Minimum Hitting Sets and Beyond

Professor W. Zang, Department of Mathematics

Background. Let X be a finite set with a nonnegative integral weight $w(x)$ on each element x, and let $\mathcal{E}$ be a family of subsets of X. A *hitting set* of $\mathcal{E}$ is a subset Y of X such that $Y \cap Z \neq \emptyset$ for all $Z \in \mathcal{E}$. The *hitting set problem* (HSP) is to find a hitting set of $\mathcal{E}$ with minimum total weight. As shown by Chvátal, this *NP*-hard problem is approximable within a factor of $\ln|X|+1$ using a greedy algorithm. Feige further proved that $(1-o(1))\ln|X|$ is a threshold below which the HSP cannot be approximated efficiently, unless *NP* has slightly superpolynomial time algorithms. So the focus of this project is on some fundamental combinatorial optimization problems contained in the HSP as special cases.

Observe that if X is the arc (resp. vertex) set of a directed graph G and $\mathcal{E}$ is the family of all cycles (directed) in G, then a hitting set of $\mathcal{E}$ is actually a *feedback arc* (resp. *vertex*) *set* of G; that is, an arc (resp. a vertex) subset X of G such that $G \backslash X$ contains no cycles. Thus the HSP reduces to the *feedback arc* (resp. *vertex*) *set problem*. If X is the vertex set of an undirected graph G and $\mathcal{E}$ is the family of all odd cycles in G, then the HSP amounts to the problem of finding a minimum-weight vertex subset X in G such that $G \backslash X$ is bipartite, which is often called the *graph bipartization problem*. The feedback set and graph bipartization problems arise in a rich variety of applications and have great theoretical interest in their own right, so they have attracted tremendous research efforts over the past five decades; they are also the major concern of this project. Prof. Zang proposes to study them and a closely related one, also of hitting-set type, using the linear programming and structure-driven approaches, establish some important and beautiful min-max relations enjoyed by them, and devise efficient combinatorial algorithms for solving them accordingly.

Objectives

(i) To characterize all tournaments $T = (V, A)$ such that the maximum size of a cycle packing is equal to the minimum total weight of a feedback arc set in T for any weight function $\boldsymbol{w} \in \mathbb{Z}_+^A$;

(ii) To characterize all tournaments $T = (V, A)$ such that the maximum size of a feedback arc set packing is equal to the minimum total weight of a cycle in T for any weight function $\boldsymbol{w} \in \mathbb{Z}_+^A$;

(iii) To show that the minimum total weight of a feedback vertex set of a planar digraph $G = (V, A)$ is at most five times the maximum size of a cycle packing for any weight function $\boldsymbol{w} \in \mathbb{Z}_+^V$; and

(iv) To characterize all graphs $G = (V, E)$ such that the maximum size of an odd cycle packing is equal to the minimum total weight of an odd cycle cover in G for any weight function $\boldsymbol{w} \in \mathbb{Z}_+^V$.

Research Findings

(i) A tournament $T = (V, A)$ is called *cycle Mengerian* (CM) if for any nonnegative integral function w defined on A, the maximum number of cycles in G such that each arc e is used at most $w(e)$ times is equal to the minimum of $\sum\{w(e) : e \in B\}$, where the sum is taken over all $B \subseteq A$ such that deleting B from T results in an acyclic digraph. As part of this GRF project, Prof. Zang and his collaborators have successfully characterized all CM tournaments; their proof also yields polynomial-time algorithms for the corresponding optimization problems.

(ii) Let $G = (V, E)$ be a graph. The matching polytope of G, denoted by $P(G)$, is the convex hull of the incidence vectors of all matchings in G. As proved by Edmonds in 1965, $P(G)$ is determined by the following linear system $\pi(G)$:

- $x(e) \geq 0$ for each $e \in E$;
- $x(\delta(v)) \leq 1$ for each $v \in V$;
- $x(E[U]) \leq \lfloor \frac{1}{2}|U| \rfloor$ for each $U \subseteq V$ with $|U|$ odd.

In 1978, Cunningham and Marsh strengthened this theorem by showing that $\pi(G)$ is always totally dual integral. In 1984, Edmonds and Giles initiated the study of graphs G for which $\pi(G)$ is box-totally dual integral. As part of this GRF project, Prof. Zang and his collaborators have obtained a structural description of all such graphs, and developed a general and powerful method for establishing box-total dual integrality.

Further Reading

(i) X. Chen, G. Ding, W. Zang, and Q. Zhao, *Ranking Tournaments with No Errors*, in preparation.

(ii) G. Ding, L. Tan, and W. Zang, *When is the Matching Polytope Box-totally Dual Integral?* submitted to Mathematics of Operations Research.