College of Computer Science
Northeastern University
COM 3391: Advanced Algorithms

Lecture 8
17 November 1999
Scribe: Ivan Baev

The Integer Multicommodity Flow Problem

The following presentation is based on Section 11.2 of Chapter 11, due to Rajeev Motwani and Prabhakar Raghavan, in the book [Hoc95] edited by Dorit Hochbaum.

1 Problem definition

The multicommodity flow problem is a generalization of the maximum-flow problem. We are given a directed network (graph) $G = (V, E)$, a capacity $c(e)$ for each arc $e \in E$, and k commodities that need to be shipped across the network. Each commodity i is described by a triple (s_i, t_i, d_i) for its source node, destination node, and demand, respectively. A solution is a set of flows $f_i(e)$ denoting the amount of commodity i flowing on arc e. The flows are subject to the following constraints.

1. capacity: for all e, $\sum_{i=1}^{k} f_i(e) \le c(e)$.
2. flow conservation: for any node v different from s_i and t_i,

$$\sum_{e\in\delta^-(v)} f_i(e) = \sum_{e\in\delta^+(v)} f_i(e),$$

where $\delta^-(v)$ and $\delta^+(v)$ are the sets of arcs into and out of v, respectively.

3. demand: $\sum_{e\in\delta^+(s_i)} f_i(e) \ge d_i$.
4. integrality: all the flows $f_i(e)$ are integral.

As stated above, the problem may be infeasible if the capacities available do not suffice to satisfy the demands. This lecture considers the integer multicommodity flow problem where we do not enforce any demand constraints, the flow shipped for any commodity is less or equal to 1, and the goal is to maximize the total flow over all commodities.

2 An LP-based approximation algorithm

The integer multicommodity flow problem can be written as the following integer linear program (ILP).

$$\max F = \sum_{i=1}^{k} \sum_{e\in\delta^+(s_i)} f_i(e)$$

subject to

$$\begin{aligned}
\sum_{i=1}^{k} f_i(e) \le c(e) \qquad & e \in E\\
\sum_{e\in\delta^-(v)} f_i(e) = \sum_{e\in\delta^+(v)} f_i(e) \qquad & i = 1,\dots,k\\
\sum_{e\in\delta^+(s_i)} f_i(e) \le 1 \qquad & i = 1,\dots,k\\
f_i(e) \in \{0,1\} \qquad & e \in E,\ i = 1,\dots,k.
\end{aligned}$$

By relaxing the integrality constraint $f_i(e) \in \{0,1\}$ for all i, e, we obtain the linear program (LP) in which $f_i(e) \in [0,1]$ for all i, e. This LP, which can be solved efficiently, forms the basis for an approximation algorithm.

Definition 1 (Algorithm $\mathcal{A}$) *We perform the following three steps.*
0. (Preprocessing) Set the capacity of every arc e to $(1-\epsilon)c(e)$.
1. Solve the modified LP optimally to obtain the fractional solutions $f_i(e)$ and the corresponding total flow F.
2. For each commodity i start from source s_i and perform the following random walk:
- at s_i, pick an arc e out of s_i with probability $f_i(e)$;
- at any node v, pick an arc e out of v with probability $\frac{f_i(e)}{\sum_{e \in \delta^+(v)} f_i(e)}$.
The walk terminates on reaching t_i.

3 Analysis

We now proceed to investigate the quality of the solution produced by algorithm $\mathcal{A}$. First, we need the following two results.

Lemma 1 (Chernoff bound) *If $X_1, X_2, \ldots, X_n$ is a sequence of independent Bernoulli trials and $X = \sum_{i=1}^n X_i$, then*

$$\Pr[|X - E[X]| \geq \beta E[X]] < e^{-\frac{\beta^2 E[X]}{3}}$$

Lemma 2 *The probability that the random walk for commodity i ($rw(i)$) goes through arc e is equal to $f_i(e)$.*

Proof: (By induction on topological order of the graph)
Basis: If e is out of s_i, then it is easy to see that the claim is true.
Induction step: Let's assume that for all arcs e' into v, the starting endnode of e, $\Pr[rw(i)$ goes through $e'] = f_i(e')$. Then,

$$\begin{aligned}
\Pr[rw(i) \text{ goes through } e] &= \Pr[rw(i) \text{ goes through } v]\frac{f_i(e)}{\sum_{e'' \in \delta^+(v)} f_i(e'')} \\
\text{(by induction hypothesis)} &= \left(\sum_{e' \in \delta^-(v)} f_i(e')\right)\frac{f_i(e)}{\sum_{e'' \in \delta^+(v)} f_i(e'')} \\
\text{(by conservation constraint)} &= f_i(e).
\end{aligned}$$

■

Let random variable $X_{e,i}$ equal 1 if rw(i) goes through e, and 0 otherwise. For arc e, let random variable X_e denote the number of random walks that go through e. Then $X_e = \sum_i X_{e,i}$, and by linearity of expectation and Lemma 2, $E[X_e] = \sum_i E[X_{e,i}] = \sum_i f_i(e) \leq c(e)$. Note that $X_{e,i}$ form a sequence of independent Bernoulli trials. Using the Chernoff bound, $\Pr[X_e \geq c(e)(1+\beta)] < e^{-\frac{\beta^2 c(e)}{3}}$. Since we solve LP with capacities $c(e)(1-\epsilon)$, actually $E[X_e] \leq c(e)(1-\epsilon)$, and $\Pr[X_e \geq c(e)(1-\epsilon)(1+\beta)] < e^{-\frac{\beta^2 c(e)(1-\epsilon)}{3}}$. Select $\beta = \epsilon$. Then,

$$\Pr[X_e \geq c(e)(1-\epsilon^2)] < e^{-\frac{\epsilon^2 c(e)(1-\epsilon)}{3}}. \tag{1}$$

Let F^{OPT} denote the total flow in the optimal ILP solution. Then, $F \geq (1-\epsilon)F^{OPT}$, and by similar arguments as above the expectation of the total flow after rounding (i.e. the total flow produced by algorithm $\mathcal{A}$) is at least $(1-\epsilon)F^{OPT}$. Finally, let $m = |E|$, and we are ready for the main result.

Theorem 1 *If all $c(e)$ are "large" (say, at least $48 \ln m$), then "with high probability" (at least $1 - 1/m$), algorithm $\mathcal{A}$ yields a feasible (all capacity constraints satisfied) integer solution. Furthermore, the value of this solution, i.e. the total flow, is at least $(1-\epsilon)F^{OPT}$ "with high probability" (at least $1 - e^{-\frac{\epsilon^2 F^{OPT}(1-\epsilon)}{3}}$).*

Proof: Algorithm $\mathcal{A}$ always produces an integer solution, so it can only fail to meet the guarantees of the theorem because of the two types of failure:

1. The rounded flow violates the capacity constraint for some arc.
2. The rounded flow is less than $(1-\epsilon)F^{OPT}$.

We show that the probability of the former is bounded by $1/m$, and the latter by $e^{-\frac{\epsilon^2 F^{OPT}(1-\epsilon)}{3}}$.

By (1), $\Pr[X_e \geq c(e)] < \Pr[X_e \geq c(e)(1-\epsilon^2)] < e^{-\frac{\epsilon^2 c(e)(1-\epsilon)}{3}}$. Then, Pr[capacity constraint is violated for some arc] is less than $me^{-\frac{\epsilon^2 c(e)(1-\epsilon)}{3}}$, which for $c(e) \geq 48 \ln m$ is bounded by $1/m$.

We have already observed that the expectation of the total flow after rounding is at least $(1-\epsilon)F^{OPT}$. Using the Chernoff bound with $\beta = \epsilon$ gives the probability bound for the second type of failure. ■

Note that we can further decrease the failure probability of algorithm $\mathcal{A}$ by independent repetitions, and checking for the feasibility of flows and the lower bound on the value of total flow.

References

[Hoc95] D. S. Hochbaum, editor. *Approximation Algorithms for NP-hard Problems.* PWS Publishing Company, Boston, MA, 1995.