MINIMUM COST MULTIFLOWS IN UNDIRECTED NETWORKS †

A.V.Karzanov

Institute for Systems Analysis
9, Prospect 60 Let Oktyabrya, 117312 Moscow, Russia

Abstract. Let $N = (G, T, c, a)$ be a network, where G is an undirected graph, T is a distinguished subset of its vertices (called *terminals*), and each edge e of G has nonnegative integer-valued *capacity* $c(e)$ and *cost* $a(e)$. The *minimum cost maximum multi*(commodity)*flow* problem (∗) studied in this paper is to find a c-admissible multiflow f in G such that: (i) f is allowed to contain partial flows connecting any pairs of terminals, (ii) the total value of f is as large as possible, and (iii) the total cost of f is as small as possible, subject to (ii). This generalizes, on one hand, the undirected version of the classical minimum cost maximum flow problem (when $|T| = 2$), and, on the other hand, the problem of finding a maximum fractional packing of T-paths (when $a \equiv 0$). Lovász and Cherkassky independently proved that the latter has a half-integral optimal solution.

In [1] a pseudo-polynomial algorithm for solving (∗) was developed and, as its consequence, the theorem on the existence of a half-integral optimal solution for (∗) was obtained. In the present paper we give a direct, shorter, proof of this theorem. Then we prove the existence of a half-integral optimal solution for the dual problem. Finally, we show that half-integral optimal primal and dual solutions can be designed by a combinatorial strongly polynomial algorithm, provided that some optimal dual solution is known (the latter can be found, in strongly polynomial time, by use of a version of the ellipsoid method).

Key words: Network, Multicommodity Flow, Minimum Cost Flow, Edge-disjoint Paths

1. Introduction

Throughout, by a *graph* (*digraph*) we mean a finite undirected (directed) graph without loops and multiple edges. VG is the *vertex-set* and EG is the *edge-set* (*arc-set*) of a graph (digraph) G. An edge of a graph with end vertices x and y is denoted by xy. A *path*, or an $x_0 - x_k$ *path*, in a graph (digraph) G is a sequence $P = (x_0, e_1, x_1, ..., e_k, x_k)$ with $x_i \in VG$ and $e_i = x_{i-1}x_i \in EG$ (respectively, $e_i = (x_{i-1}, x_i) \in EG$).

We deal with an undirected network $N = (G, T, c, a)$, where G is a graph; T is a subset of its vertices, called *terminals* in N; and each edge $e \in EG$ is provided with a *capacity* $c(e) \in \mathbf{Z}_+$ and a *cost* $a(e) \in \mathbf{Z}_+$.

† This work was supported in part by a grant from Mairie de Grenoble, France.

Let $\mathcal{P} := \mathcal{P}(G,T)$ denote the set of simple $s-t$ paths in G for distinct $s,t \in T$. A (c-admissible) *multicommodity flow*, or, briefly, a *multiflow*, in N is a mapping $f : \mathcal{P} \to \mathbf{Q}_+$ satisfying the capacity constraint

$$(1) \qquad \zeta^f(e) := \sum(f(P) \,:\, e \in P \in \mathcal{P}) \leq c(e) \quad \text{for all } e \in EG$$

(here, writing $e \in P$, we consider a path as an edge-set). Sometimes it will be convenient to think of f as consisting of *flows* f_{st} ($s,t \in T$, $s \neq t$), where f_{st} is the restriction of f to the set of $s-t$ paths. The *total value* v_f of f is $\sum(f(P) \,:\, P \in \mathcal{P})$, and the *total cost* a_f is $\sum(f(P)a(P) \,:\, P \in \mathcal{P})$, or $\sum(a(e)\zeta^f(e) \,:\, e \in EG)$. [For $g : S \to \mathbf{Q}$ and $S' \subseteq S$, $g(S')$ denotes $\sum(g(e) \,:\, e \in S')$.] We say that f is a *maximum* multiflow if its total value v_f is as large as possible.

The following problem will be the focus of the present paper:

(2) *given $N = (G,T,c,a)$, find a maximum multiflow f in N whose total cost a_f is minimum.*

This problem has two well-known special cases.

(i) When T consists of two terminals, s and t say, (2) turns into the (undirected) minimum-cost maximum-*flow* problem: find a maximum flow from s to t whose total cost is minimum. A classical result in network flow theory is that this problem has an integral optimal solution [2]. Moreover, such a solution can be found in strongly polynomial time [3] (see also [4] for a purely combinatorial strongly polynomial algorithm, and [5] for a survey).

(ii) When $a = 0$, we get the maximum multiflow problem, which is, in fact, the fractional relaxation of the problem on finding a maximum packing of T-paths (paths connecting arbitrary pairs in T). Lovász [6] and Cherkassky [7] independently proved the existence of an optimal solution f that is *half-integral* (that is, $2f$ is integer-valued). Moreover, in [7] a strongly polynomial algorithm was designed to find such a solution (a faster, of complexity $O(\eta_n \log |T|)$, algorithm was developed in [8], where η_n is the time required to find a maximum flow in a network with n vertices).

Fig. 1

Figure 1 illustrates an instance of problem (2) with $|T| > 2$ for which no integral optimal solution exists (here $T := \{s_1, \ldots, s_6\}$, and $c(e) = a(e) = 1$ for all edges e.) Nevertheless, the following is true.

Theorem 1 [1]. *Problem (2) has a half-integral optimal solution.*

Instead of (2), it is convenient to consider a more general problem, namely:

(3) *Given $p \in \mathbf{Q}_+$, find a multiflow f in N which maximizes the objective function $pv_f - a_f$.*

Theorem 2 [1]. *For any $p \geq 0$ problem (3) has a half-integral optimal solution.*

By standard linear programming arguments, (2) and (3) are equivalent whenever p is large enough (moreover, the existence of a half-integral optimal solution for (3) easily implies that taking p to be $2a(EG)c(EG) + 1$ is sufficient). Thus, Theorem 1 immediately follows from Theorem 2.

In its turn, Theorem 2 was obtained in [1] as a consequence of an algorithm developed there, which constructs a sequence $f_1, f_2, \ldots, f_M$ of half-integral multiflows in N together with a sequence $0 < p_1 < p_2 < \ldots < p_M$ of rationals so that for $i = 1, \ldots, M$, f_i is an optimal solution of (3) for any p such that $p_i \leq p < p_{i+1}$, assuming p_{M+1} to be ∞. This algorithm is pseudo-polynomial, the number of elementary operations in it (over numbers of $O(Q \log(\widehat{c}+\widehat{a}))$ digits in binary notation) is bounded by the minimum of $\widehat{c}P_1$, $\widehat{a}P_2$ and 2^{P_3}, where $\widehat{c} := c(EG) + 1$, $\widehat{a} := a(EG) + 1$, and Q, P_1, P_2, P_3 are polynomials in $|VG|$.

The goals of the present paper are:

(i) to give a direct, shorter, proof of Theorem 2 (Section 2);

(ii) to show that the problem dual to (3) has a half-integral optimal solution, provided that p is an integer, and to design a combinatorial strongly polynomial algorithm for finding half-integral optimal primal and dual solutions for (3), provided that some optimal dual solution is given (Section 3).

[Note that an optimal dual solution for (3) can be found in strongly polymonial time by use a general approach due to Tardos [9] based on the ellipsoid method [10]; see Section 4 for more explanations.] Assign to an edge $e \in EG$ a variable $l(e) \in \mathbf{Q}$. Then the linear program dual to (3) is:

(4) minimize $cl := \sum(c(e)l(e) \,:\, e \in EG)$, provided that
   (i) $l \geq 0$, and
   (ii) $l(P) \geq p - a(P)$ for any $P \in \mathcal{P}$.

In conclusion of the Introduction let us consider a more general concept of the minimum cost maximum multiflow problem. More precisely, let $H = (T, U)$ be a graph, called the *commodity graph*, whose edges are to indicate the pairs of terminals which are allowed to connect by flows; in particular, the above definition of a multiflow was concerned with H to be the *complete* graph on T. Now we define a multiflow as a corresponding function on the set of simple $s-t$ paths in G such that $\{s,t\}$ is an edge of H. According to this, we speak of a maximum multiflow and pose problem (2) with respect to H. E.g., if $|U|=2$, we obtain the minimum cost maximum *two-commodity-flow* problem. A natural question arises: given H, what is the minimum integer $k:=k(H)$ such that, for any G with $VG \supseteq T$, c (integral) and a, problem (2) for G, H, c, a has an optimal solution f with kf integer-valued?

In particular, $k(H)=1$ if $|U|=1$, and $k(H)=2$ if H is the complete graph K_m with $m \ge 3$ vertices (by Theorem 1). Theorem 1 can be easily generalized as follows (cf. [11]): if H is a complete m-partite graph with $m \ge 3$ then $k(H)=2$ (while $k(H)=1$ if $m=2$). [H is called m-partite if there is a partition $\{T_1,\ldots,T_m\}$ of T such that $st \in U$ if and only if $s \in T_i$ and $t \in T_j$ for $i \ne j$.] On the other hand, it was shown in [11] that $k(H)=\infty$ unless H is a complete m-partite graph (e.g., $k(H)=\infty$ if U consists of two non-adjacent edges).

2. Proof of Theorem 2.

For $\lambda \in \mathbf{Q}_+^{EG}$ and $x,y \in VG$, let $\text{dist}_\lambda(x,y)$ denote the λ-*distance* between vertices x and y, that is, the minimum λ-*length* $\lambda(P)$ of an $x-y$ path P in G. Obviously, the system (ii) in (4) can be rewritten in a more compact form, namely,

(5) $$\text{dist}_{a+l}(s,t) \ge p \quad \text{for any } s,t \in T, s \ne t.$$

The linear programming duality theorem applied to (3)-(4) implies that a (c-admissible) multiflow f and a vector $l \in \mathbf{Q}_+^{EG}$ satisfying (5) are optimal solutions of (3) and (4), respectively, if the following (complementary slackness) conditions hold:

(6) if $P \in \mathcal{P}$ and $f(P)>0$ then $a(P)+l(P)=p$; in particular, P is an $(a+l)$-*shortest* path in G (that is, a shortest path with respect to the length $a+l$);

(7) if $e \in EG$ and $l(e)>0$ then e is *saturated* by f, that is, $\zeta^f(e)=c(e)$.

We at first prove Theorem 2 for the case when the cost function a is *positive*, that is, $a(e)>0$ for all $e \in EG$. The proof will follow from a series of auxiliary statements (Claims 1-5), some of them, as well as the idea to design the "doubly covering" digraph Γ defined below, occurred in [1]. We need some terminology and notation.

For brevity, a path $P=(x_0,e_1,x_1,\ldots,e_k,x_k)$ in G may be denoted by $x_0x_1\ldots x_k$ (it is not confusing because G has no multiple edges). For $0 \le i \le j \le k$, $P(x_i,x_j)$ is the part $x_ix_{i+1}\ldots x_j$ of P from x_i to x_j. The reverse path $x_kx_{k-1}\ldots x_0$ is denoted by P^{-1}.

Let us be given a positive function λ on EG. Put

(8) $$p' := p_\lambda := \min\{\mathrm{dist}_\lambda(s,t)\ :\ s,t\in T, s\neq t\}.$$

A path P connecting two distinct terminals and having λ-length exactly p' is called a *geodesic* for λ, or a λ-*geodesic*; in particular, P is λ-shortest. Let G_λ be the subgraph of G whose edges belong to λ-geodesics and vertices belong to λ-geodesics or T.

Consider a vertex $v\in VG_\lambda$. Define the *potential* $\pi(v) := \pi_\lambda(v)$ of v to be the λ-distance from v to T. In particular, $\pi(v)=0$ if $v\in T$.

Claim 1. *Let v belong to a geodesic P from s to t. Then $\pi(v)$ is the minimum of two lengths $q:=\lambda(P(s,v))$ and $r:=\lambda(P(v,t))$.*

Proof. Assume for definiteness that $q\le r$. Then $\pi(v)\le q\le p'/2$ (since $q+r=\lambda(P)=p'$). Suppose that $\pi(v)<q$, and let s' be a terminal such that $\pi(v)=\mathrm{dist}_\lambda(s',v)$. Choose $s''\in\{s,t\}$ such that $s''\neq s'$. Now, $\mathrm{dist}_\lambda(s',s'')\le \mathrm{dist}_\lambda(s',v)+\mathrm{dist}_\lambda(v,s'')\le \pi(v)+r<q+r=p'$; a contradiction. •

From Claim 1 it follows immediately that $\pi(v)\le p'/2$. For $s\in T$ define $V^s := V^s_\lambda$ to be $\{v\in VG_\lambda\ :\ \mathrm{dist}_\lambda(s,v)<p'/2\}$; and define $C:=C_\lambda:=\{v\in VG_\lambda\ :\ \pi(v)=p'/2\}$; a vertex in C is called *central.* Then the sets V^s ($s\in T$) and C are pairwise disjoint and give a partition of VG_λ. Also Claim 1 shows that if $P=v_0v_1\ldots v_k$ is a geodesic from $s=v_0$ to $t=v_k$, then there are i and j such that $v_0,\ldots,v_i\in V^s$, $v_j,\ldots,v_k\in V^t$, and either $j=i+1$, or $j=i+2$ and $v_{i+1}\in C$. Let E^s (respectively, $E^{\{s,t\}}$, where $t\in T-\{s\}$) denote the set of edges in G_λ with one end in V^s and the other in $V^s\cup C$ (respectively, in V^t).

Claim 2. *Let $e=uv\in EG_\lambda$. Then either $e\in E^s$ for some $s\in T$, and $|\pi(v)-\pi(u)|=\lambda(e)$; or $e\in E^{\{s,t\}}$ for some $s,t\in T$, and $\pi(u)+\lambda(e)+\pi(v)=p'$. In particular, the sets E^s ($s\in T$) and $E^{\{s,t\}}$ ($s,t\in T$) give a partition of EG_λ, and no edge of G_λ connects two central vertices.*

Proof. By definition of G_λ, there is a geodesic P from s to t passing u,e,v (in this order). Then $\lambda(P(s,u))+\lambda(e)+\lambda(P(v,t))=p'$. One may assume that $q:=\lambda(P(s,u))\le\lambda(P(v,t))=:r$. We know that $\lambda(e)>0$ (as λ is positive), whence $q<p'/2$. This implies $q=\pi(u)$ (by Claim 1) and $u\in V^s$. Suppose that $r\ge p'/2$. Then $q':=\lambda(P(s,v))\le p'/2$, whence $q'=\pi(v)$, by Claim 1. This implies that $v\in V^s\cup C$, $e\in E^s$ and $\pi(v)-\pi(u)=\lambda(e)$. Now suppose that $r<p'/2$. Then $\pi(v)=r$, and we obtain $v\in V^t$, $e\in E^{\{s,t\}}$ and $\pi(u)+\lambda(e)+\pi(v)=p'$. •

The following claim describes geodesics in terms of potentials.

Claim 3. *Let $P=v_0v_1\ldots v_k$ be an $s-t$ path in G_λ with $s,t\in T$ and $s\neq t$. The following are equivalent:*

(i) P is a geodesic;

(ii) there is q, $0 \le q < k$, such that $\pi(v_i) - \pi(v_{i-1}) = \lambda(v_{i-1}v_i)$ for $i = 1, \dots, q$ and $\pi(v_i) - \pi(v_{i+1}) = \lambda(v_i v_{i+1})$ for $i = q+1, \dots, k-1$.

Proof. (i)→(ii) follows from Claim 2. To show (ii)→(i), put $u := v_q$, $v := v_{q+1}$ and $e := uv$. Observe that $\pi(u) = \lambda(P(s,u))$ and $\pi(v) = \lambda(P(v,t))$. Consider an $s' - t'$ geodesic Q passing u, e, v (in this order). Then $\lambda(Q) = \lambda(Q(s',u)) + \lambda(e) + \lambda(Q(v,t'))$, $\lambda(Q(s',u)) \ge \pi(u)$ and $\lambda(Q(v,t')) \ge \pi(v)$, whence $p' = \lambda(Q) \ge \lambda(P)$, and therefore, P is a geodesic. •

Now, based on Claims 2 and 3, we design the so-called *doubly covering digraph* $\Gamma = \Gamma_\lambda$ for G_λ. Each non-central vertex v of G_λ generates two vertices v^1 and v^2 in Γ. If $v \in VG_\lambda$ is central, it generates $2|T(v)|$ vertices v_s^i ($s \in T(v)$, $i = 1,2$) in Γ, where $T(v) := \{s \in T : \mathrm{dist}_\lambda(s,v) = p'/2\}$. The arcs of Γ are defined as follows:

(9) (i) an edge $uv \in E^s$ ($s \in T$) with $\pi(v) - \pi(u) = \lambda(uv)$ induces two arcs (u^1, v^1) and (v^2, u^2) (or (u^1, v_s^1) and (v_s^2, u^2) when v is central) in Γ, each of capacity $c(uv)$;

(ii) an edge $uv \in E^{\{s,t\}}$ ($s, t \in T$) induces two arcs (u^1, v^2) and (v^1, u^2) in Γ, each of capacity $c(uv)$;

(iii) a central vertex $v \in C$ induces $|T(v)|(|T(v)| - 1)$ arcs (v_s^1, v_t^2) in Γ for all distinct $s, t \in T(v)$, each of capacity ∞;

[See Fig. 2 for illustration; here $T = \{s, t, q\}$ and the number on the edges e indicate $\lambda(e)$.] The positivity of λ makes Γ well-defined; furthermore, one can see that Γ is acyclic.

Fig. 2

It is convenient to keep the same notation c for the capacities of arcs in Γ. The arcs of a subgraph in Γ arising from a central vertex $v \in C$ are called *central*. We think of $S := \{s^1 : s \in T\}$ (respectively, $S' := \{s^2 : s \in T\}$) as the set of *sources* (respectively, *sinks*) of Γ. Let $P = x_0 x_2 \dots x_k$ be a path in Γ (that is, $(x_{i-1}, x_i) \in E\Gamma$

for $i = 1, \dots, k$). We say that P is an $S - S'$ *path* if $x_0 \in S$ and $x_k \in S'$.

The construction of Γ determines a natural mapping τ of $V\Gamma \cup E\Gamma$ onto $VG_\lambda \cup EG_\lambda$ which brings $v^i \in V\Gamma$ (or $v_s^i \in V\Gamma$) to the vertex v, brings a non-central arc $(x, y) \in E\Gamma$ to the edge $\tau(x)\tau(y)$, and brings a central arc (v_s^1, v_t^2) to the vertex v.

The mapping τ is naturally extended to paths in Γ and G_λ. Namely, for a path $P = x_0 x_1 \dots x_k$ in Γ, let $\tau(P)$ be the path in G_λ induced by the sequence $\tau(x_0), \tau(x_1), \dots, \tau(x_k)$ of vertices (in which repeated vertices going in succession are deleted).

We also define the mapping $\vartheta : (V\Gamma \cup E\Gamma) \to (V\Gamma \cup E\Gamma)$ such that a vertex v_i (or v_s^i) is mapped to v^{3-i} (or v_s^{3-i}), and an arc (x, y) to $(\vartheta(y), \vartheta(x))$. This gives a (skew) symmetry of Γ. For an $s - t$ path $P = x_0 x_1 \dots x_k$ in Γ, $\vartheta(P)$ is the symmetric $\vartheta(t) - \vartheta(s)$ path $\vartheta(x_k)\vartheta(x_{k-1}) \dots \vartheta(x_0)$; obviously, the path $\tau(P)$ in G is opposite to $\tau(\vartheta(P))$.

Claim 4. *τ determines a one-to-one correspondence between the set of $S - S'$ paths in Γ and the set of λ-geodesics in G.*

Proof. Claim 3 and definition (9) show that for a geodesic P one can form (uniquely) the $S - S'$ path P' in Γ such that $P = \tau(P')$. Conversely, consider an $S - S'$ path $P = x_0 x_1 \dots x_k$ in Γ. From (9) one can see that P contains exactly one arc, $e = (x_q, x_{q+1})$ say, such that either e is central, (v_s^1, v_t^2) say, or $\tau(e) \in E^{\{s,t\}}$ for some distinct $s, t \in T$; let for definiteness $\tau(x_q) \in V^s$. Moreover, $\tau(x_j) \in V^s$ for $j = 0, \dots, q-1$ and $\tau(x_j) \in V^t$ for $j = q+2, \dots, k$, and either $\tau(x_q) = \tau(x_{q+1}) \in C$, or $\tau(x_q) \in V^s$ and $\tau(x_{q+1}) \in V^t$. Now (9), Claim 3, and the fact that $s \neq t$ imply that $\tau(P)$ is a geodesic. •

The above correspondence of geodesics in G and $S - S'$ paths in Γ naturally generates a relationship between certain multiflows in N and $S - S'$ flows in Γ (whose arcs e have the capacities $c(e)$ defined in (9)). We say that a multiflow f in N *goes along λ-geodesics* if $f(P) > 0$ implies that P is a λ-geodesic.

For a function $g : E\Gamma \to \mathbf{Q}_+$ and a vertex $x \in V\Gamma$ define

$$\text{div}_g(x) := \sum_{y:(x,y)\in E\Gamma} g(x, y) - \sum_{y:(y,x)\in E\Gamma} g(y, x). \tag{10}$$

We say that g is a (c-admissible) *flow* from S to S', or $S - S'$ *flow*, if it satisfies the conservation condition $\text{div}_g(x) = 0$ for all $x \in V\Gamma - (S \cup S')$ as well as the capacity constraint $g(e) \leq c(e)$ for all $e \in E\Gamma$. The *value* v_g of a flow g is $\sum(\text{div}_g(x) : x \in S)$; g is called maximum if v_g is as large as possible.

A routine fact is that a flow g as above can be represented as the sum of elementary flows along paths (taking into account that Γ is acyclic). More precisely, there are $S - S'$ paths $P_1, P_2, \dots, P_m$ ($m \leq |E\Gamma|$) and positive rationals $\alpha_1, \alpha_2, \dots, \alpha_m$ such that:

$$\sum(\alpha_i : e \in P_i) = g(e) \text{ for any } e \in E\Gamma. \tag{11}$$

From (11) it follows that $v_g = \sum(\alpha_i : i = 1,\dots,m)$. We say that $\mathcal{D} := \{(P_i,\alpha_i) : i = 1,\dots,m)\}$ is a *decomposition* of g. If g is integral then there exists a decomposition with all α_i's integral; such a decomposition can be found by a trivial procedure of complexity $O(|V\Gamma||E\Gamma|)$ (cf. [2]). A decomposition $\mathcal{D}$ determines a multiflow $f := f_{\mathcal{D}}$ in N by setting $f(\tau(P_i)) := \alpha_i/2$ for $i = 1,\dots,m$, and $f(P) := 0$ for the remaining paths P in $\mathcal{P}$. Using (11) we observe that for any non-central $e \in E\Gamma$,

$$\zeta^f(\tau(e)) = \frac{1}{2}(g(e) + g(\vartheta(e))) \le \frac{1}{2}(c(e) + c(\vartheta(e))) = c(\tau(e)),$$

that is, f is c-admissible. Moreover, f goes along geodesics, and $2v_f = v_g$. Also a converse (in a sense) property is true. More precisely, for a (c-admissible) multiflow f in N going along geodesics, define the function $g = g_f$ on $E\Gamma$ so that for $e \in E\Gamma$, $g(e)$ is the sum of values $f(\tau(P))$ over all $S - S'$ paths P in Γ containing e or $\vartheta(e)$. Using the definition of c on $E\Gamma$ (in particular, the fact that $c(e) = \infty$ if e is central), one can check that g_f is a (c-admissible) $S - S'$ flow in Γ, and $v_{g_f} = 2v_f$. These observations are summarized as follows.

Claim 5. (i) *If g is an $S - S'$ flow in Γ and $\mathcal{D} = (P_i,\alpha_i)$ is a decomposition of g, then $f = f_{\mathcal{D}}$ is a multiflow in N going along geodesics, and $v_g = 2v_f$. Moreover, if all α_i's are integral, then f is half-integral.*

(ii) *If f is a multiflow in N going along geodesics, then $g = g_f$ is an $S - S'$ flow in Γ, and $v_g = 2v_f$.* •

Since c is integral, there exists an integral maximum $S - S'$ flow in Γ. This gives the following corollary of Claim 5, which is not used later on, but interesting in its own right.

Claim 6. *If c is integral then there exists a half-integral maximum multiflow in N going along geodesics.* •

Now we are able to prove Theorem 2 (under the assumption of positivity of a). Given $p \ge 0$, suppose that f and l are optimal solutions of (3) and (4), respectively. Let $\lambda := a + l$; then λ is positive. We may assume that $p = p_\lambda$ (since $p \le p_\lambda$, by (5), and if $p < p_\lambda$ then $f = 0$, by (6)). By (6), f is a multiflow in N going along λ-geodesics; in particular, $\zeta^f(e) > 0$ holds only if e is in G_λ. Consider the $S - S'$ flow $g := g_f$ in $\Gamma := \Gamma_\lambda$. We say that an arc $e \in E\Gamma$ is *feasible* if either e is central or $l(\tau(e)) = 0$; let A denote the set of feasible arcs. By (7), for each $e \in E\Gamma - A$ the edge $\tau(e)$ in G is saturated by f (since $l(\tau(e)) > 0$), which implies that $g(e) = c(e)$, and hence g is integral on $E\Gamma - A$. Let γ be the restriction of g to A. We say that a function $h : A \to \mathbf{Q}_+$ is *compatible* with γ if h is c-admissible, that is, $h(e) \le c(e)$ for $e \in A$, and

$$\mathrm{div}_{A,h}(x) = \mathrm{div}_{A,\gamma}(x) \quad \text{for all } x \in V\Gamma - (S \cup S'),$$

where for a function b, $\mathrm{div}_{A,b}$ is defined as in (10) with respect to A rather than $E\Gamma$.

The value $\mathrm{div}_{A,\gamma}(x) = -\mathrm{div}_{E\Gamma - A,g}(x)$ is an integer for each $x \in V\Gamma - (S \cup S')$. Hence, there is an *integral* h which is compatible with γ. Define g' by $g'(e) := h(e)$ for $e \in A$ and $g'(e) := g(e)$ for $e \in E\Gamma - A$. Then g' is an integral $S - S'$ flow in Γ. Now the multiflow $f' := f_{\mathcal{D}}$ gives a half-integral multiflow in N, where $\mathcal{D} = \{(P_i, \alpha_i)\}$ is a decomposition of g with all α_i's integral. Since g' can differ from g only on arcs in A, f' satisfies (7). Furthermore, (i) in Claim 5 implies (6) for f'. Hence, f' is an optimal solution of (3), and the theorem follows.

Now suppose that a is a *nonnegative* cost function on EG. Put $Z := \{e \in EG : a(e) = 0\}$. To prove the theorem for a, we apply obvious perturbation techniques, replacing a by appropriate positive cost functions. More precisely, for a rational number $\delta > 0$, put $a^{\delta}(e) := a$ for $e \in EG - Z$ and $a^{\delta}(e) := \delta$ for $e \in Z$, and consider an infinite sequence $\delta_1 > \delta_2 > \ldots$ of positive rationals approaching zero. By the above proof, problem (3) for G, T, c, a^{δ_i} has a half-integral optimal solution f_i. Moreover, the number of different half-integral c-addmissible multiflows for G, T is finite (as $\mathcal{P}(G,T)$ is finite and c is bounded), hence, we may assume that all the f_i's are the same, f say. Now trivial arguments yield that f must be an optimal solution of (3) for G, T, c, a.

This completes the proof of Theorem 2. ••

Remark 2.1. Assuming, without loss of generality, that p is an integer (and that a is integer-valued, as before), we observe that taking δ as above to be $(4c(Z)+1)^{-1}$ ensures that any half-integral optimal solution f for (3) with a^{δ} is an optimal solution for (3) with a. Indeed, suppose for a contradiction that there is a (c-admissible) multiflow f' such that $d := \phi(f', a) - \phi(f, a) > 0$, where $\phi(f'', a'')$ denotes $pv_{f''} - a''_{f''}$. The integrality of a and p together with the half-integrality of f and f' implies that $d \geq \frac{1}{2}$. On the other hand,

$$|\phi(f,a) - \phi(f,a^{\delta})| = \delta \sum_{e \in Z} \zeta^f(e) \leq \delta c(Z) < \frac{1}{4}.$$

(by the choice of δ), and similarly, $|\phi(f', a) - \phi(f', a^{\delta})| < \frac{1}{4}$. Then $\phi(f', a^{\delta}) - \phi(f, a^{\delta}) > d - \frac{1}{4} - \frac{1}{4} \geq 0$, contrary to the optimality of f for a^{δ}.

3. Dual half-integrality and algorithm

Theorem 3. *Let p be a nonnegative integer. Then (4) has a half-integral optimal solution.*

This theorem follows from Theorem 2 and the general fact that for a system of inequalities the "totally dual $1/k$-integrality" implies the "totally primal $1/k$-integrality" (this is a natural generalization of the well-known result on TDI systems due to Edmonds and Jiles [12]). More precisely, to our purposes, it suffices to utilize the following simple fact (see, e.g., [13, Statement 1.1]).

Statement 3.1. *Let A be a nonnegative $m \times n$-matrix, let b be an integral m-vector, and let k be a positive integer. Suppose that the program $D(c) := \max\{yb : y \in \mathbf{Q}_+^m,\ yA \le c\}$ has a $1/k$-integral optimal solution for every nonnegative integral n-vector c such that $D(c)$ has an optimal solution. Then for every nonnegative integral n-vector c, the program $P(c) := \min\{cx : x \in \mathbf{Q}_+^n,\ Ax \ge b\}$ has a $1/k$-integral optimal solution whenever it has an optimal solution.* •

In our case, one should put $k := 2$ and take as A (b) the constraint matrix (respectively, the right hand size vector) of the system (ii) in (4). Then b is integral, and $D(c)$ is just problem (3), whence the result follows. •

Thus, the polyhedron $Q = \{l \in \mathbf{Q}^E : l$ satisfies (i),(ii) in (4)$\}$ is half-integral (that is, every face of Q contains a half-integral point), by Theorem 3. In addition, the separation problem for Q is obviously reduced to finding $a + l$-shortest paths in G connecting pairs of terminals. Hence, a half-integral optimal solution of (4) (with integral a and p) can be found in polynomial time by use of the ellipsoid method and arguments as in [14].

Furthermore, since (ii) in (4) is equivalent to (5), to find *some* optimal solution l of (4) is a linear program whose constraint matrix has entries only 0,1,-1 and consists of $O(|T||EG|)$ rows and $O(|T||VG| + |EG|)$ columns. Thus, l can be found in strongly polynomial time by use of a method due to Tardos [9]. However, without a more careful analysis, it is unclear whether any (or even some) optimal basis solution for the latter program is half-integral; so we cannot argue that combining approaches developed in [9] and [14] would enable us to find a half-integral optimal solution of (4) in strongly polynomial time. Nevetherless, this task can be fulfilled by involving certain combinatorial techniques.

More precisely, given c, a, p integral, let $\overline{a}$ be the function a^δ with $\delta := \min\{(4c(Z)+1)^{-1}, (2n^2t+1)^{-1}\}$ (cf. Remark 2.1), where $n := |VG|$ and $t := |T|$. We design a strongly polynomial algorithm consisting of three parts:

(S1) find a (fractional) optimal solution l of (4) for $G, T, c, \overline{a}, p$;

(S2) using l, find a half-integral optimal solution f of (3) for G, T, c, a, p;

(S3) using l and f, find a half-integral optimal solution l' of (4) for G, T, c, a, p.

[Note that (S3) will provide an alternative, combinatorial, proof of Theorem 3.] We use terminology and notation as in Section 2. Part (S1) has been explained above. To solve (S2), we construct $\Gamma := \Gamma_\lambda$ for $\lambda := \overline{a} + l$, and determine an integral $S - S'$ flow g in Γ satisfying $g(u) = c(u)$ for all $u \in E\Gamma - A$ (where A is the set of central arcs and arcs u such that $l(\tau(u)) = 0$). Such a g exists, as it was shown in Section 2. Then $f := f_{\mathcal{D}}$ is a half-integral optimal solution of (3) for $\overline{a}$, where $\mathcal{D}$ is an integral decomposition of g. By Remark 2.1 and the choice of δ in the definition of $\overline{a}$, f is an optimal solution of (3) for a, as required.

To solve (S3), consider the digraph $\Gamma = \Gamma_\lambda$ for $\lambda := \overline{a} + l$ and the $S - S'$ flow $g = g_f$ in Γ; note that $|V\Gamma| \le 2t + 2t(n-t) \le 2nt - 2$. The vertex-set $V\Gamma$ of Γ is partitioned into sets W_s, W'_s ($s \in T$). Here W_s is formed by the vertices v^1 generated by $v \in V^s$ and the vertices v^1_s generated by the central vertices $v \in C$ with $s \in T(v)$; and $W'_s := \vartheta(W_s)$. For $x \in W_s$ ($x \in W'_s$) define $\rho(x)$ to be $\pi(\tau(x))$ (respectively, $p - \pi(\tau(x))$), where π is the potential function on VG_λ defined as in Section 2 for given λ. For $u \in E\Gamma$ put $a(u) := a(\tau(u))$ and $\overline{a}(u) := \overline{a}(\tau(u))$ if u is non-central, and put $a(u) := \overline{a}(u) := 0$ otherwise. Let $U^+ := \{u \in E\Gamma : \zeta^f(u) = c(u)\}$. We know that

$$\text{(12)} \qquad \begin{gathered} \rho(x) + \rho(x') = p \ \text{ for } x \in V\Gamma \text{ and } x' = \vartheta(x); \quad \rho(s) = 0 \ \text{ for } s \in S; \\ \rho(x) \le p \ \text{ for } x \in W_s,\ s \in T; \quad \rho(x) = p/2 \ \text{ if } \tau(x) \in C; \end{gathered}$$

$$\text{(13)} \qquad \begin{aligned} \rho(y) - \rho(x) &\ge \overline{a}(u) \qquad \text{for } u = (x,y) \in U^+, \\ &\overline{a}(u) \qquad \text{for } u = (x,y) \in E\Gamma - U^+ \end{aligned}$$

((13) follows from (7)). Expand Γ by adding new arcs representing, in a sense, the part of G outside G_λ, as follows. Let x, y be distinct vertices in G_λ connected by a path P with all edges in $EG - EG_\lambda$, and let $a\langle x,y\rangle$ ($\overline{a}\langle x,y\rangle$) be the minimum cost $a(P)$ (respectively, $\overline{a}(P)$) among all such paths P.

(14) (i) If $x \in V^s$ and $y \in V^s \cup C$, add to Γ the arcs $(x^1,y^1), (y^1,x^1), (x^2,y^2), (y^2,x^2)$;
(ii) if $x \in V^s$ and $y \in V^t$ ($s \ne t$), add to Γ the arcs (x^1,y^2) and (y^1,x^2);
(iii) for each arc u in (i),(ii), put $a(u) := a\langle x,y\rangle$ and $\overline{a}(u) := \overline{a}\langle x,y\rangle$

(we need not add new arcs to Γ when both x, y are central). The set of these new arcs is denoted by U^0. Obviously, (6) implies:

$$\text{(15)} \qquad \rho(y) - \rho(x) \le a(u) \qquad \text{for any } u = (x,y) \in U^0.$$

Claim. *There is a function ρ' on $V\Gamma$ such that:*

$$\text{(16)} \qquad \rho'(s) = 0 \ \textit{ for } s \in S, \qquad \textit{and } \rho'(s') = p \textit{ for } s' \in T;$$

$$\text{(17)} \qquad \rho'(x) \le \rho'(\vartheta(x)) \ \textit{ for any } x \in W_s,\ s \in T;$$

$$\text{(18)} \qquad \begin{aligned} \rho'(y) - \rho'(x) &\ge a(u) \qquad \textit{if } u = (x,y) \in U^+, \\ &= a(u) \qquad \textit{if } u = (x,y) \in E\Gamma - U^+, \\ &\le a(u) \qquad \textit{if } u = (x,y) \in U^0. \end{aligned}$$

Proof. The existence of ρ' is equivalent to the fact that the digraph H whose edges are weighted by b has no negative circuits. Here H and b are designed as follows:

(i) $VH = V\Gamma \cup \{q, q'\}$; q is connected with each $s \in S$ by arcs (q, s) and (s, q) of weight 0; q' is connected with each $s' \in S'$ by arcs (q', s') and (s', q') of weight 0; while q and q' are connected by an arc (q, q') of weight p and an arc (q', q) of weight $-p$;

(ii) each $x \in W_s$ $(s \in T)$ is connected with $x' := \vartheta(x)$ by an arc (x', x) of weight 0;

(iii) each arc $u = (x, y)$ in U^+ $(U^0;\ E\Gamma - U^+)$ induces in H an arc (y, x) of weight $-a(u)$ (respectively, an arc (x, y) of weight $a(u)$; an arc (x, y) of weight $a(u)$ and an arc (y, x) of weight $-a(u)$).

Suppose that there is a simple circuit Q in H with $b(Q) < 0$. Let $\bar{b}$ be the weighting on EH defined as above for $\bar{a}$ rather than a. The existence of ρ satisfying (12)-(13) implies that H has no negative circuit with respect to $\bar{b}$, hence, $\bar{b}(Q) \ge 0$. Since a and p are integral, $b(Q) \le -1$, whence there is an arc u in Q such that $\bar{b}(u) - b(u) \ge (2nt)^{-1}$ (taking into account that Q is simple, and $|VH| = |V\Gamma| + 2 \le 2nt$). On the other hand, we know that $|\bar{a}(e) - a(e)|$ is at most δ for any $e \in E\Gamma$ and at most $n\delta$ for any $e \in U^0$. Hence, $\bar{b}(u) - b(u) \le n\delta < (2nt)^{-1}$, by the definition of δ; a contradiction. •

Since the system (16)-(18) is solvable, it has an *integral* solution ρ'; it can be found, e.g., by applying a shortest path algorithm to H and b as above. Now for $v \in VG_\lambda$ define $\pi'(x) := \frac{1}{2}(\rho'(x) + p - \rho'(x'))$, where $\tau(x) = \tau(x') = v$, $x \in W_s$ and $x' \in W'_s$ (for the corresponding s). Then π' is half-integral. (17) implies that $\pi'(v) \le p/2$ for each $v \in V^s$, $s \in T$, and (17)-(18) imply that $\pi'(v) = p/2$ for $v \in C$ (taking into account that $a(u) = 0$ for any central arc u in Γ).

Finally, for $e = xy \in EG_\lambda$ define

$$\begin{aligned} l'(e) &:= |\pi'(x) - \pi'(y)| - a(e) \qquad \text{if } x, y \in V^s \cup C, s \in T; \\ &:= p - \pi'(x) - \pi'(y) - a(e) \qquad \text{if } x \in V^s, y \in V^t, s \ne t; \end{aligned}$$

and define $l'(e) := 0$ for $e \in EG - EG_\lambda$. Then l' is half-integral. Using (14)-(18), a routine examination shows that f and l' satisfy (6)-(7) (we leave it to the reader). Thus, l' is as required in (S3).

Note from March 1993. Recently A.V. Goldberg and the author found two purely combinatorial algorithms for finding a half-integral optimal solution to (3) (in: "Transitive fork environments and minimum cost multiflows", Report No. STAN-CS-93-1476, Stanford University, Stanford, 1993). Both algorithms are polynomial (but not, in general, strongly polynomial); the first applies scaling on capacities, and the second applies scaling on costs.

REFERENCES

[1] A.V. Karzanov, A minimum cost maximum multiflow problem, in: *Combinatorial Methods for Flow Problems* (Institute for System Studies, Moscow, 1979, issue 3) pp. 138-156 (Russian).

[2] L.R. Ford and D.R. Fulkerson, *Flows in networks* (Princeton Univ. Press, Princeton, NJ, 1962).

[3] E. Tardos, A strongly polynomial minimum cost circulation algorithm, *Combinatorica* **5** (3) (1985) 247-255.

[4] A.V. Goldberg and R. Tarjan, Finding minimum-cost circulations by canceling negative cycles *J. of the ACM* **36** (4) (1989) 873-886.

[5] A.V. Goldberg, E. Tardos, and R.E. Tarjan, Network flow algorithms, in: B. Korte, L. Lovász, H.J. Prömel, and A. Schrijver, eds., *Paths, Flows, and VLSI-Layout* (Springer, Berlin et al, 1990) pp. 101-164.

[6] L. Lovász, On some connectivity properties of Eulerian graphs, *Acta Math. Akad. Sci. Hung.* **28** (1976) 129-138.

[7] B.V. Cherkassky, A solution of a problem on multicommodity flows in a network, *Ekonomika i Matematicheskie Metody* **13** (1)(1977) 143-151 (Russian).

[8] A.V. Karzanov, Fast algorithms for solving two known undirected multicommodity flow problems, in: *Combinatorial Methods for Flow Problems* (Institute for System Studies, Moscow, 1979, issue 3) pp. 96-103 (Russian).

[9] E. Tardos, A strongly polynomial algorithm to solve combinatorial linear programs, *Operations Research* **34** (1986) 250-256.

[10] L.G. Khachiyan, Polynomial algorithms in linear programming, *Zhurnal Vychislitelnoj Matematiki i Matematicheskoi Fiziki* **20** (1980) 53-72 (Russian).

[11] A.V. Karzanov, Unbounded fractionality in maximum-value and minimum-cost maximum-value multiflow problems, in: A.A. Fridman, ed., *Problems of Discrete Optimization and Methods to Solve Them* (Central Economical and Mathematical Inst., Moscow, 1987) pp.123-135 (Russian).

[12] J. Edmonds and R. Giles, A min-max relation for submodular functions on directed graphs, in: P.L. Hammer, E.L. Johnson, and B.H. Korte, eds., *Studies in Integer Programming* (Annals of Discrete Mathematics, v. 1, 1977) 185-204.

[13] A.V. Karzanov, Polyhedra related to undirected multicommodity flows, *Linear Algebra and Its Applications* **114/115** (1989) 294-328.

[14] M. Grötshel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization* (Springer, Berlin et al, 1988).