Separation Problem for *k*-parashuties

Inna Urazova[1] and Ruslan Simanchev[1,2]

[1] Omsk State University, 55a Mira avenue, 644077 Omsk, Russia
`urazovainn@mail.ru`
[2] Omsk Scientific Center of SB RAS, 15 Marksa avenue, 644024 Omsk, Russia
`osiman@rambler.ru`

Abstract. This article continues the work [16] in which polyhedral setting of graph approximation problem is provided, support inequalities to polytope are built. In this work NP-hard of the separation problem of k-parachutes relative to M-graph polytope is proved.

Keywords: polytope, facet inequality, separation problem.

1 Introduction

Let $\mathbf{K}_n = (V, E)$ be a complete unoriented graph without loops and multipe edges. Spanning subgraph $H \subset \mathbf{K}_n$ is called M-graph if each of its connected components is a clique or single-vertex graph. We denote a set of all M-graphs in $\mathbf{K}_n$ through $\mu(V)$.

Let $G \subset \mathbf{K}_n$ be some a priori set spanning subgraph. Approximation problem of G consists in finding M-graph H minimizing the functional

$$\rho_G(H) = |EG \cup EH| - |EG \cap EH| \tag{1}$$

on set $\mu(V)$. Here EG and EH are sets of edges of G and H, respectively. In other words, it is necessary to find such a set of pairwise non-overlapping cliques at V which is as far as possible less (in terms of edges) different from G.

Harary was the first to state graph approximation problem in [10] in 1955. In the 1960s-1970s in a number of works non-trivial classes of graphs on which the problem is polynomially solvable [7, 18] . In 1986 Krivanek and Moravek considered graph approximation problem as a particular case of the tree clusterization and proved that it was NP-hard [11] . Systematic studies of graph approximation problem began in the last decade when the problem was re-discovered under different names (correlation clustering; Cluster editing) by different groups of authors [3, 4, 13]. In particular, the NP-hardness of its various options was established [1, 3, 13] and suggested first approximate algorithms with the guaranteed accuracy evaluation [3, 5, 8]. The best of the currently known graph approximation algorithms finds a guaranteed solution by factor of max. 2.5 worse than the optimum one [2, 17].

In this work which continues the work [16] we consider polyheldral setting of the graph approximation problem. The polyhedral approach to the solution of extreme combinatorial problems consists in the correlation of the problem with a special polytope set as a convex hull of incidence vectors of the admissible solutions and, consequently, the use of convex analysis and integer programming means. In this route a special role belongs to the task of polytope description as a set of solutions to the system of linear equations and inequalities. If a full linear description of the polytope is available, the extreme combinatorial problem reduces to the linear programming problem (possibly - with exponential number of constraints) which quite often enables obtaining effective algorithms for the solution thereof. One of the most widely known examples of such a situation is problem of maximum weighted matching [6, 12]. As a rule, for NP-hard problems a full liner description of the polytope is not known (see, e.g., [9]). Nevertheless, the availability of a partial description, i.e., classes of valid, support or facet inequalities enables building evaluations of the target function optimum value, develop special cutting plane procedures.

To state the results of this article, we will introduce the following designations and notions. For any graph $D \subset \mathbf{K}_n$ via VD and ED we will denote a set of its vertices and edges, respectively. For an edge $e \in E$ we will also use uv notation, where u and v are vertices from V incident to edge e. For $D \subseteq \mathbf{K}_n$ and $u \in V$ via $\delta_D(u)$ we will denote the set of edges of the D incident to vertex u. If $D = \mathbf{K}_n$, then in this notation we will omit index D. Each set of edges $R \subset E$ induces in $\mathbf{K}_n$ some subgraph T in which $ET = R$ and VT is a set of vertices incident to edges from R. If no ambiguity arises, a graph induced by a set of edges R will be denoted via R. For subgraphs $D, F \subseteq \mathbf{K}_n$ we assume

$$D \cup F = ED \cup EF, D \cap F = ED \cap EF.$$

We will correlate graph $\mathbf{K}_n$ with a Euclidean space $\mathbf{R}^E$ of dimensions $\frac{n^2-n}{2}$ by correlating each edge with a coordinate axis in $\mathbf{R}^E$. This space may be considered as a space of column vectors with the components indexed by elements from E. If $x \in \mathbf{R}^E$ and $R \subseteq E$, then we denote linear form $\sum_{e \in R} x_e$ via $x(R)$. The incidence vector of the arbitrary graph $D \subseteq \mathbf{K}_n$ without isolated vertices is called vector $x^D \in \mathbf{R}^E$ with components $x_e^D = 1$ at $e \in ED$ and $x_e^D = 0$ at $e \notin ED$. The latter rule, obviously, sets a one-to-one correspondence between the set of all subgraphs without isolated vertices of $\mathbf{K}_n$ and the set of vertices of the unit cube in $\mathbf{R}^E$.

A set $P \subset \mathbf{R}^E$ is called a polytope if it is a convex hull of a finite number of points. Linear inequality $a^T x \leq a_0$ ($a, x \in \mathbf{R}^E, a \neq 0, a_0 \in \mathbf{R}$, "$T$" transposition sign) is called support to P if it is fulfilled for any point from P and there exists at least one point from P converting it into an equality. Any inequality support to P generates the set $\{x \in P | a^T x = a_0\}$ which is called the face of the polytope P. Maximum by the inclusion faces of the polytope are called facets. It is clear that a facet is the face and only the face the dimensionality of which is by 1 smaller than the dimensionality of the polytope itself. A support inequality generating a facet will, consequently, be referred to as a facet inequality.

Facet inequalities are present (with the equivalency accuracy) in any linear description of the polytope [14, 12]. Besides, they proved to be good as cutting planes for the solution of high dimensionality problems (see, e.g., [15, 9]).

During the development of cutting plane procedures using support inequalities the separation problem goes to the foreground. It consists in the following. What is given is a class of inequalities L support to polytope P and point $\bar{x} \in \mathbf{R}^E$. It is required to find an inequality in the class L strictly separating the point $\bar{x}$ from polytope P, or prove that in L such an inequality does not exist. Examples of extreme combinatorial problems and inequality classes for which the separation problem is polynomially solvable, for example, travelling salesman problem and subtour elimination inequalities [9], the problem of building a multi-processor schedule and inequality class induced by the paths in the precedence graph [15] etc. are known. In this work NP-hardness of the separation problem of k-parachute class inequalities relative to M-graph polytope is proved.

2 M-graph polytope and k-parachutes

We will refer to the set

$$P_n = conv\{x^H \in \mathbf{R}^E | H \in \mu(V)\}$$

as an M-graph polytope. Here "conv" means "convex hull".

In [16] it was prove that $(0, 1)$-vector $x \in \mathbf{R}^E$ is the vector of incidences of M-graph if and only if it satisfies the system:

$$\begin{aligned} -x_{uv} + x_{uw} + x_{vw} &\leq 1 \\ x_{uv} - x_{uw} + x_{vw} &\leq 1, \\ x_{uv} + x_{uw} - x_{vw} &\leq 1 \end{aligned} \quad (2)$$

where $u, v, w \in V$ are all kinds of sets of three pairwise different vertices,

$$x_{uv} \geq 0 \ for \ all \ uv \in E. \quad (3)$$

Besides, in the same work it was demonstrated that the target function (1) in terms of $\mathbf{R}^E$ may be written as

$$\rho_G(x) = |EG| + x(E\bar{G}) - x(EG). \quad (4)$$

Therefore, graph approximation problem may be stated as a problem of integer linear programming (2)–(4).

Let $U = \{u_1, u_2, \ldots, u_k\}$ and $W = \{v_1, v_2, \ldots, v_p\}$ be nonempty subsets of set V, $U \cap W = \varnothing$, $k \geq 1$, $p \geq 2$. We will designate star in $\mathbf{K}_n$ with the centre in vertex u_i and arms $u_i v_j, j = 1, 2, \ldots, p$ through $T_i, i = 1, 2, \ldots, k$. We will denote the clique on the set of vertices W through K_p. Let us assume that $T = \bigcup_{i=1}^{k} T_i$. Graph $T \cup K_p$ is called a k-parachute. We will correlate inequality

$$x(ET) - x(EK_p) \leq \frac{k^2 + k}{2}.$$

with this graph. In [16] it was proved that this inequality induced by k-parachute $T \cup K_p$ is support inequality to polytope P_n if and only if $p \geq k$, and facet inequality if and only if $k = 1$.

3   Separation Problem

Let L be a class of support inequalities to polytope P_n induced by k-parachutes. Whereas polytope P_n completely lies in the unit cube of $\mathbf{R}^E$, we will state the separation problem stronger than it was done in the Introduction.

   Problem A. Let $\bar{x} \in \mathbf{R}^E$ and $0 \leq \bar{x} \leq 1$. Does L class contain such an inequality for which $\bar{x}(ET) - \bar{x}(EK_p) > \frac{k^2+k}{2}$?

   We need auxiliary fact, which is easily proved by induction.

Lemma 1. *If any t edges with $t < n$ are withdrawn from the complete n-vertex graph, the remaining graph will contain clique of the order $n - t$.*

Theorem 1. *The separation problem for the inequalities induced by k-parachutes relative to the polytope of the graph approximation problem is NP-hard.*

Proof. Let us prove that this problem is NP-hard already at $k = 1$, i.e., at $U = \{u\}$. In this case Problem A may be re-stated as follows:

   Problem B. Let $\bar{x} \in \mathbf{R}^E$ and $0 \leq \bar{x} \leq 1$. Is there among inequalities of the type $\sum_{j=1}^{p} x_{uv_j} - \sum_{i=1}^{p-1} \sum_{j=2, j>i}^{p} x_{v_i v_j} \leq 1$ such an inequality which is violated by point $\bar{x}$?

   It is clear that in this problem u vertex may be fixed. Now at preset point $\bar{x} \in \mathbf{R}^E$, $0 \leq \bar{x} \leq 1$ and vertex $u \in V$ we will define vector $c \in \mathbf{R}^E$:

$$c_e = \begin{cases} \bar{x}_e, & e \in \delta(u); \\ -\bar{x}_e, & e \in E \setminus \delta(u). \end{cases}$$

   We will denote this edge weighted graph via $\mathbf{K}_n(c, u)$. It is not difficult to see that Problem B is equivalent to the following problem.

   Problem C. Does this edge weighted graph $\mathbf{K}_n(c, u)$ contain such clique K that $c(EK) > 1$?

   Let us reduce CLIQUE problem stated as follows to Problem C: what is given is graph G and natural number $s < |VG|$. Does graph G contain clique of the order larger than s?

   So, let what is given be graph G, $|VG| = m$ and a natural number $s < m$. Let us assume that $n = m + 1$ and build a complete graph $\mathbf{K}_n$ on a set of vertices $VG \cup \{u\}$ where u is an added vertex. Let us assign the following weights to $\mathbf{K}_n$ edges:

$$c_e = \begin{cases} \frac{1}{s}, & e \in \delta(u); \\ 0, & e \in EG; \\ -\frac{1}{s}, & e \in E\bar{G}, \end{cases}$$

where $\bar{G}$ is the complement of graph G.

Let us note that if in graph G there is clique K of the order larger than s, then for the clique $K' \subset \mathbf{K}_n(c, u)$ on the set of vertices $VK \cup \{u\}$ we have $c(EK') > 1$.

Let us suppose that in graph $\mathbf{K}_n(c, u)$ there is clique $\bar{K} \subset \mathbf{K}_n(c, u)$ such that $c(E\bar{K}) > 1$. Because only edges from $\delta(u)$ have positive weights, then necessarily $u \in V\bar{K}$. It means that clique $\bar{K}$ may be presented as $\bar{K} = \delta_{\bar{K}}(u) \cup K$ where K is clique in $G \cup \bar{G}$. Clique K, in it turn may be represented as $K = (K \cap G) \cup (K \cap \bar{G})$. Therefore, $\bar{K} = \delta_{\bar{K}}(u) \cup (K \cap G) \cup (K \cap \bar{G})$. Because edge sets of these three graphs do not overlap pairwise, then

$$c(E\bar{K}) = c(\delta_{\bar{K}}(u)) + c(E(K \cap G)) + c(E(K \cap \bar{G})) =$$
$$= \frac{1}{s}|VK| - \frac{1}{s}|E(K \cap \bar{G})| = \frac{1}{s}(|VK| - |E(K \cap \bar{G})|) > 1.$$

Hence, particularly, it follows that $|VK| > |E(K \cap \bar{G})|$. Let us note that the withdrawal of edges $E(K \cap \bar{G})$ from clique K gives exactly $K \cap G$ graph which, first, lies fully in G, and second, by virtue of the lemma, contains clique of the order $|VK| - |E(K \cap \bar{G})| > s$.

The theorem is proved.

4   Conclusion

This article continues the work [16] in which facet inequalities to the graph approximation problem polytope is built. The next step in this direction is the inclusion of the received facet inequalities into polyhedral type algorithms. The effectiveness of this inclusion depends essentially on the computational difficulties of the separation problem of inequalities. In this work NP-hardness of the separation problem of k-parachutes relative to M-graph polytope is proved. The most natural direction for further research is the development of heuristics to separation problem.

In conclusion, we will announce preliminary results of our numerical experiment. The aim of the experiment is the evaluation of the use of 1-parachutes in cutting plane and branch-bound algorithms to solve the graph approximation problem. Below are brief results which are as follows. We solve two integer linear programm. The first problem have function (4) as an objective function and have polyhedron (2)–(3) as a relaxation set. In the second problem, we change the polyhedron, adding to the constraints (2)–(3) all 1-parachute inequalities. This, of course, imposes serious restrictions on the dimension of the problem. To solve this problem we used IBM ILOG CPLEX Optimization Studio package and Intel(R) Celeron(R) CPU N2830 2.16GHz. Solution time was limited to 2 hours. More than a hundred problems with different objective functions of the form (4) for $n \leq 25$ were solved. Table 1 shows the average solving time for the first type and second type problems (columns SR and PR, respectively). As the table shows, the addition of 1-parachutes greatly reduces the time to solve the problem.

Table 1.

n	Number of variables	SR(mean time, sec)	PR(mean time, sec)
10	45	0.16	0.23
15	105	0.66	1.99
20	190	46.91	35.35
25	300	6566.36	1464.68

References

1. Ageev A.A., Il'ev V.P., Kononov A.V., Talevnin A.S.: Computational complexity of the graph approximation problem. J. of Applied and Industrial Mathematics. vol. 1, Issue 1, 1–8 (2007)
2. Ailon N., Charikar M., Newman A.: Aggregating inconsistent information: Ranking and clustering. J. ACM. 55, 5, 1–27 (2008)
3. Bansal N., Blum A., Chawla S.: Correlation clustering. Machine Learning. 56, 89–113 (2004)
4. Ben-Dor A., Shamir R., Yakhimi Z.: Clustering gene expression patterns. J. Comput. Biol. 6, 3–4, 281–297 (1999)
5. Charikar M., Guruswami V., Wirth A.: Clustering with qualitative information. J. Comput. Syst. Sci. 71, 3, 360–383 (2005)
6. Edmonds J.: Maximum matching and a polyhedron with 0,1 - vertices. J. of Research National Bureau of Standards. Section B, 69, 125–130 (1965)
7. Fridman G. Sh.: A Problem of Graph Approximation (Russian). Upravlyaemye Sistemy. 8, 73–75 (1971)
8. Giotis I., Guruswami V.: Correlation clustering with a fixed number of clusters. Theory of Computing. 2, 1, 249–266 (2006)
9. Grotschel M., Holland O.: Solution of large-scale symmetric travelling salesman problems. Mathematical Programming. 51, 2, 141–202 (1991)
10. Harary, F.: On the notion of balance of a signed graph. Michigan Mathematical Journal. 2, 143–146 (1955)
11. Křivánek M., Morávek J.: NP-hard problems in hierarchical-tree clustering. Acta informatica. 23, 311–323 (1986)
12. Schrijver, A.: Combinatorial Optimization. Polyhedra and Efficiency. Springer, Heidelberg (2004)
13. Shamir R., Sharan R., Tsur D.: Cluster graph modification problems. J. Discrete Applied Mathematics. 144, 1–2, 173–182 (2004)
14. Simanchev R.Yu.: Convex Polytope and Facet inequalities (Russian). Omsk State University, Omsk (1999)
15. Simanchev R.Yu., Urazova I.V.: Scheduling unit-time jobs on parallel processors polytope (Russian). Diskretnyi Analiz i Issledovanie Operatsii. 18(11), 85–97 (2011)
16. Simanchev, R.Yu., Urazova, I.V.: On the Polytope Faces of the Graph Approximation Problem. J. of Applied and Industrial Mathematics. vol. 9, 2, 283–291 (2015)
17. van Zuylen A., Williamson D.P.: Deterministic Pivoting Algorithms for Constrained Ranking and Clustering Problems. Mathematics of Operations Research. 34, 3, 594–620 (2009)
18. Zahn C. T.: Approximating symmetric relations by equivalence relations. J. of the Society for Industrial and Applied Mathematics. 12, 4, 840–847 (1964)