On the star forest polytope for cactus graphs

Thanh Loan Nguyen[1], Viet Hung Nguyen[1] et Minh Hieu Nguyen[1]

Laboratoire LIMOS, UMR CNRS 7024, INP et Université Clermont Auvergne,

thanh_loan.nguyen@uca.fr, viet_hung.nguyen@uca.fr, minh_hieu.nguyen@uca.fr

Mots-clés : *Combinatorial optimization, polyhedral combinatorics, star forest.*

1 Introduction

This study is dedicated to investigating the Maximum Weight Star Forest Problem (MWSFP). Given an undirected graph $G = (V, E)$ with a weight vector $c \in \mathbb{R}_+^{|E|}$ associated with its edges, a *star* in G is defined as either an isolated node or a subgraph where all edges converge at a common point known as the *center*. A *star forest* refers to a collection of disjoint stars within G. The weight of a star forest is the aggregate sum of the weights of its edges. The objective of the MWSFP is to identify a star forest in G with the maximum possible weight. This problem finds applications in diverse fields such as computational biology [5] and the automobile industry [2]. Notably, the MWSFP is known to be NP-hard [5], as it can be reduced to the minimum dominating set problem.

Within the existing literature, a limited number of scholarly works have delved into examining the polyhedron associated with the MWSFP. To the best of our knowledge, the comprehensive description of the star forest polytope on a tree or a cycle has been exclusively addressed by Aider et al. in their work [1]. Furthermore, Nguyen [4] has introduced a linear-time algorithm for solving the MWSFP in the context of cactus graphs, which represent a generalization of trees and cycles. This instills confidence that a complete description of the MWSFP polyhedron is within reach. As such, the primary goal of this work is to provide a detailed description of the MWSFP polyhedron, highlighting the cactus graph's unique features.

2 Star forest polytope for cactus graphs

A graph G is a *cactus* if each edge belongs to at most one cycle. Let $SFP(G)$ be the polytope of the MWSFP on G. Our strategy to fully characterize $SFP(G)$ draws inspiration from the polytope of the uncapacitated facility location problem (UFLP). This problem entails locating a subset of facilities to minimize the sum of facility opening costs and the costs associated with serving client demands. Consider a directed version $\overrightarrow{G} = (V, A)$ of $G = (V, E)$, where each edge $uv \in E$ corresponds to two directed arcs (u, v) and (v, u) in A, with arc costs given by $c(u, v) = c(v, u) = c_{uv}$. Define a center as a potential facility location, an adjacent node as a client, and an arc (u, v) as an assignment from client u to facility v. A feasible solution of the MWSFP on G precisely aligns with a feasible solution of the $UFLP$ on $\overrightarrow{G}$. Therefore, the projection of $UFLP(\overrightarrow{G})$ onto the variables x_{uv}, where x_{uv} represents the occurrence of edge uv in the maximum spanning star forest, precisely forms $SFP(G)$.

We define the variable $\overrightarrow{x}(u, v) = 1$ if there (u, v) is an assignment of a feasible solution of the $UFLP$ and $\overrightarrow{x}(u, v) = 0$ otherwise. The variable $y(u)$ indicates the presence of facility u in the UFLP solution. Utilizing these variables, the polytope for $UFLP(\overrightarrow{G})$, as showed by Baiou et al. [3], incorporates bidirected cycle inequalities, lifted g-odd cycle inequalities, and assignment inequalities. Our next step is to project these constraints onto the variables x_{uv}.

2.1 Projection of bidirected cycle inequalities

The bidirected cycle inequality [3] has the following form : $x(A(\overrightarrow{C})) \leq \lfloor \frac{2|V(\overrightarrow{C})|}{3} \rfloor$, where $\overrightarrow{C}$ is a bidirected cycle [3] in $\overrightarrow{G}$ and $A(\overrightarrow{C})$ is the set of all arcs in $\overrightarrow{C}$. Let C be the cycle in

G corresponding to a bidirected cycle $\overrightarrow{C}$ in $\overrightarrow{G}$. We observe that $x(A(\overrightarrow{C})) = x(E(C))$ and $|V(\overrightarrow{C})| = |V(C)|$. Thus, we obtain the projection of the bidirected cycle inequalities into the variables x_{uv}, which is *cycle inequalities*.

2.2 Projection of the assignment inequalities

The assignment inequalities have the following forms :

$$y(u) + \sum_{(u,v)\in A} \overrightarrow{x}(u,v) \leq 1,\ \forall u \in V,$$
$$\overrightarrow{x}(u,v) \leq y(v),\ \forall (u,v) \in A.$$

We use the Fourier-Motzkin technique to eliminate the variables $y(u)$ results in the derived inequality $x_{uv} + \sum_{(u,w)\in A, w\neq v} \overrightarrow{x}(u,w) \leq 1, \forall u \in V, uv \in E$. To project these constraints on the variables x_{uv}, we employ a technique called *correct and non-redundant collection*. The results are inequalities defined on cactus subgraphs, where every non-pendant node is connected to precisely one pendant node unless it forms a triangle or is part of a cycle of length four. These subgraphs are termed *MV-cactus* graphs. Then, the projection of the assignment inequalities is called *MV-cactus inequalities*, given by $x(\tau_{MV}) \leq |MV|$. Here τ_{MV} is a MV-cactus, and MV is a bipartite subgraph (U,S) of τ_{MV} such that nodes in part U have a maximum degree of two and part S is connected in τ_{MV}.

2.3 Projection of lifted g-odd cycle inequalities

According Baiou et al.[3], the g-odd cycle inequality takes the form below :

$$\textstyle\sum_{(u,v)\in A(C)} \overrightarrow{x}(u,v) + \sum_{(u,v)\in A'(C)} \overrightarrow{x}(u,v) - \sum_{u\in\hat{C}} y(u) \leq \frac{|\hat{C}|+|\tilde{C}|-1}{2},$$

where C is a g-odd cycle in $\overrightarrow{G}$ associated with three node partitions $\{\hat{C}, \dot{C}, \tilde{C}\}$ [3], $A(C)$ is the set of all arcs in C and $A'(C)$ is a lifting set [3] of C. In a similar manner, we utilize the Fourier-Motzkin technique to eliminate the variables $y(u)$. Afterwards, we identify the characterization of collections of inequalities that are both correct and non-redundant. Our findings reveal that such collections must solely contain a g-odd cycle inequality and assignment inequalities. Then, the resulting projection of the lifted g-odd cycle inequalities is *MV-partition inequalities*, which are of the following format :

$$x(E(C) \setminus P(C)) + 2x(P(C)) + x(\tau_{MVC}) \leq |P(C)| + \tfrac{|C|+|\hat{C}|-1}{2} + |MVC|,$$

where τ_{MVC} is a set of MV-cacti associated with g-odd cycle C, MVC is a set of bipartite graphs associated with τ_{MVC} and $P(C)$ is a subset of $E(C)$ associated with $\{\hat{C}, \tilde{C}\}$.

As a result, we have arrived at the main theorem of this report.

Theorem 1. *Let G be a cactus graph, then the system of the cycle inequalities, the MV-cactus inequalities, and the MV-partition inequalities completely describe $SFP(G)$.*

Références

[1] M. Aider, L. Aoudia, M. Baiou, A. R. Mahjoub, V. H. Nguyen. On the star forest polytope for trees and cycles. Operations Research, 53 (5), 1763-1773, 2019.

[2] A. Agra, D. Cardoso, O. Cerdeira, and E. Rocha. A spanning star forest model for the diversity problem in automobile industry, ECCO XVIII, Minsk, Belarus, May 26-28, 2005.

[3] M. Baiou and F. Barahona, The dominating set polytope via facility location. Combinatorial Optimization. ISCO 2014. Lecture Note Computer Sciences 38–49, 2014.

[4] V. H. Nguyen. The maximum weight spanning star forest problem on cactus graphs. Discrete Mathematics, Algorithms and Applications, World Scientific Publishing, 2015.

[5] C.T. Nguyen, J. Shen, M. Hou, L. Sheng, W. Miller, and L. Zhang, Approximating the spanning star forest problem and its applications to genomic sequence alignment. SIAM Journal on Computing, 946-962, 2008.