UNIVERSITÀ DI PAVIA

Università della Svizzera italiana

Lower bounds for the Integrality Gap of the Metric Steiner Tree Problem via a novel formulation

ISCO 2024

Ambrogio Maria Bernardelli Eleonora Vercesi
Janos Barta Luca Maria Gambardella Stefano Gualandi Monaldo Mastrolilli

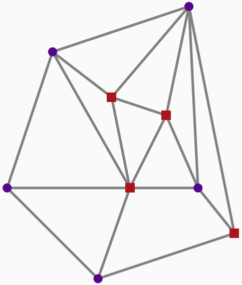
May 24, 2024

Problem definition

Steiner Tree Problem

Given an undirected, edge-weighted graph $G = (V, E)$, $|V| = n$, and a subset of vertices $T \subset V$, $|T| = t$, the Steiner Tree Problem (STP) involves finding the minimum-cost tree that spans T.

The STP is NP-Hard and the corresponding decision problem is NP-Complete [Kar10].

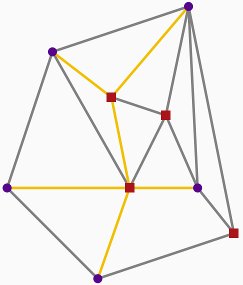


Problem definition

Steiner Tree Problem

Given an undirected, edge-weighted graph $G = (V, E)$, $|V| = n$, and a subset of vertices $T \subset V$, $|T| = t$, the Steiner Tree Problem (STP) involves finding the minimum-cost tree that spans T.

The STP is NP-Hard and the corresponding decision problem is NP-Complete [Kar10].

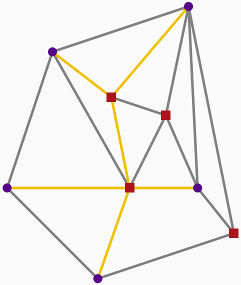


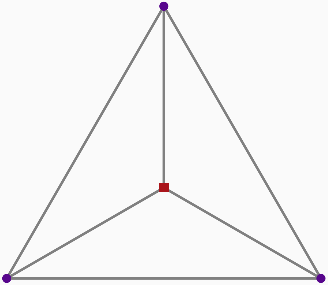
Problem definition

Steiner Tree Problem

Given an undirected, edge-weighted graph $G = (V, E)$, $|V| = n$, and a subset of vertices $T \subset V$, $|T| = t$, the Steiner Tree Problem (STP) involves finding the minimum-cost tree that spans T.

The STP is NP-Hard and the corresponding decision problem is NP-Complete [Kar10].



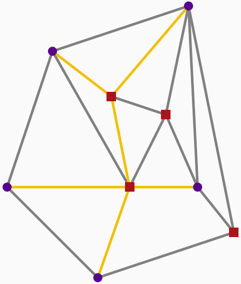


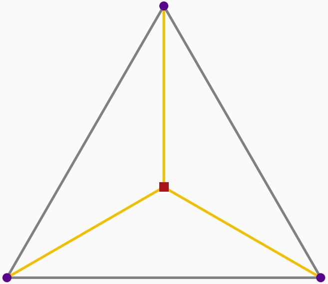
Problem definition

Steiner Tree Problem

Given an undirected, edge-weighted graph $G = (V, E)$, $|V| = n$, and a subset of vertices $T \subset V$, $|T| = t$, the Steiner Tree Problem (STP) involves finding the minimum-cost tree that spans T.

The STP is NP-Hard and the corresponding decision problem is NP-Complete [Kar10].





Integer Linear Programming

Bidirected cut formulation (DCUT)

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & \sum_{e=\{i,j\}\in E} c_e(x_{ij}+x_{ji}) && && \text{(1a)}\\
\text{s.t.} \quad & x_{ij}+x_{ji} \le 1, && e=\{i,j\}\in E, && \text{(1b)}\\
& \sum_{(i,j)\in A\,:\,j\in W, i\notin W} x_{ij} \ge 1, && W\subset V\setminus\{r\},\ W\cap T\neq\emptyset, && \text{(1c)}\\
& x_{ij}\in\{0,1\}, && (i,j)\in A. && \text{(1d)}
\end{aligned}
$$

Integer Linear Programming

Bidirected cut formulation (DCUT)

$$\min_{\mathbf{x}} \quad \sum_{e=\{i,j\}\in E} c_e(x_{ij} + x_{ji}) \tag{1a}$$

$$\text{s.t.} \quad x_{ij} + x_{ji} \leq 1, \qquad e = \{i,j\} \in E, \tag{1b}$$

$$\sum_{(i,j)\in A\,:\, j\in W, i\notin W} x_{ij} \geq 1, \qquad W \subset V \setminus \{r\},\ W \cap T \neq \emptyset, \tag{1c}$$

$$x_{ij} \in \{0,1\}, \qquad (i,j) \in A. \tag{1d}$$

Substituiting Constraint (1d) with $0 \leq x_{ij} \leq 1$ leads to the linear relaxation of DCUT, which we denote with RDCUT. The constraints of RDUCT define the polytope $P_{\text{DCUT}}(n,t)$.

Measuring integerness

Definition 1 (Integrality gap).

The integrality gap is defined as the supremum on all the instances of the ration between the optimal integer value and the linear relaxation:

$$\alpha_{\mathsf{DCUT}}(n,t) = \sup_{\substack{G=(V,E)\in \mathsf{STP} \\ |V|=n, |T|=t}} \frac{\mathsf{DCUT}(G)}{\mathsf{RDCUT}(G)}.$$

We have $36/31 \leq \alpha_{\mathsf{DCUT}} \leq 2$ [BGRS13].

Measuring integerness

Definition 1 (Integrality gap).

The integrality gap is defined as the supremum on all the instances of the ration between the optimal integer value and the linear relaxation:

$$\alpha_{\mathrm{DCUT}}(n,t) = \sup_{\substack{G=(V,E)\in \mathrm{STP} \\ |V|=n,\,|T|=t}} \frac{\mathrm{DCUT}(G)}{\mathrm{RDCUT}(G)}.$$

We have $36/31 \leq \alpha_{\mathrm{DCUT}} \leq 2$ [BGRS13].

Theorem 1.

Given G' the metric closure of G, we have that

$$\frac{\mathrm{DCUT}(G)}{\mathrm{RDCUT}(G)} = \frac{\mathrm{DCUT}(G')}{\mathrm{RDCUT}(G')}.$$

Vertices

One can study the integrality gap through vertices [BB08].

The gap problem

Given $\bar{x}$ vertex of $P_{\mathrm{DCUT}}(n,t)$, we define

$$
\begin{aligned}
\frac{1}{Gap(\bar{x})} = \min\ & c^T\bar{x} \\
\text{s.t. }\ & 0 \le c_{ij} \le c_{ik} + c_{jk}, && \forall \{i,j\},\{i,k\},\{j,k\} \in E, \\
& \mathrm{DCUT}(c) \ge 1,\ \text{Duality constraints, Slackness compatibility conditions.}
\end{aligned}
$$

Vertices

One can study the integrality gap through vertices [BB08].

The gap problem

Given $\bar{x}$ vertex of $P_{\mathrm{DCUT}}(n,t)$, we define

$$
\begin{aligned}
\frac{1}{Gap(\bar{x})} = \min\, & c^T \bar{x} \\
\text{s.t. } & 0 \le c_{ij} \le c_{ik} + c_{jk}, && \forall \{i,j\},\{i,k\},\{j,k\} \in E, \\
& \mathrm{DCUT}(c) \ge 1, \text{ Duality constraints, Slackness compatibility conditions.}
\end{aligned}
$$

Equivalence

$$
\sup_{\substack{\bar{x} \text{ vertex of} \\ P_{\mathrm{DCUT}}(n,t)}} Gap(\bar{x}) = \sup_{\substack{G=(V,E)\in \mathrm{STP} \\ |V|=n,\, |T|=t \\ G \text{ metric and complete}}} \frac{\mathrm{DCUT}(G)}{\mathrm{RDCUT}(G)}.
$$

Vertices problem

For the vertex enumeration, a software like Polymake [AGH+17] can be used. The pipeline described before for the gap problem is then executed as in [VGMG23].

Vertices problem

For the vertex enumeration, a software like Polymake [AGH+17] can be used. The pipeline described before for the gap problem is then executed as in [VGMG23].

n	t	time for vertices generation	# feas problems	gap
4	3	0.04	70/256	1.000
5	3	4563.57	3655/28345	1.000
5	4	2798.17	3645/24297	1.000

Table: Results obtained via Polymake for the DCUT formulation. Number of nodes, number of terminals, time for generating the vertices, number of feasible problems and maximum gap found.

A novel formulation

Complete metric (CM) formulation

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & \sum_{e=\{i,j\}\in E} c_e(x_{ij} + x_{ji}) && && \text{(2a)}\\
\text{s.t.} \quad & x_{ij} + x_{ji} \leq 1, && e = \{i,j\} \in E, && \text{(2b)}\\
& \sum_{\substack{(i,j)\in A\\ j\in W, i\notin W}} x_{ij} \geq 1, && W \subset V \setminus \{r\},\ W \cap T \neq \emptyset, && \text{(2c)}\\
& \sum_{i\neq r} x_{ir} \leq 0, && && \text{(2d)}\\
& \sum_{i\neq j} x_{ij} \leq 1, && j \in V \setminus \{r\}, && \text{(2e)}\\
& 2 \cdot \sum_{i\neq j} x_{ij} - \sum_{k\neq j} x_{jk} \leq 0, && j \in V \setminus T, && \text{(2f)}\\
& x_{ij} \in \{0,1\}, && (i,j) \in A. && \text{(2g)}
\end{aligned}
$$

Properties

n	t	time for vertices generation	# feas problems	gap
4	3	0.732	4/4	1.000
5	3	44.62	5/5	1.000
5	4	37.01	44/44	1.000

Lemma 1.

It holds that the integrality gap of the CM formulation is a lower bound for the integrality gap of the DCUT formulation, i.e.,

$$\alpha_{\mathrm{CM}}(n,t) \leq \alpha_{\mathrm{DCUT}}(n,t).$$

Properties

n	t	time for vertices generation	# feas problems	gap
4	3	0.732	4/4	1.000
5	3	44.62	5/5	1.000
5	4	37.01	44/44	1.000

Moreover, we were able to prove interesting results for the CM formulation, both for integer and fractional vertices, regarding connectedness, number of edges, constraint reduction, and vertex redundancy.

Lemma 1.

It holds that the integrality gap of the CM formulation is a lower bound for the integrality gap of the DCUT formulation, i.e.,

$$\alpha_{\text{CM}}(n,t) \leq \alpha_{\text{DCUT}}(n,t).$$

Theorem 2.

Let x be an integer point of $P_{\text{CM}}(n,t)$. Then x is an optimal solution for the CM formulation with the metric cost $c_{ij} = 2 - (x_{ij} + x_{ji}) \in \{1, 2\}$.

Algorithm One-Two-Costs (OTC) heuristic

$\mathbb{G} = \{G = (V, E) \mid G \text{ connected}, \ |V| = n, \ n \leq |E| \leq n \cdot t - t^2\}$ [MP14]
$\mathbb{T} = \{T \mid T \subset \{1, \ldots, n\}, \ |T| = t\}$
$\mathfrak{G}, \mathcal{V} = \emptyset$
for $G \in \mathbb{G}, \ T \in \mathbb{T}, \ r \in T$ **do**
    $G_{T,r}$ = node-colored graph with G as its support graph, r colored as `root`, i colored as `terminal` $\forall i \in T \setminus \{r\}$, j colored as `steiner` $\forall i \notin T$
    if $H \ncong G_{T,r} \ \forall H \in \mathfrak{G}$ **then**
        add $G_{T,r}$ to $\mathfrak{G}$
    end if
end for
for $G_{T,r} \in \mathfrak{G}$ **do**
    obtain the STP instance (G, T, r) from $G_{T,r}$ with $c_{ij} = 1$ if $\{i, j\} \in G_{T,r}$ and $c_{ij} = 2$ otherwise; solve (2a) - (2f)
    if a solution x is found and it is a non-integer vertex of $P_{\mathrm{CM}}(n, t)$ **then**
        add x to $\mathcal{V}$
    end if
end for

Special vertices

Pure half-integer vertices

Given x a non-integer vertex of $P_{CM}(n,t)$, we say that x is

- half integer (HI) if $x_{ij} \in \{0, 1/2, 1\}$ $\forall (i,j) \in A$,
- pure half integer (PHI) if $x_{ij} \in \{0, 1/2\}$ $\forall (i,j) \in A$.

Special vertices

Pure half-integer vertices

Given x a non-integer vertex of $P_{CM}(n,t)$, we say that x is

- half integer (HI) if $x_{ij} \in \{0, 1/2, 1\}\ \forall (i,j) \in A$,
- pure half integer (PHI) if $x_{ij} \in \{0, 1/2\}\ \forall (i,j) \in A$.

Theorem 3 (PHI theorem).

Let x be a PHI vertex of $P_{\text{CM}}(n,t)$, $t \geq 3$, and let it also be a vertex of $P_{\text{DCUT}}(n,t)$ optimum for a metric cost. Suppose that $x \ncong y$ for every y vertex of $P_{\text{CM}}(n-1,t)$. Define G_x as the support graph of x. In the hypothesis that the indegree of every non-terminal node in G_x is exactly 1, the followings hold:

- the indegree of every terminal in G_x is exactly 2;
- G_x is a connected graph with n nodes;
- G_x has exactly $n+t-2$ edges.

Algorithm PHI heuristic

$\mathbb{G} = \{G = (V, E) \mid G \text{ connected},\ deg(i) \geq 2\, \forall i \in V,\ |V| = n,\ |E| = n + t - 2\}$

di$\mathbb{G}, \mathcal{V} = \emptyset$

for $G = (V, E) \in \mathbb{G}$ **do**

  if $|\{i \in V \mid deg(i) = 2\}| \leq t$ **then**

    add to di$\mathbb{G}$ every non-isomorphic orientation of G s.t. every edge can be oriented in only one way and every node has a maximum indegree of 2

  end if

end for

for di$G = (V, A) \in$ di$\mathbb{G}$ **do**

  $x_{ij} = 1/2$ iff $(i, j) \in A$ is a solution of $P_{\mathrm{CM}}(n, t)$ with

    $\cdot\ \{r\} = \{i \in V \mid indeg(i) = 0\}$

    $\cdot\ V \setminus T = \{i \in V \mid indeg(i) = 1\}$

    $\cdot\ T \setminus \{r\} = \{i \in V \mid indeg(i) = 2\}$

  if x is a feasible vertex of $P_{\mathrm{CM}}(n, t)$ **then**

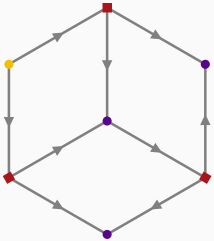
    add x to $\mathcal{V}$

  end if

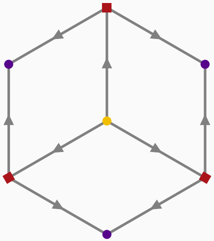
end for

		PHI			OTC		
n	t	# vert.	max gap	# vert. max. gap	# vert.	Gap	# vert. max. gap
	3	0	-	-	0	-	-
6	4	1	1/1	1	0	-	-
	5	7	1/1	7	0	-	-
	3	0	-	-	0	–	
7	4	2	10/9	2	11	10/9	2
	5	46	1/1	46	19	1/1	19
	6	71	1/1	71	8	1/1	8
	3	0	-	-	0	-	-
	4	0	-	-	19	10/9	2
8	5	89	12/11	15	195	10/9	14
	6	1070	1/1	1070	239	1/1	239
	7	758	1/1	758	0	-	-

Some vertices

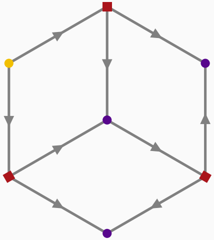


$(n, t) = (7, 4)$, gap $= 10/9$.

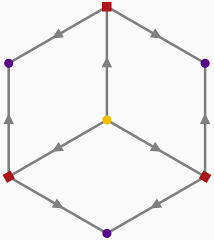


$(n, t) = (7, 4)$, gap $= 10/9$.

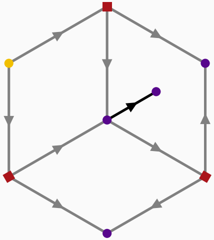
Some vertices



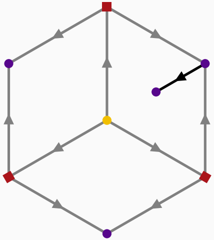
$(n, t) = (7, 4)$, gap $= 10/9$.



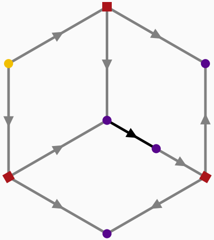
$(n, t) = (7, 4)$, gap $= 10/9$.



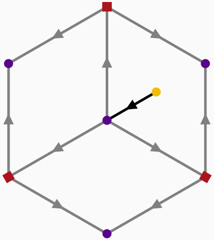
$(n, t) = (8, 5)$, gap $= 10/9$.



$(n, t) = (8, 5)$, gap $= 10/9$.

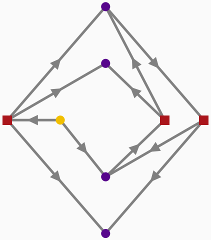


$(n, t) = (8, 5)$, gap $= 10/9$.

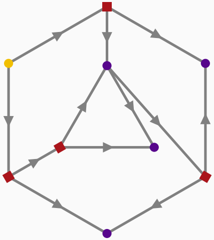


$(n, t) = (8, 5)$, gap $= 10/9$.

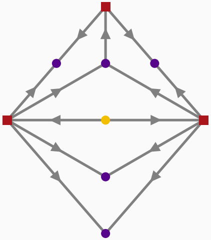
Some more vertices



$(n, t) = (8, 5)$, gap $= 12/11$.

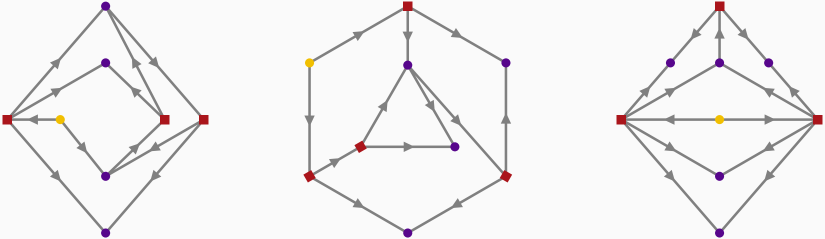


$(n, t) = (9, 5)$, gap $= 10/9$.



$(n, t) = (9, 6)$, gap $= 14/13$.

Some more vertices



$(n, t) = (8, 5)$, gap $= 12/11$. $(n, t) = (9, 5)$, gap $= 10/9$. $(n, t) = (9, 6)$, gap $= 14/13$.

- All the values of integrality gap found by PHI heuristic: $10/9$, $12/11$, $14/13$, $16/15$, $18/17$, $20/19$, $22/21$, $24/23$.
- Note how the PHI(n, t) can be generalized to vertex attaining values in the set $\{0, 1/m\}$ just by changing some parameters.

Conclusions and future work

- A novel and stricter formulation with some interesting properties.
- A problem regarding the maximization of the integrality gap.
- Two heuristics for generating vertices.

Conclusions and future work

- A novel and stricter formulation with some interesting properties.
- A problem regarding the maximization of the integrality gap.
- Two heuristics for generating vertices.

- Exploit the OTC heuristic, adding some constraints derived from numerical experiments or theoretical results.
- Same as above but with the PHI heuristic.
- Prove or disprove some conjectures we made along the way.

Fin.

For other things I do → ambrogiomb.github.io

References

[AGH+17] Assarf, B. and Gawrilow, E. and Herr, K. and Joswig, M. and Lorenz, B. and Paffenholz, A. and Rehn, T., Computing convex hulls and counting integer points with polymake. *Mathematical Programming Computation*, **9**(1), pp. 1-38, 2017.

[BB08] Benoit, G., and Boyd, S., Finding the exact integrality gap for small traveling salesman problems. *Mathematics of Operations Research*, **33**(4), pp. 921–931, 2008.

[BGRS13] Byrka, J., and Grandoni, F., and Rothvoß, T., and Sanità, L., Steiner tree approximation via iterative randomized rounding, *Journal of the ACM*, **60**(1), pp. 1–33, 2013.

[Kar10] Karp, R. M., Reducibility among combinatorial problems, *Springer Berlin Heidelberg*, 2010.

[KPT11] Könemann, J. and Pritchard, D. and Tan, K., A partition-based relaxation for Steiner trees, *Mathematical Programming*, **127**(2), pp. 345-370, 2011.

[Lju21] Ljubić, I., Solving Steiner trees: Recent advances, challenges, and perspectives, *Networks*, **77**(2), pp. 177-204, 2021.

[MP14] McKay, B. D. and Piperno, A., Practical graph isomorphism, II, *Journal of symbolic computation*, **60**(1), pp. 94-112, 2014.

[VGMG23] Vercesi, E. and Gualandi, S. and Mastrolilli, M. and Gambardella, L. M., On the generation of Metric TSP instances with a large integrality gap by branch-and-cut, *Mathematical Programming Computation*, **15**(2), pp. 389-416, 2023.

Properties

Lemma 2.

The support graph of any feasible point of RCM is a connected graph.

Properties

Lemma 2.

The support graph of any feasible point of RCM is a connected graph.

Lemma 3.

Let x be a feasible solution for the CM formulation for a graph with $|V| = n$ nodes and $|T| = t$ terminals. Then x verifies

$$\sum_{i,j} x_{ij} \leq \min(n-1, 2t-3). \tag{3}$$

Let $t \leq \frac{n}{2} + 1$ and so $\min(n-1, 2t-3) = 2t-3$. Then, our solution is a tree with at most $2t-3$ edges, so it has $2t-3+1 = 2t-2$ nodes, with $t-2$ being Steiner vertices. Thus, it suffices to write Constraints (2c) only for

$$W = W_1 \sqcup W_2, \quad W_1 \subset T \setminus r, \; |W_1| \geq 1, \quad W_2 \subset V \setminus T, \; |W_2| \leq t-2, \tag{4}$$

instead of writing it for any $W = W_1 \sqcup W_2$, $W_2 \subset V \setminus T$.

Avoid redundancy

Lemma 4.

Let x be a vertex of $P_{\mathrm{CM}}(n,t)$. Then

$$y_{ij} = \begin{cases} x_{ij}, & \text{if } i,j \neq n+1, \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

is a vertex of $P_{\mathrm{CM}}(n+1,t)$.

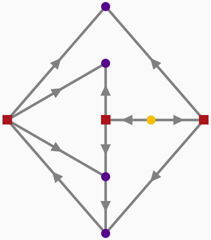
Lemma 5.

Let y be a vertex of $P_{\mathrm{CM}}(n,t)$ of the form

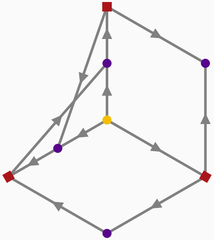
$$y_{ij} = \begin{cases} x_{ij}, & \text{if } i \neq k \neq j, \\ 0, & \text{else,} \end{cases} \tag{6}$$

for a certain $k \in V \setminus T$. Then x is a vertex of $P_{\mathrm{CM}}(n \setminus \{k\}, t) \cong P_{\mathrm{CM}}(n-1,t)$.

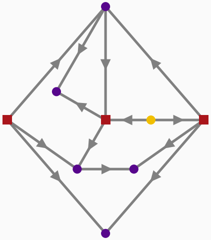
Other vertices



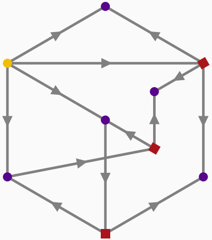
$(n, t) = (8, 5)$, gap $= 14/13$.



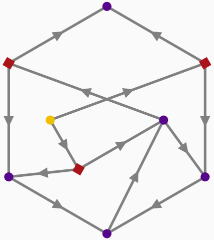
$(n, t) = (8, 5)$, gap $= 18/17$.



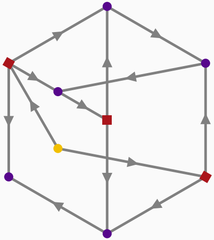
$(n, t) = (9, 6)$, gap $= 16/15$.



$(n, t) = (9, 6)$, gap $= 20/19$.

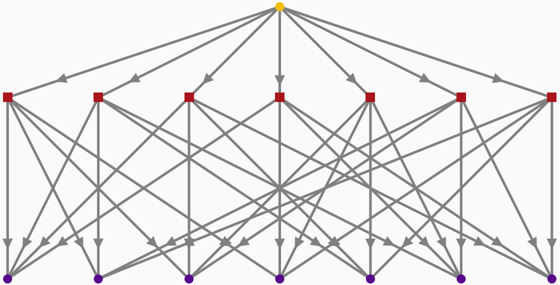


$(n, t) = (9, 6)$, gap $= 22/21$.



$(n, t) = (9, 6)$, gap $= 24/23$.

PHI generalization



If we have $x_{ij} \in \{0, 1/m\}$, the indegree of the terminal nodes must now be m, while the indegree of the Steiner nodes is again 1. This gives us $n + (m-1) \times t - m$ edges. In addition, every node has degree at least $\min(3, m)$; if $m > 3$ the number of nodes with degree 3 is at most $n - t$; there must exist one node of indegree 0, $n - t$ nodes of indegree 1, and $t - 1$ nodes of indegree m. We would have been able to find the vertex above of gap 8/7. [KPT11]