A lower bound on the integrality gap for minimum multicut in directed networks

Michael Saks [*] Alex Samorodnitsky [†] Leonid Zosin[‡]

June 13, 2002

Abstract

Given a directed edge-weighted graph and k source-sink pairs, the *Minimum Directed Multicut Problem* is to find an edge subset with minimal weight, that separates each source-sink pair. Determining the minimum multicut in directed or undirected graphs is NP-hard. The fractional version of the minimum multicut problem is dual to the maximum multicommodity flow problem. The integrality gap for an instance of this problem is the ratio of the minimum weight multicut to the minimum weight fractional multicut; trivially this gap is always at least 1 and it is easy to show that it is at most k. In the analogous problem for undirected graphs this upper bound was improved to $O(\log k)$.

In this paper, for each k an explicit family of examples is presented each with k source-sink pairs for which the integrality gap can be made arbitrarily close to k. This shows that for directed graphs, the trivial upper bound of k can not be improved.

[*]Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019. This work was supported in part by NSF grant CCR-9700239 and by DIMACS. E-mail: `saks@math.rutgers.edu`.

[†]School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem 93708, Israel. This work was done while a postdoctoral fellow at DIMACS. E-mail: `salex@cs.huji.ac.il`.

[‡]ITG Inc. 44 Farnsworth Str. Boston, MA 02210. This work was done while a postdoctoral fellow at DIMACS. E-mail: `lzosin@itginc.com`.

1 Introduction

The *minimum multicut problem* is a well studied problem in combinatorial optimization. The setting of this problem is a *k-commodity network*, which consists of a directed graph (V, E), an edge capacity function $c : E \to R^+$, a list of k vertices $s_1, \ldots, s_k$ called the *sources*, and a list of k vertices $t_1, \ldots, t_k$ called the *sinks*. Each pair (s_i, t_i) is called a *commodity*. A vertex that is a source or a sink is called a *terminal*. We abuse notation and refer to both the network and its underlying directed graph by the same name. A *multicut* in a network G is a set of edges that intersects every directed path joining an s_i to the corresponding t_i. The capacity of a multicut is the sum of the capacities of its edges. The minimum multicut problem is to find the multicut for which the sum of the capacities is minimum.

The minimum multicut problem is naturally formulated as an integer linear program. We seek to assign $d(e) \in \{0, 1\}$ to each edge $e \in E$, so as to

$$\textbf{minimize} \textstyle\sum_{e \in E} d(e) \cdot c(e)$$

subject to the constraints:

Path covering: For each commodity i and each path P_i from s_i to t_i:

$$\sum_{e \in P_i} d(e) \geq 1 \tag{1}$$

As usual, with integer linear programs, we can consider the fractional relaxation of this problem, where we replace the constraint $d(e) \in \{0, 1\}$ by $d(e) \geq 0$. A feasible solution to the relaxed LP is called a *fractional multicut*. The dual linear program corresponds to the maximum multicommodity flow problem, which is to construct a flow between each source-sink pair so that the total flow across each edge obeys the capacity constraint, and the sum of the k individual flow values is maximized.

Let $\kappa(G)$ denote the minimum capacity of an integral multicut and $\kappa^*(G)$ denote the minimum capacity of a fractional multicut. Then $\kappa(G) \geq \kappa^*(G)$. As with many optimization problems that can be formulated as integer linear programs, it is of great interest to determine when the integral optimum coincides with the fractional optimum. In proving their celebrated max-flow min-cut theorem for single commodity networks, Ford and Fulkerson also showed that for such networks $\kappa(G) = \kappa(G^*)$. In 1963, T.C. Hu [Hu63] showed $\kappa(G) = \kappa^*(G)$ for undirected graphs in the 2-commodity case. For more than two commodities one can construct very simple examples to show that equality does not hold. However, equality has been established for special classes of graphs; a survey of these results appears in Section 8 of [Fr95].

For cases where equality does not hold, it is natural to consider the *integrality gap* which is the ratio $\kappa(G)/\kappa^*(G)$. In a seminal work, Leighton and Rao [LR88] introduced the notion of an approximate min-cut max-flow theorem. They observed that there are examples (namely, bounded degree expanders) where the integrality gap is $\Omega(\log k)$, which they showed was tight for some cases. For undirected graphs, Garg, Vazirani and Yannakakis [GVY93] showed that the integrality gap is $O(\log k)$. Tardos and Vazirani [TV93] obtain further results for graphs with excluded minors. In particular, the integrality gap of planar graphs is constant.

For directed graphs, there is a trivial upper bound of k on the integrality gap in any k-commodity network: let κ_i^* denote the minimum capacity of a fractional cut that separates s_i from t_i. By definition $\kappa^* \geq \max_i \kappa_i^*$ and since the integrality gap for $k = 1$ is 1, there is an integral cut separating s_i from t_i of value κ_i^* and so $\kappa \leq \sum_i \kappa_i^* \leq k \max_i \kappa_i^* \leq k\kappa^*$.

In this note we establish that the trivial upper bound of k is best possible by giving, for each $k \geq 2$, a family of instances of the directed k-commodity multicut problem for which the integrality gap can be made arbitrarily close to k.

2 A family of examples with large integrality gap

It is somewhat more convenient to formulate our examples as networks in which capacities are assigned to the non-terminal vertices rather than to the edges. In this formulation, the variables are $\{w_v : v \in V \backslash T\}$, where T is the set of terminals, and the objective is to minimize $\sum_{v \in V} c_v w_v$, where c_v is the capacity of vertex v. The "path covering" constraints are replaced by constraints that say the sum of w_v on any path from s_i to t_i is at least 1. This problem is called the *vertex multicut problem.* There is a standard reduction from an instance I of the vertex multicut problem to an instance I' of the edge multicut problem so that I and I' have the same integrality gap: replace each non-terminal vertex v in the digraph for I by two vertices v_{in} and v_{out} with a directed edge (v_{in}, v_{out}) of capacity c_v. All incoming edges to v go into v_{in} and all outgoing edges from v go from v_{out}. All edges in the network other than (v_{in}, v_{out}) edges are given infinite (or sufficiently large) capacity.

We now define the family of examples. For integer m, $[m]$ denotes the set $\{1, \ldots, m\}$. For positive integers n and k, let $C(n,k)$ be the undirected graph with vertex set $[n]^k$, where $x = (x_1, \ldots, x_k)$ is a neighbor of y iff $x_i - y_i \in \{-1, 0, 1\}$ for every i. Let $G(n,k)$ be the directed network obtained from $C(n,k)$ by replacing each undirected edge by two oppositely directed edges, and by adding k source-sink pairs $(s_1, t_1), \ldots, (s_k, t_k)$ and directed edges for every $i \in \{1, \ldots, k\}$ from the source s_i to every vertex $(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_k)$ and from every vertex $(x_1, \ldots, x_{i-1}, n, x_{i+1}, \ldots, x_k)$ to the sink t_i. Note that the vertices $s_1 ... s_k, t_1 ... t_k$ are new and distinct. Every non-terminal vertex $(x_1, \ldots, x_k)$ has unit capacity.

Theorem 2.1: *The integrality gap for the minimum multicut problem on $G(n,k)$ is at least $k(1-\frac{1}{n})^{k-1}$. Thus for $n = \Omega(k)$, the gap is $\Theta(k)$, and tends to k as n tends to infinity.*

3 Proof of the Gap theorem

Assigning weight $1/n$ to each non-terminal vertex gives a fractional cut of $G(n,k)$, since each source-sink path has at least n vertices. The capacity of this cut is n^{k-1}, so the theorem will follow from the following lower bound on the capacity of an integral cut.

Lemma 3.1: *Any integral multicut of $G(n,k)$ has cost at least $k \cdot (n-1)^{k-1}$.*

Proof: We start with some definitions. For $j \in [k]$, two points $x, y \in [n]^k$ are *j-collinear* if $x_i = y_i$ for all $i \neq j$. This is an equivalence relation, and the equivalence classes are called *j-lines.* Each j-line L consists of n points $x^1, \ldots, x^n$ where $x_i^1 = \ldots = x_i^n$ for $i \in [k] \setminus j$ and $x_j^t = t$ for $t \in [n]$. A j-line is *unexposed* if for each $i \neq j$, the common value of coordinate i for points on the line is at most $n-1$. Let $W(j)$ denote the union of all unexposed j-lines, i.e., the set of points whose coordinates other than j are bounded from above by $n-1$.

We define the set $F(j) = \{y \in \{0,1\}^k : y_j = 0\}$; $F(j)$ is a (lower) face of the unit k-dimensional hypercube. For $x \in [n]^k$, $F(x,j)$ denotes the translation of $F(j)$ by x, i.e., the set $\{x + y : y \in F(j)\}$. Note $F(x,j) \subseteq [n]^k$ if and only if $x \in W(j)$.

It might be helpful to visualize $[n]^k$ as an n-cube in the k-dimensional grid, and then $F(x,j)$ is a $(k-1)$-dimensional unit subcube of $[n]^k$, parallel to the j-face of the big cube, with a distinguished 'lower left' corner x.

Let $P(n,k) = \{(x,j) : x \in [n]^k, j \in [k]\}$. The reader should think of an element of $P(n,k)$ as a point of $[n]^k$ with a distinguished coordinate. For $Q \subseteq P(n,k)$, Q_j denotes the set $\{x : (x,j) \in Q\}$.

Now let C be a vertex multicut. Let $Z \subseteq P(n,k)$ be the set $\{(x,j) : F(x,j) \subseteq C\}$. Observe that for each $j \in [k]$, $Z_j \subseteq W(j)$.

Proposition 3.2: *For $j \in \{1,\ldots,k\}$, Z_j contains at least one point from each unexposed j-line.*

Proof: Suppose for contradiction that $Z_j \cap L = \varnothing$, where $L = x^1,\ldots,x^n$ is an unexposed j-line. For each $t \in [n]$ there is a point $y^t \in F(x^t,j) \setminus C$, and $y^t \in [n]^k$ since L is unexposed. Then $s_j \to y^1 \to \ldots \to y^n \to t_j$ is a path in $G(n,k) \setminus C$, contradicting the fact that C is a multicut. ∎

By the proposition, Z has a subset Y such that for each j, Y_j has *exactly* one point from each unexposed j-line. $|Y_j|$ equals the number of unexposed j-lines, which is $(n-1)^{k-1}$, so $|Y| = k(n-1)^{k-1}$. Let $U = \cup_{(x,j)\in Y} F(x,j)$; clearly $U \subseteq C$. We claim $|U| \geq |Y| (= k(n-1)^{k-1})$, which will give $|C| \geq |U| \geq k(n-1)^{k-1}$ as required.

We use linear algebra to prove the claim. Consider the space $(\mathbf{Z}_2)^U$ of functions ϕ mapping U to $\mathbf{Z}_2$. For each $(x,j) \in Y$, let $\phi_{x,j} \in (\mathbf{Z}_2)^U$ denote the characteristic function of $F(x,j)$. It now suffices to show that $\{\phi_{x,j} : (x,j) \in Y\}$ is a linearly independent set, since this would imply $|U| = \dim (\mathbf{Z}_2)^U \geq |Y|$. So the following proposition will complete the proof of the lemma and the gap theorem.

Proposition 3.3: *Let $Y \subseteq P(n,k)$ satisfy that for each $j \in [k]$, $Y_j \subseteq W(j)$ and Y_j has at most one point from any unexposed j-line. Let $\Phi_j = \{\phi_{x,j} : x \in Y_j\}$. Then $\Phi = \Phi_1 \cup \ldots \cup \Phi_k$ is linearly independent.*

Proof: First, we define the *degree* of $y \in [n]^k$, $d(y)$, to be the number of $(x,j) \in Y$ such that $y \in F(x,j)$.

Suppose the proposition is false, and let Y be a counterexample of minimum size. Then Φ is linearly dependent and any subset of Φ is linearly independent. Over $\mathbf{Z}_2$, this means that $\sum_{\phi\in\Phi} \phi \equiv 0 \bmod 2$; equivalently $d(y)$ is even for every $y \in [n]^k$. We now show that this is impossible.

Fix $h \in [k]$ such that $Y_h \neq \varnothing$ and among all $x \in Y_h$, choose x^* minimizing $\sum_{j\neq h} x_j$. Let L be the h-line containing x^*. We now show that $\sum_{y\in L} d(y) \equiv 1 \bmod 2$ which will contradict $d(y) \equiv 0 \bmod 2$ for all $y \in [n]^k$.

We have $\sum_{y\in L} d(y) = \sum_{(x,j)\in Y} |F(x,j) \cap L|$. For $(x,j) \in Y$ with $j \neq h$, $|F(x,h) \cap L|$ is 0 or 2, and thus $\sum_{y\in L} d(y) \equiv \sum_{x\in Y_h} |F(x,h) \cap L| \bmod 2$. If x is a point such that $F(x,h) \cap L \neq \varnothing$ then $x_j \leq x_j^*$ for all $j \neq h$. Since $x \in Y_h$, the choice of x^* implies $\sum_{j\neq h} x_j \geq \sum_{j\neq h} x_j^*$ and thus $x_j = x_j^*$ for each $j \neq h$ and so $x \in L$. By the hypothesis of the lemma, x^* is the only point in $L \cap Y_h$, and so $x = x^*$ and $F(x,h) \cap L = \{x\} = \{x^*\}$. Thus $\sum_{(x,j)\in Y} |F(x,j) \cap L| \equiv 1 \bmod 2$ which is the desired contradiction. ∎

4 Conclusions

One of the motivations for studying this question is to obtain better approximation algorithms for the minimum multicut problem. A standard approach to this is to solve the fractional problem and then round it to get an integral solution, and to use the integrality gap to bound the approximation ratio of the algorithm. Our examples indicate that such an approach by itself can not give an approximation ratio better than k.

References

[AMO] R.K. Ahuja, T.L. Magnanti, and J. Orlin, *Network Flows*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[ADK75] G.M. Adelson-Welsky, E.A. Dinits, and A.V. Karzanov. *Flow Algorithms* (in Russian), Nauka, Moscow, 1975.

[FF56] L.R. Ford and D.J. Fulkerson, "Maximal Flow through a network", Canadian J. of Math, 8, 399-404.

[Fr95] A. Frank, "Connectivity and Network Flows", in *Handbook of Combinatorics*, Volume I, R.L. Graham, M. Grötschel, L. Lovász, Eds., Elsevier Science B.V., Amsterdam, pp. 111-178, 1995.

[GVY93] N. Garg, V.V. Vazirani and M. Yannakakis, "Approximate max-flow min-(multi) cut theorems and their applications," *25th STOC,* pp. 698–707, 1993. To appear in SIAM Journal on Computing.

[Hu63] T.C. Hu, "Multicommodity network flows", *Operations Research*, 11:344-360, 1963.

[LR88] T. Leighton and S. Rao, "An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms," *29th FOCS*, pp. 422–431, 1988.

[Lo85] M. Lomonosov, "Combinatorial approaches to multiflow problems", *Discrete Applied Mathematics*, 11:1-93, 1985.

[OS89] H. Okamura and P.D. Seymour, "Multicommodity flows in planar graphs", *Journal of Combinatorial Theory (B)*, 31:75-81, 1989.

[RW66] B.Rothschild and A. Whinston, "On two-commodity network flows", *Operations Research*, 14:377-387, 1966.

[Se80] P. D. Seymour, "Four terminous flows", *Networks*, 10:79-86, 1980.

[Se81] P. D. Seymour, "Matroids and multicommodity flows", *European Journal of Combinatorics*, 2:257-290, 1981.

[TV93] E. Tardos and V. Vazirani, "Improved bounds for the max-flow min-multicut ratio for planar and $K_{r,r}$-free graphs", *Inform. Process. Lett.*, 47, 2:77-80, 1993.