k_{t_1} = 1$ and $k_{t_2} = 2$.

- Integral solution: $6\beta + 6$
- Fractional solution: $5\beta + 7$

Integrality gap: $\frac{6}{5}$
Toward an APX-hardness proof.

**Theorem (Berman, Karpinski, Scott)**

For every $0 < \varepsilon < 1$, it is \NP-hard to approximate \Max-3-Sat where each literal appears exactly twice, within an approximation ratio smaller than $\frac{1016-\varepsilon}{1015}$. 
Reduction for two terminals

\[ \begin{align*}
\text{X}_1 & \\
\text{X}_2 & \\
\text{X}_3 & \\
\text{X}_4 & \\
\text{X}_n & \\
\end{align*} \]
Analysis (two terminals problem)

Using $\text{OPT}_\phi \geq \frac{7q}{8}$, we get:

$$\rho \geq \frac{13n + (q - \text{APP}_\phi)}{13n + (q - \text{OPT}_\phi)} = 1 + \frac{\text{OPT}_\phi - \text{APP}_\phi}{13n + q - \text{OPT}_\phi}$$

$$\geq 1 + \frac{7 \text{OPT}_\phi - \text{APP}_\phi}{79 \text{OPT}_\phi} = 1 + \frac{7}{79} \left(1 - \gamma^{-1}\right)$$

and finally

$$\rho \geq 1 + \frac{7}{80264} - \xi, \text{ for any } \xi > 0.$$
Analysis (two terminals problem)

Using \( \text{OPT}_\phi \geq \frac{7q}{8} \), we get:

\[
\rho \geq \frac{13n + (q - \text{APP}_\phi)}{13n + (q - \text{OPT}_\phi)} = 1 + \frac{\text{OPT}_\phi - \text{APP}_\phi}{13n + q - \text{OPT}_\phi}
\]

\[
\geq 1 + \frac{7 \text{OPT}_\phi - \text{APP}_\phi}{79 \text{OPT}_\phi} = 1 + \frac{7}{79} \left(1 - \gamma^{-1}\right)
\]

and finally

\[
\rho \geq 1 + \frac{7}{80264} - \xi, \text{ for any } \xi > 0.
\]

Easy \( k \)-approximation when only \( k \) terminals.
Outline

1. $k$-DRC with $O(1)$ terminals.
2. Hardness of $k$-DRC (directed graph).
3. Hardness of $k$-URC (undirected graphs).
4. Integrality gap of $k$-DRC.
General directed rooted connectivity

**Theorem**

The directed and undirected rooted $k$-connectivity problem are at least as hard to approximate as the label cover problem $(2^{\log^{1-\varepsilon} n})$.

**Proof:** Approximation-preserving reduction from Directed Steiner Forest (Dodis, Khanna) (pairs $(s_i, t_i)$ to connect)

Undirected version by a reduction of Lando and Nutov.
Reduction (directed Steiner Forest)
Reduction (directed Steiner Forest)

\[ \text{red arcs cost} = \text{green arcs cost} \]

Diagram showing a network with multiple sources and sinks connected by directed arcs.
Reduction (directed Steiner Forest)

\[ \text{red arcs cost} = 0 \]
Reduction (directed Steiner Forest)

\[ s_1 \rightarrow \cdots \rightarrow s_{k-1} \rightarrow r \]
\[ t_1 \rightarrow \cdots \rightarrow t_{k-1} \rightarrow \]
Reduction (directed Steiner Forest)

Red arcs cost $= 0 = \text{green arcs cost}$
Reduction (directed Steiner Forest)

red arcs cost $= 0 = \text{green arcs cost}$

$\begin{align*}
\text{red arcs cost} &= 0 \\
\text{green arcs cost} &= 0
\end{align*}$
Theorem

The directed rooted $k$-connectivity problem cannot be approximated to within $O(k^\varepsilon)$, for some constant $\varepsilon > 0$, assuming that $\text{NP}$ is not contained in $\text{DTIME}(n^{\text{polylog}(n)})$.

Proof: Reduction from a label cover instance obtained from $\text{Max-3-Sat}(5)$ with $l$ repetition (Chakraborty, Chuzhoy, Khanna).
Label Cover problem

- $G = (U, W, E)$ bipartite graph,
- $L$ set of labels,
- constraints $\Pi_e \subseteq L \times L$ for all $e \in E$,
- assign labels to every vertex to cover every edge $(\forall uw \in E, \Pi_{uw} \cap (f(u) \times f(w)) \neq \emptyset)$,
- minimize the number of labels assigned $\sum_{u \in U \cup W} |f(u)|$.

Instances obtained from $\text{MAX-3-SAT}(5)$ with $l$ repetition:

$$|U| = |W| = O(N^{O(l)}), \quad |L| = 10^l, \quad d = 15^l$$
Reduction from label cover

\[ \text{cost}(\text{label cover}) = 1, \text{cost}(\text{others}) = 0 \]
Reduction from label cover

\[ \text{cost}(\text{label cover}) = 1, \text{cost(other labels)} = 0 \]

\[
\begin{align*}
U & \quad W \\
\{u_1, u_2, u_3, u_4\} & \quad \{w_1, w_2, w_3, w_4\} \\
\end{align*}
\]
Reduction from label cover

\[ \text{cost}(\rightarrow) = 1, \text{cost(others)} = 0 \]
Reduction from label cover

\[
\text{cost}(\rightarrow) = 1, \quad \text{cost}(\text{others}) = 0
\]
Reduction from label cover

\[ \text{cost(\quad)} = 1, \text{cost(others)} = 0 \]
Reduction from label cover

\[
\text{cost( } r \text{ )} = 1, \text{ cost(others) } = 0
\]
Reduction from label cover

\[ \text{cost}(\rightarrow) = 1, \text{cost(others)} = 0 \]
Reduction from label cover

\[ \text{cost}(\rightarrow) = 1, \text{cost(others)} = 0 \]
Reduction from label cover

\[ \text{cost}(\rightarrow) = 1, \text{cost}(\text{others}) = 0 \]
Getting the hardness ratio

Theorem (Parallel repetition theorem, Raz)

There exists a constant $\gamma > 0$ (independent of $l$) such that the minimum total label cover problem obtained from instances of MAX-3SAT(5) with $l$ repetitions cannot be approximated within a factor of $2^{\gamma l}$.

In our reduction, $k = d = 15^l$, hence the $k^\varepsilon$-hardness!
Outline

1. $k$-DRC with $O(1)$ terminals.
2. Hardness of $k$-DRC (directed graph).
3. Hardness of $k$-URC (undirected graphs).
4. Integrality gap of $k$-DRC.
Adapting the reduction to undirected graphs
Adapting the reduction to undirected graphs

\begin{align*}
U & \\
W & \\
\end{align*}

\begin{align*}
r & \\
U & \\
W & \\
\end{align*}

\begin{align*}
t_{2,1} & \\
t_{1,1} & \\
t_{2,2} & \\
t_{3,3} & \\
t_{1,3} & \\
t_{4,4} & \\
t_{3,4} & \\
\end{align*}

\begin{align*}
u_1 & \\
u_2 & \\
u_3 & \\
u_4 & \\
\end{align*}

\begin{align*}
w_1 & \\
w_2 & \\
w_3 & \\
w_4 & \\
\end{align*}
Adapting the reduction to undirected graphs
Forbidding illegal paths
Forbidding illegal paths
Forbidding illegal paths
Forbidding illegal paths
Forbidding illegal paths
Forbidding illegal paths
There are illegal paths,
Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
Very informal description

- There are illegal paths,
- add padding edges to remove illegal paths,
- this creates new illegal paths,
There are illegal paths,
add padding edges to remove illegal paths,
this creates new illegal paths,
add more padding edges to remove the new illegal paths,
There are illegal paths,
add padding edges to remove illegal paths,
this creates new illegal paths,
add more padding edges to remove the new illegal paths,
the second padding set does not induce new illegal paths.
There are illegal paths,
add padding edges to remove illegal paths,
this creates new illegal paths,
add more padding edges to remove the new illegal paths,
the second padding set does not induce new illegal paths.

We are done!
The undirected rooted $k$-connectivity problem cannot be approximated to within $O(k^\varepsilon)$, for some constant $\varepsilon > 0$, assuming that $\text{NP}$ is not contained in $\text{DTIME}(n^{\text{polylog}(n)})$. 

- Improved from $\Omega(\log^{\Theta(1)} n)$,
- Best known approximation ratios are $\tilde{O}(k)$. 
Outline

1. $k$-DRC with $O(1)$ terminals.
2. Hardness of $k$-DRC (directed graph).
3. Hardness of $k$-URC (undirected graphs).
4. Integrality gap of $k$-DRC.
**Theorem**

The natural LP relaxation of the directed rooted $k$-connectivity problem has an integrality ratio of $\Omega \left( \frac{k}{\log k} \right)$.

$$\min \sum_{e \in E} c_e x_e \quad \text{s.t.} \quad \sum_{e \in \delta^+(R)} x_e \geq k \quad (\forall R, r \in R, T \notin R)$$

$$0 \leq x \leq 1$$

**Proof:** we follow a construction of Chakraborty, Chuzhoy, Khanna for \textbf{SNDP} integrality gap.
The construction

\[
\text{cost}(\rightarrow) = 1
\]

\[
\text{cost(others)} = 0
\]

\( k \): connectivity req.

\( q = k \)

\( |A_i| = |B_j| = k^2 \)
The construction

cost(→) = 1
cost(others) = 0

$k$: connectivity req.
$q = k$
$|A_i| = |B_j| = k^2$
The construction

cost(→) = 1
cost(others) = 0

\( k: \) connectivity req.

\( q = k \)

\( |A_i| = |B_j| = k^2 \)
The construction

cost(→) = 1

k: connectivity req.

q = k

|A_i| = |B_j| = k^2
The construction

\[ \text{cost}(\rightarrow) = 1 \]
\[ \text{cost(others)} = 0 \]

\( k: \) connectivity req.

\( q = k \)

\( |A_i| = |B_j| = k^2 \)
The construction

\[
\begin{align*}
\text{cost}(\rightarrow) &= 1 \\
\text{cost(others)} &= 0 \\
q &= k \\
|A_i| &= |B_j| = k^2
\end{align*}
\]
The construction

cost(→) = 1

k: connectivity req.

$q = k$

$|A_i| = |B_j| = k^2$
Computing the gap

- Fractional solution:
  - \( x_e = \frac{1}{k^2} \) for each \( e \in E \) with \( c(e) = 1 \).
  - Total cost: \( 2q = 2k \)
Computing the gap
Computing the gap

**Fractional solution:**
- \( x_e = \frac{1}{k^2} \) for each \( e \in E \) with \( c(e) = 1 \).
- Total cost: \( 2q = 2k \)

**Integral solution:**
- Consider a subset \( S \) of arcs of cost \( \leq \frac{\gamma k^2}{\log k} \),
- prove \( p_S = \Pr[S \text{ is an integral solution}] \) is very very small,
- deduce \( \sum_S p_S < 1 \).
- There is an instance without solution of cost \( \leq \frac{\gamma k^2}{\log k} \).
Computing the gap
Computing the gap
Computing the gap

- Fractional solution:
  - \( x_e = \frac{1}{k^2} \) for each \( e \in E \) with \( c(e) = 1 \).
  - Total cost: \( 2q = 2k \)

- Integral solution:
  - Consider a subset \( S \) of arcs of cost \( \leq \frac{\gamma k^2}{\log k} \),
  - prove \( p_S = \Pr[S \text{ is an integral solution}] \) is very very small,
  - deduce \( \sum_S p_S < 1 \).
  - There is an instance without solution of cost \( \leq \frac{\gamma k^2}{\log k} \).
Computing the gap

- Fractional solution:
  - $x_e = \frac{1}{k^2}$ for each $e \in E$ with $c(e) = 1$.
  - Total cost: $2q = 2k$

- Integral solution:
  - Consider a subset $S$ of arcs of cost $\leq \frac{\gamma k^2}{\log k}$,
  - prove $p_S = \Pr[S \text{ is an integral solution}]$ is very very small,
  - deduce $\sum_S p_S < 1$.
  - There is an instance without solution of cost $\leq \frac{\gamma k^2}{\log k}$.

- Integrality gap is $\Omega \left( \frac{k}{\log k} \right)$
Other result:
- Subset Connectivity problem.

Open questions:
- approximability when $\sum k_i = O(1)$?
- inapproximability when $k = O(1)$? (No better result known than DST)