MAX-PLANCK-INSTITUT FÜR INFORMATIK

Minimum Base of Weighted k–Polymatroid and Steiner Tree Problem

A. Zelikovsky

MPI–I–92–121 June 1992

INFORMATIK

Im Stadtwald
66123 Saarbrücken
Germany

Minimum Base of Weighted k–Polymatroid and Steiner Tree Problem

A. Zelikovsky

MPI–I–92–121 June 1992

"Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie (Betreuungskennzeichen ITS 9103) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor."

Minimum Base of Weighted k-Polymatroid and Steiner Tree Problem

Alexander Zelikovsky

Max-Plank-Institut für Informatik
Im Stadtwald, 6600 Saarbrücken
Germany
Institute of Mathematics of Moldavian Academy of Sciences
Akademicheskaya 5, Kishinev, 277024 Moldova

May 20, 1992

Abstract

A generalized greedy approximation algorithm for finding the lightest base of a weighted k-polymatroid and its applications to the Steiner tree problem is presented.

1 Introduction

We consider k-polymatroids which appear from matroids in the following way. Let $M = (X, r)$ be a matroid with a rank function r. Given a family P of closed subsets of X, called *coalitions*, define the *rank function* $f : 2^P \to N$ such that

$$f(W) = r(\bigcup_{p\in W} p)$$

Let every element of P has rank at most k, then (P, f) is k-polymaroid. The set $\cup\{p|p \in W\}$ will be denoted W^*

This is the well-known definition [5]. Now our goal is to assign weights to elements of k-polymatroid (P, f) in the same way as for elements of the underlying matroid (M, r) and to compare these weights. So further we assume, that $X \subset P$.

Definition. *A weighted $k-$polymatroid $M_P = (P, f, d)$, is defined to be a $k-$polymatroid (P, f) with a weight function $d : 2^P \to R^+$, such that*

$$d(\bigcup_{p \in W)} p) = \sum_{p \in W)} d(p),$$

for any $W \subset P$.

Let W be a subset of a weihgted polymatroid (P, f, d). We define the *span* of W to be the largest subset of P which includes W and has the same rank as W. We say that W is spanning if it spans the whole set P, that is, if $f(W) = f(P) = r(X)$, since we put $X \subset P$. We deal with a problem of finding a minimum weighted spanning set (i.e. *base*) of a weighted $k-$polymatroid (WPP). The weight of the WPP solution will be denoted by $w(M_P)$. If $k = 1$, then WPP requires the lightest base of a matroid. The unweighted variant of this problem includes the famous Matroid Parity Problem which requires a maximum set of coalitions $W \subset P$ such that W^* is independent, where P consists of all elements of a given partition of X into subsets of cardinality two. Since MPP is NP-hard (see [4]), WPP is also NP-hard even for $k = 2$. But the unweighted problem has an exact polynomial solution if the matroid M is linear [5]. For $k = 3$, the unweighted problem is NP-hard even for linear underlying matroids [3]. So approximation algorithms for this problem may be investigated.

Since $X \subset P$, we may consider greedy algorithm for underlying weighted matroid M (GAM) as an approximation algorithm for WPP. Let $w(M)$ denote the weight of the resulted spanning set. It is known that it is the best solution if $P = X$.

In this paper, we present an approximation algorithm for WPP which generelizes and improves GAM by consideration of coalitions with rank more than one (Section 1). Then an approximation bound of this algorithm is considered in Section 2. The last section is devoted to applications of this algorithm to the Steiner Tree Problem.

2 Algorithm

Let $M = (X, r, d)$ be the weighted underlying matroid for M_P, *to contract* an element $\{e\}$ means to reduce its weight to zero. We use $M[e] = (X, r, d')$ to denote the resulting weighted matroid. For any coalition p, define the matroid $M[p]$ to be the result of contraction of all elements of p.

Our algorithm goes as follows:

Algorithm.
(1) $F \leftarrow M; W \leftarrow \emptyset$
(2) *repeat forever*
  (a) *find* $p \in P$ *which maximizes*
    $win = w(F) - w(F[p]) - d(p)$
  (b) *if* $win \leq 0$, *then exit repeat*
  (c) $F \leftarrow F[p]; W \leftarrow W \cup \{p\}$
(3) *Find* B, *a minimum base of* F, *using GAM*
    $B^* \leftarrow B \cup W$

In other words, Algorithm chooses a coalition, which can decrease the weight of base produced by GAM by the value of win. Then the best coalition is contracted and Algorithm is ready to look for the next such coalition. Note, that the choice of a coalition is greedy, i.e. every time it looks for the greatest win. So we can consider this Algorithm as generalized greedy algorithm.

3 The approximation bound

The following theorem compares approximations given by the usual greedy algorithm and generalized one and shows that the latter is better.

Theorem 1. *Let M_P be a weighted k-polymatroid, M be its underlying matroid and B^* be a base resulted by Algorithm. Then*

$$nd(B^*) \leq w(M_P) + (n-1)w(M)$$

where n is the maximum cardinality of a coalition in W.

Proof. Let $F = (P, f, d)$ be a a weighted k-polymatroid matroid and F be its underlying weighted matroid. For a subset $A \subset X$, we define $F[A]$

recursively: $F[\emptyset] = F$, and $F[A \cup e] = F[A][e]$. (For brevity, we denote a singleton $\{x\}$ as x.)

Let $B = \{b_1, ..., b_r\}$ be a "greedy" order of a minimum base of F, i.e., $d(b_i) \leq d(b_{i+1})$ for all $i = 1, ..., r-1$. Put $B_0 = \emptyset$, $B_i = \{b_1, ..., b_i\}$, $F_i = F[B_i]$ for $i = 1, ..., r$.

Let $e \in X$ and b_k be an element of the maximum index such that $e \cup B - b_k \in I$. The index k will be denoted by $ind(e)$. It is obvious, that the set $e \cup B_{k-1}$ can be chosen by the greedy algorithm by the first k steps. Since the unique cycle in $e \cup B$ belongs to $e \cup B_k$, the closure of $e \cup B_{k-1}$ coincides with the closure of B_k. Thus, $e \cup B - b_k = \{e, b_1, ..., b_{k-1}, b_{k+1}, ..., b_r\}$ is the greedy order of a minimum base of $F[e]$.

For a set $A \subset X$, define $s_F(A) = w(F) - w(F[A])$. Further we will often use the following obvious equality

$$s_F(a \cup b) = s_F(a) + s_{F[a]}(b). \tag{1}$$

Lemma 1. *The function s_F has the following properties:*
(i) For any $e \in X$, $s_F(e) = d(b_{ind(e)})$;
(ii) For any two elements $a, b \in X$,
$s_{F[b]}(a) = s_F(a)$ if $ind(a) \neq ind(b)$,
$s_{F[b]}(a) \leq s_F(a)$ if $ind(a) = ind(b)$.

Proof. The first two equalities are obvious. Assume that $ind(a) = ind(b) = k$. The closure of B_k contains a and coincides with the closure of $b \cup B_{k-1}$. Therefore $s_{F[b]}(a) \leq d(b_{k-1}) \leq d(b_k) = s_F(a)$. []

For a set of coalitions $Z \subset P$, we define

$$win_F(Z) = s_F(Z) - d(Z) = w(F) - w(F[Z]) - d(Z).$$

Lemma 2. *For any $A, B \subset X$ and $Z \subset P$,*

$$s_{F[B]}(A) \leq s_F(A), win_F(Z) \leq win_{F[B]}(Z).$$

Proof. It follows directly from Lemma 1 (ii).[]

Lemma 3. *Let Z be a nonempty set of coalitions. Then for every element e, there exists $z \in Z$ such that*

$$win_{F[e]}(Z - z) \geq win_{F[z]}(Z - z). \tag{2}$$

Proof. We claim that Lemma 2 holds if $e \in Z^*$. Indeed, there exists a coalition $z \in Z$ such that $e \in z$. Then the inequality (2) follows from Lemma 2. So we may assume that $e \notin Z^*$.

We will prove the existence of $z \in Z$ such that the value

$$win_{F[e]}(Z - z) - win_{F[z]}(Z - z) = s_{F[e]}(Z^* - z) - s_{F[z]}(Z^* - z) =$$
$$w(F[e]) - w(F[e \cup Z^* - z]) - w(F[z]) + w(F[Z^* - z]) =$$
$$s_F(e \cup Z^* - z) - s_F(e) - s_F(Z^* - z) + s_F(z) \tag{3}$$

is nonnegative.

It follows from Lemma 2 that value (3) equals

$$s_{F_{k-1}}(e \cup Z^* - z) - s_{F_{k-1}}(e) - s_{F_{k-1}}(Z^* - z) + s_{F_{k-1}}(z), \tag{4}$$

where $k = min\{ind(a) | a \in Z^* \cup e\}$.

Let $Y = \{y \in Z^* \cup e | ind(y) = k\}$. Lemma 1 and the equality (1) imply that for any nonempty subsets $Y' \subset Y$ and $A \subset X$,

$$s_{F_k}(Y') = 0, s_{F_k}(A \cup Y') = s_{F_k}(A),$$
$$s_{F_{k-1}}(Y') = s_{F_{k-1}}(y) + s_{F_{k-1}[y]}(Y' - y) = d(b_k).$$

Now we can rewrite (4) as follows:

$$[s_{F_{k-1}}(Y - z) - s_{F_{k-1}}(e \cap Y) - s_{F_{k-1}}(Y - (z \cup e)) + s_{F_{k-1}}(z \cap Y)]$$
$$+ s_{F_k}(e \cup Z^* - z) - s_{F_k}(e) - s_{F_k}(Z^* - z) + s_{F_k}(z) \tag{5}$$

We claim that (5) is nonnegative if $e \in Y$. Indeed, in this case (5) equals

$$s_{F_{k-1}}(z \cap Y) - s_{F_{k-1}}(Y - (z \cup e)) + s_{F_k}(z).$$

If $Y \cap Z = \emptyset$, then $s_{F_{k-1}}(Y - (z \cup e)) = 0$. Otherwise, we can choose z such that $z \cap Y \neq \emptyset$ and $s_{F_{k-1}}(z \cap Y) = d(b_k) \geq s_{F_{k-1}}(Y - (z \cup e))$.

We now turn to the case $e \notin Y$. Then the expression in square brackets (see (5)) equals $s_{F_{k-1}}(z \cap Y) \geq 0$ and therefore (3) is no less than

$$s_{F_k}(e \cup Z^* - z) - s_{F_k}(e) - s_{F_k}(Z^* - z) + s_{F_k}(z).$$

Similarly, we may contract $b_{k+1}, b_{k+2}, ..., b_{ind(e)}$ without increasing the value of (3). Let $F' = F_{ind(e)}$. As mentioned above,

$$s_{F'}(e \cup Z^* - z) - s_{F'}(e) - s_{F'}(Z^* - z) + s_{F'}(z) \geq 0.\Box$$

Let $z_1, ..., z_m$ be a sequence of coalitions. We say that this sequence is *greedy in* F if it satisfies the following conditions:
(1) if $win_F(z) \leq 0$ for every $z \in P$, then $m = 0$,
otherwise, $win(z_1) \geq win(z)$ for every coalition z;
(2) the sequence $z_2, ..., z_m$ is greedy in $F[z_1]$.

Lemma 4. *If H is the set of elements of a greedy sequence of coalitions and n is the maximum cardinality of elements of H, then for every set of coalitions Z,*

$$nwin_F(H) \geq win_F(Z). \tag{6}$$

Proof. We shall prove (6) by induction on $\#H$. If $H = \emptyset$, then $win_F(Z) \leq 0$. Indeed, in this case for every coalition z, $win_F(z) \leq 0$ and $win_F(Z - z) = win_F(Z) - win_F(z)$.

In the inductive step, let h be the first element of a greedy sequence. The cardinality of h, say m, is at most n. By Lemma 2, there exists a subset Y of Z with at most m elements, such that

$$win_{F[h]}(Z) \geq win_{F[Y]}(Z - Y) = win_F(Z) - win_{F[Z-Y]}(Y) \geq$$

$$win_F(Z) - win_F(Y) \geq win_F(Z) - nwin_F(h).$$

The last inequality follows from trivially if Y is empty or a singleton set. If $Y = \{y_1, y_2, ..., y_m\}$, then, by Lemma 2,

$$win_F(Y) = win_F(y_1) + win_{F[y_1]}(y_2) + ... + win_{F[Y-y_m]}(y_m) \leq$$

$$win_F(y_1) + win_F(y-2) + ... + win_F(y_m) = mwin_F(h) \leq nwin_F(h).$$

Note that $H - h$ is the set of elements of a greedy sequence in $F[h]$. By inductive hypothesis and inequality $win_{F[h]}(Z) \geq win_F(Z) - nwin_F(h)$, we conclude

$$nwin_F(H) = nwin_{F[h]}(H - h) + nwin_F(h) \geq$$

$$win_{F[h]}(Z) + nwin_F(h) \geq win_F(Z). []_{Lemma4}$$

Theorem follows from Lemma 3 and the next two equalities:

$$max\{win(Z)|Z \subset P\} = w(M) - w(M),$$

$$win_M(W) = w(M) - d(B^*).$$

Since W is the greedy sequence of coalitions (see Algorithm), the last equality holds.[]

If Algorithm was applied to a k-polymatroid, then the weight of the resulted base is denoted by gam_k and we may put $gam_1 = w(M)$. Denote the weight of the minimum base of K-polymatroid by opt_k.

Corollary 1. $kgam_k \leq opt_k + (k-1)gam_1$

4 Steiner Tree Problem

At first, we formulate the Steiner Tree Problem (SP). Given a distance graph $G = (V, E, d)$ and a set S of distinguished vertices, SP requires the shortest tree within G which spans S. We denote this minimum Steiner tree by $SMT(S)$.

To show the connection between WPP and SP let consider a complete graph $G' = (S, E', d')$ in which the length of every edge equals to the weight of the shortest path between its ends in G. This graph is the underlying matroid. The subsets of vertices correspond to the closed sets of the family P. The *Steiner length* of a set of vertices $A \in S$ is defined to be the length of the shortest tree within G which spans A. Thus, SP requires the minimum length base of this weighted polymatroid.

Note, that this formulation gives nothing, since we cannot find the Steiner length of large sets. So we restrict our attention to small Steiner trees and introduce some additional notations.

$SMT(S)$ may in general contain vertices of $V \backslash S$. So $SMT(S)$ contains the set S of *given vertices* and and some additional vertices. The Steiner tree is called a *full Steiner tree* if all given vertices are the leaves of $SMT(S)$. If $SMT(S)$ is not full, then we can split it into the union of edge-disjoint full Steiner components. $SMT(S)$ is called $k-$*restricted* if the size of full Steiner components is at most k. We may consider the problem of finding the shortest $k-$restricted Steiner tree for the set S, i.e., component-size bounded Steiner Problem (SPk). For example, the 2-restricted Steiner tree coinsides with the minimum length spanning tree of the graph G'.

Thus, we have an important example of a weighted $k-$polymatroid (corresponding to the graph G') with the Steiner length as a weight function. Note, that the rank of k-subset of S equals to $k-1$. Theorem 1 has the obvious following

Corollary 2 [1]. *Let sp_k be the length of the exact solution of SPk. Then*

$$kgam_k \leq sp_{k+1} + (k-1)gam_1$$

Let sp denotes the length of SMT(S). Note, that $gam_1 = sp_2$, since GAM finds the exact minimum base of a weighted matroid. It was proved that $sp_2/sp \leq 2$ [4] and $sp_3/sp \leq 5/3$ [6].

Corollary 3 [6]. *The approximation bound for Algorithm applied to SP3 equals to*

$$gam_2/sp \leq 11/6$$

If the graph G arises from the rectilinear metrics (Rectilinear Steiner Problem), then denote the length of SMT(S) by rsp. The correspondig constants are [2,7] :

$$gam_1/rsp \leq 3/2$$

$$rsp_3/rsp \leq 5/4$$

$$gam_2/rsp \leq 11/8$$

References

[1] Ding-Zhu Du, Yanjun Zhang and Qing Feng. On better Heuristic for euclidian Steiner minimum trees. *Proceedings, 32nd FOCS*, 431-439, 1991.

[2] F. K.Hwang. On Steiner minimal trees with rectilinear distance. *SIAM J.Appl.Math.*, 30:104-114, 1976.

[3] R. M.Karp. Reducibility among combinatorial problems. In Miller and Tatcher (eds.), *Complexity of Computer Computations*, Plenum Press, New York 85-103, 1972.

[4] H. Takahashi and A. Matsuyama. An approximate solution for the Steiner problem in graphs. *Math. Japonica*, 24: 573-577, 1980.

[5] L. Lovasz and D. Plummer. Matching theory. Bp. Acad. Kiado 1986, 544 p.

[6] A. Z.Zelikovsky. An 11/6-approximation algorithm for the Steiner problem on graphs. In Complexity and computation, *Proceedings of International Symposium on combinatorics, Prachatice (Chechslovakia), June 10-16, 1990* (to appear)

[7] A. Z.Zelikovsky. An 11/8-approximation algorithm for the Steiner problem in networks with rectilinear distance. *Proceedings of J.Bolyai math. soc. conference "Sets, Graphs and Numbers", Budapest (Hungary), January 20-26, 1991* (to appear).