Strong Lower Bounds for a Survivable Network Design Problem

Markus Leitner[1] Günther R. Raidl[2]

Institute of Computer Graphics and Algorithms,
Vienna University of Technology, Vienna, Austria

Abstract

We consider a generalization of the Prize Collecting Steiner Tree Problem on a graph with special redundancy requirements on a subset of the customer nodes suitable to model a real world problem occurring in the extension of fiber optic communication networks. We strengthen an existing connection-based mixed integer programming formulation involving exponentially many variables using a recent result with respect to the orientability of two-node connected graphs. The linear programming relaxation of this model is then solved by means of column generation. We show that our new model is theoretically stronger than a previously described one and present promising preliminary computational results.

Keywords: mixed integer programming, column generation, survivable network design

1 Introduction

We consider a generalization of the Prize Collecting Steiner Tree Problem (PCSTP) on a graph suitable to model the extension of real world fiber optic

[1] Email: `leitner@ads.tuwien.ac.at`
[2] Email: `raidl@ads.tuwien.ac.at`

networks on the last mile. We are given an undirected graph $G = (V, E)$ in which the existing fiber optic infrastructure is represented by a single root node r. Each edge $e = (u, v) \in E$ corresponding to a potential fiber optic route between its end points $u, v \in V$ is given with its length $l_e \geq 0$ and installation cost $c_e \geq 0$. The node set $V = S \cup C \cup \{r\}$ is the disjoint union of Steiner nodes S, customer nodes C with associated prizes $p_k \geq 0$, $\forall k \in C$, – i.e. the expected return of invest when supplying customer k – and the root node r. The set of customers $C = C_1 \cup C_2$ is partitioned into type-1 customer nodes C_1 without specific redundancy requirements and type-2 customer nodes C_2 that need to be redundantly connected by means of two node disjoint paths to the root node r. Since full redundancy is often too expensive and might not pay off we are further given a maximum branch line length $b_{\max}(k) \geq 0$, $\forall k \in C_2$, relaxing above mentioned redundancy requirements: We allow a non-redundant, i.e. single path (*branch line*) from a type-2 customer $k \in C_2$ to some intermediate node $v \in V$ (*branch node*) of maximum length $b_{\max}(k)$ while v in turn must be redundantly connected to the root node. In the following $\mathcal{B}(k)$ denotes the set of potential branch nodes for a customer $k \in C_2$, i.e. those nodes reachable from k by a path no longer than $b_{\max}(k)$, also including node k itself. In the light of this special redundancy concept, we refer to our problem as the $b_{\max}$-*Survivable Network Design Problem* ($b_{\max}$-SNDP).

A solution $G' = (V', E')$, $V' \subseteq V$, $E' \subseteq E$, to an instance of $b_{\max}$-SNDP is a connected subgraph of G feasibly – i.e. respecting the given redundancy requirements – connecting a set of customers $C' \subseteq C$; see Figure 1 for an exemplary solution. Similarly to the PCSTP, we aim at identifying the most profitable solution eventually connecting only a subset of all customers, i.e. we minimize $o(G') = \sum_{e \in E'} c_e + \sum_{k \in C \setminus C'} p_k$. $b_{\max}$-SNDP obviously is NP-hard, since the PCSTP is a special case of it.

2 Previous Work

$b_{\max}$-SNDP has been introduced by Bachhiesl et al. [2]. Ljubić [10] pointed out the relatedness to $\{0, 1, 2\}$-SNDP [7] which corresponds to $b_{\max}$-SNDP

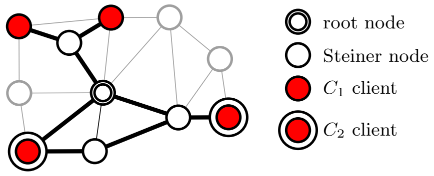

Fig. 1. An exemplary solution to $b_{\max}$-SNDP.

if $b_{\max}(k) = 0$, $\forall k \in C_2$. Wagner et al. presented mixed integer programming (MIP) approaches for $b_{\max}$-SNDP based on multicommodity flows [12] and connection cuts [11]; they are, however, only suitable to solve relatively small instances. The current authors heuristically approached medium-sized instances of $b_{\max}$-SNDP by means of Lagrangian decomposition (LD), variable neighborhood search, greedy randomized adaptive search as well as by hybrid methods combining LD with variable neighborhood descent [8]. Subsequently, we presented a large MIP formulation involving variables for all feasible connections of customers and showed how to efficiently solve its linear relaxation using alternative dual optimal solutions during column generation [9]. Modeling redundant connections by pairs of reversely oriented paths, Chimani et al. [5,4] further came up with strong formulations for $\{0,1,2\}$-SNDP based on multi-commodity flows and directed connection cuts, theoretically dominating those of Wagner et al. [12,11] for the case of $b_{\max}(k) = 0$, $\forall k \in C_2$.

3 The Directed Connection Formulation

Chimani et al. [5] showed that any feasible solution to $\{0,1,2\}$-SNDP can be transformed into a directed graph with a simple path from r to each connected type-1 customer and two oppositely directed, internally node disjoint paths between r and any connected type-2 customer $k \in C_2$. Interpreting a feasible connection to some customer $k \in C_2$ with $b_{\max}(k) > 0$ as two independent connections – a non-redundant from r to k and a fully redundant connection to its branching node $v \in \mathcal{B}(k)$ – the orientability of any solution to $b_{\max}$-SNDP follows from the result of Chimani et al. The model (dCol) introduced in the following strengthens our previous connection-based model from [9] by exploiting this orientability. Let $A = \{(u,v),(v,u) \mid (u,v) \in E\}$ consist of two oppositely directed arcs for each original edge $(u,v) \in E$. To model $b_{\max}$-SNDP we utilize variables $x_{u,v} \in \{0,1\}$, $\forall (u,v) \in A$, indicating whether or not arc $(u,v) \in A$ is part of the (oriented) solution. Variables $y_k \in \{0,1\}$, $\forall k \in C$, specify whether a customer is feasibly connected according to its redundancy requirements or not. We further use variables $f_p^k \in \{0,1\}$, $\forall k \in C$, $\forall p \in P_k$, where P_k is the set of all feasible directed connections for customer $k \in C$, indicating if the corresponding connection is realized or not. For type-1 customers $k \in C_1$, P_k simply corresponds to the set of all simple directed paths from the root node r to k, i.e. $P_k = \{p \subsetneq A \mid p \text{ forms a directed path from } r \text{ to } k\}$. For type-2 customers $k \in C_2$, P_k is defined as $P_k = \{p \subsetneq A \mid p$ forms two oppositely directed, internally node disjoint paths from r to some node $v \in \mathcal{B}(k)$ and a directed path from v to k of maximum length $b_{\max}(k)\}$.

$$
\begin{aligned}
&(1) \qquad \text{(dCol)} \quad z = \min \sum_{(u,v)\in A} c_{u,v} x_{u,v} + \sum_{k\in C} p_k(1-y_k) && \\
&(2) \qquad \text{s.t.} \quad \sum_{p\in P_k} f_p^k \geq y_k && \forall k \in C \\
&(3) \qquad \sum_{p\in P_k \mid (u,v)\in p} f_p^k \leq x_{u,v} && \forall k \in C,\ \forall (u,v) \in A \\
&(4) \qquad x_{u,v} + x_{v,u} \leq 1 && \forall (u,v) \in E \\
&(5) \qquad 0 \leq x_{u,v} \leq 1 && \forall (u,v) \in A \\
&(6) \qquad 0 \leq y_k \leq 1 && \forall k \in C \\
&(7) \qquad f_p^k \in \{0,1\} && \forall k \in C,\ \forall p \in P_k
\end{aligned}
$$

Constraints (2) ensure that a customer's prize can only be earned if it is feasibly connected to r, while constraints (3) link connection variables to arc variables. Inequalities (4) guarantee that at most one out of each pair of oppositely directed arcs is used in a solution. Note that for variables $x_{u,v}$ and y_k only lower and upper bounds are defined in (5) and (6), respectively, as they will automatically become integer.

Since there are exponentially many variables corresponding to feasible connections $F = \{f_p^k \mid k \in C \wedge p \in P_k\}$, we cannot solve (dCol) directly. As usual in column generation – see e.g. [3] – we start with a small subset of connections $\tilde{F} \subsetneq F$ considered in the restricted master problem (RMP), where also the integrality constraints (7) are replaced by their continuous relaxations, and dynamically add further variables $f \in F \setminus \tilde{F}$ by iteratively solving the pricing problem. Let $\mu_k \geq 0$, $\forall k \in C$, be the dual variables associated to constraints (2) and $\pi_{k,u,v} \leq 0$, $\forall k \in C$, $\forall (u,v) \in A$, denote the dual variables associated to constraints (3). Then, the pricing problem is to determine an $f_p^k \in F \setminus \tilde{F}$ corresponding to a connection p with minimum reduced costs $\overline{c}_{k,p} = -\mu_k - \sum_{(u,v)\in p} \pi_{k,u,v}$. As long as at least one variable with negative reduced costs exists, we add it to $\tilde{F}$ and resolve the RMP.

In other words, in the pricing problem we need to determine the cheapest connection to each customer $k \in C$ in $D = (V, A)$ with arc costs $|\pi_{k,u,v}|$, $\forall (u,v) \in A$. If the total costs of such a connection are smaller than μ_k, we include it in the RMP. Since arc costs are non-negative we can efficiently solve the pricing problem for type-1 customers, by simple shortest path calculations. For customers $k \in C_2$ with $b_{\max}(k) = 0$ we need to compute the cheapest pair of oppositely directed, internally disjoint paths (ODP) between r and k. As shown in Figure 2 any instance of the directed disjoint pair of paths problem (2DP) for two source-destination pairs (s_1, t_1), (s_2, t_2), which is known to be NP-hard [6], can be transformed into an instance of ODP for s, t by adding nodes s, t and arcs $\{(s, s_1), (t_2, s), (t_1, t), (t, s_2)\}$. We conclude that ODP as well as the pricing problem for the more general case of customers $k \in C_2$ with $b_{\max}(k) > 0$ are NP-hard.

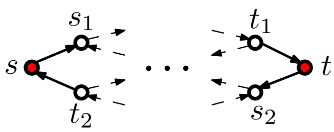

Fig. 2. Transformation of 2DP on (s_1, t_1), (s_2, t_2) into RDP on (s, t).

We solve the pricing problem for each customer $k \in C_2$ using the MIP (8)–(20), where $\mathcal{A}(k) = \{(u, v) \in A \mid u, v \in \mathcal{B}(k)\}$ denotes the set of potential edges in the customer's branch line. Each feasible connection is represented by a directed cycle containing r and at least one potential branching node $w \in \mathcal{B}(k)$ and a path from r to k using arcs not on this cycle for the branch line only. Constraints (9)–(13) ensure that variables $q_{u,v} \in \{0, 1\}$, $\forall (u, v) \in A$, describe such a cycle. Constraints (14)–(16) guarantee that variables $s_{u,v} \in \{0, 1\}$, $\forall (u, v) \in A$, form a path from r to k using arcs not on above mentioned cycle for the branch line only. Finally, constraints (17) ensure that variables $b_{u,v} \in [0, 1]$, $\forall (u, v) \in \mathcal{A}(k)$, indicate the arcs forming the branch line, whereas constraints (18) restrict the branch line's length.

$$\text{(8)} \qquad \min \sum_{(u,v)\in A} |\pi_{k,u,v}| q_{u,v} + \sum_{(u,v)\in \mathcal{A}(k)} |\pi_{k,u,v}| b_{u,v}$$

$$\text{(9)} \qquad \text{s.t.} \sum_{(u,v)\in A} q_{u,v} - \sum_{(v,w)\in A} q_{v,w} = 0 \qquad \forall v \in V$$

$$\text{(10)} \qquad \sum_{(r,v)\in A} q_{r,v} = 1$$

$$\text{(11)} \qquad q_{u,v} + q_{v,u} \leq 1 \qquad \forall (u, v) \in E$$

$$\text{(12)} \qquad \sum_{(u,v)\in A} q_{u,v} \leq 1 \qquad \forall v \in V \setminus \mathcal{B}(k)$$

$$\text{(13)} \qquad \sum_{v\in \mathcal{B}(k)} \sum_{(u,v)\in A} q_{u,v} >= 1$$

$$\text{(14)} \qquad \sum_{(u,v)\in A} s_{u,v} - \sum_{(v,w)\in A} s_{v,w} = \begin{cases} -1 & \text{if } v = r \\ 1 & \text{if } v = k \\ 0 & \text{otherwise} \end{cases} \qquad \forall v \in V$$

$$
\begin{array}{lll}
(15) & s_{u,v} + s_{v,u} \leq 1 & \forall (u,v) \in E \\
(16) & s_{u,v} \leq q_{u,v} & \forall (u,v) \in A \setminus \mathcal{A}(k) \\
(17) & b_{u,v} \geq s_{u,v} - q_{u,v} & \forall (u,v) \in \mathcal{A}(k) \\
(18) & \sum_{(u,v) \in \mathcal{A}(k)} l_{u,v} b_{u,v} \leq b_{\max}(k) & \\
(19) & q_{u,v}, s_{u,v} \in \{0,1\} & \forall (u,v) \in A \\
(20) & 0 \leq b_{u,v} \leq 1 & \forall (u,v) \in \mathcal{A}(k)
\end{array}
$$

4 Polyhedral Comparison

Let $\mathcal{P}_{\text{dcol}}$ denote the polyhedron corresponding to the set of feasible solutions to the linear programming relaxation of model (dCol) and $\mathcal{P}_{\text{col}}$ denote the corresponding polyhedron of its undirected variant (Col) from [9]. Model (Col) uses undirected edge variables $\tilde{x}_e$, $\forall e \in E$, and undirected connection variables $\tilde{f}_p^k$, $\forall k \in C$, $\forall p \in \tilde{P}_k$, but otherwise corresponds to (dCol). It is easy to see that $proj_{\tilde{x},\tilde{f}}(\mathcal{P}_{\text{dcol}}) \subseteq \mathcal{P}_{\text{col}}$ holds, if $proj_{\tilde{x},\tilde{f}}(\mathcal{P}_{\text{dcol}})$ denotes the obvious projection of $\mathcal{P}_{\text{dcol}}$ into the space of $\mathcal{P}_{\text{col}}$. On the other hand, consider the instance given in Figure 3. Its optimal undirected solution $G'_{\text{col}} \in \mathcal{P}_{\text{col}}$ has an objective value of $o(G'_{\text{col}}) = 6.5$ and variables values given by $\tilde{x}_e = 0.5$, for all shown edges e, $y_i = y_h = 1$, and $y_j = 0.5$; however the optimal directed solution $G'_{\text{dcol}} \in \mathcal{P}_{\text{dcol}}$ does not connect any customers. Thus, we conclude that (dCol) dominates its undirected variant, i.e. $proj_{\tilde{x},\tilde{f}}(\mathcal{P}_{\text{dcol}}) \subsetneq \mathcal{P}_{\text{col}}$.

5 Preliminary Computational Results

Table 1 summarizes the results of our preliminary tests on small instances constructed from real world data of a German city [1] which have already been solved to proven optimality for $\{0,1,2\}$-SNDP [5,12,11] – denoted by RED – while not all of them could be solved when using $b_{\max}(k) = 30$, $\forall k \in C_2$, denoted as BMAX [11]. Here, LB_{rel} denotes the average relative improvement

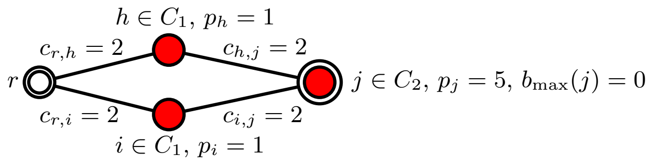

Fig. 3. An exemplary instance of $b_{\max}$-SNDP.

of lower bounds compared to the undirected formulation from [9] in percent and $\#_{\text{opt}}$ the number of instances solved to proven optimality, i.e. where the solution to the linear relaxation is integral. We conclude that the obtained lower bounds are significantly better than those of the undirected formulation from [9] and correspond to proven optimal solutions for all but one of the so far tested instances. While this certainly will not be generally the case for larger instances it indicates that significantly reducing the time needed for solving the pricing problems for type-2 customers might allow for solving significantly larger instances to proven optimality than so far existing approaches.

Table 1
Preliminary Computational Results.

Instance set characteristics						RED		BMAX	
Set	#	$\|V\|$	$\|E\|$	$\overline{C_1}$	$\overline{C_2}$	LB_{rel} [%]	$\#_{\text{opt}}$	LB_{rel} [%]	$\#_{\text{opt}}$
ClgS-I1	25	100	342	3.8	2.1	1.63	25	1.72	25
ClgS-I2	15	100	342	8.9	4.9	9.13	15	12.67	14
ClgS-I3	15	100	342	6.0	3.6	7.09	15	6.94	15

6 Conclusions and Outlook

We presented a new directed mixed integer programming formulation for a generalization of the $\{0, 1, 2\}$-SNDP based on an orientability result by Chimani et al. [5,4]. Our formulation is based on exponentially many so-called connection variables and can be solved using column generation. We theoretically analyzed the pricing subproblems and compared our new formulation to a previously proposed one. Preliminary computational results indicate that the obtained lower bounds are tight. In future work we want to significantly speed up the solution of the linear relaxation by developing heuristics for solving the NP-hard pricing subproblems. Furthermore, we plan to extend our approach to a branch-and-price approach for solving medium sized instances of $b_{\max}$-SNDP to proven optimality.

References

[1] Bachhiesl, P., *The OPT- and the SST-problems for real world access network design – basic definitions and test instances*, Working Report 01/2005, Carinthia

Tech Institute, Department of Telematics and Network Engineering, Klagenfurt, Austria (2005).

[2] Bachhiesl, P., M. Prossegger, G. Paulus, J. Werner and H. Stögner, *Simulation and optimization of the implementation costs for the last mile of fibre optic networks*, Networks and Spatial Economics **3** (2003), pp. 467–482.

[3] Barnhart, C., E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh and P. H. Vance, *Branch-and-price: Column generation for solving huge integer programs*, Operations Research **46** (1998), pp. 316–329.

[4] Chimani, M., M. Kandyba, I. Ljubic and P. Mutzel, *Orientation-based models for 0,1,2-survivable network design: Theory and practice*, Mathematical Probramming, Series B (2009), accepted.

[5] Chimani, M., M. Kandyba and P. Mutzel, *A new ILP formulation for 2-root-connected prize-collecting Steiner networks*, in: *15th Annual European Symposium on Algorithms (ESA'07)*, LNCS **4698** (2007), pp. 681–692.

[6] Fortune, S., J. Hopcroft and H. Wyllie, *The directed subgraph homeomorphism problem*, Theoretical Computer Science **10** (1980), pp. 111–121.

[7] Kerivin, H. and A. R. Mahjoub, *Design of survivable networks: A survey*, Networks **46** (2005), pp. 1–21.

[8] Leitner, M. and G. R. Raidl, *Lagrangian decomposition, metaheuristics, and hybrid approaches for the design of the last mile in fiber optic networks*, in: M. J. Blesa et al., editors, *Hybrid Metaheuristics 2008*, LNCS **5296** (2008), pp. 158–174.

[9] Leitner, M., G. R. Raidl and U. Pferschy, *Accelerating column generation for a survivable network design problem*, in: M. G. Scutellà et al., editors, *Proceedings of the International Network Optimization Conference 2009*, 2009.

[10] Ljubić, I., "Exact and Memetic Algorithms for Two Network Design Problems," Ph.D. thesis, Vienna University of Technology (2004).

[11] Wagner, D., U. Pferschy, P. Mutzel, G. R. Raidl and P. Bachhiesl, *A directed cut model for the design of the last mile in real-world fiber optic networks*, in: B. Fortz, editor, *Proceedings of the International Network Optimization Conference 2007*, Spa, Belgium, 2007, pp. 103/1–6.

[12] Wagner, D., G. R. Raidl, U. Pferschy, P. Mutzel and P. Bachhiesl, *A multi-commodity flow approach for the design of the last mile in real-world fiber optic networks*, in: K.-H. Waldmann and U. M. Stocker, editors, *Operations Research Proceedings 2006* (2007), pp. 197–202.