600.469 / 600.669 Approximation Algorithms **Lecturer:** Michael Dinitz
Topic: Randomized Rounding, Set Cover, Minimizing Conjestion **Date:** 3/10/15
Scribe: Torrey Kelly

11.1 Set Cover

Definition 11.1.1 *Given a universe U, collection $\mathcal{S} = \{S_1, \ldots, S_k\}$ with $S_i \subseteq U$ for each $i \in [k]$, and cost function $c : \mathcal{S} \to \mathbb{R}^+$: Construct a set $T \subset \mathcal{S}$ such that for all $e \in U$ there exists some $S_i \in T$ with $e \in S_i$, which minimize the total cost $\sum_{S_i \in T} c(S_i)$.*

In other words, find the set of sets in $\mathcal{S}$ that covers all elements in U with minimum cost. We can solve this using randomized rounding on a Linear Programming relaxation.

Consider linear program (**MIN-SET**) below.

$$\text{minimize:} \quad \sum_{S \in \mathcal{S}} c(S) \cdot X_S \qquad (\textbf{MIN-SET})$$

$$\text{subject to:} \quad \sum_{S : e \in S} X_S \geq 1 \quad \forall e \in U \qquad (11.1.1)$$

$$0 \leq X_S \leq 1 \quad \forall S \in \mathcal{S} \qquad (11.1.2)$$

Let λ be a constant (we will define its actual value later)

Let $\{X_S^*\}$, $S \in \mathcal{S}$ be the solution of the linear program.

The rounding algorithm works as follows: for each set $S \in \mathcal{S}$, let X_S be a random variable such that $X_S = 1$ with probability $\min(\lambda X_S^*, 1)$, else it equals 0. Then if $X_S = 1$, we add S to T

Algorithm 1 Randomized Rounding Set Cover Algorithm

Input: Universe U, collection $\mathcal{S} = \{S_1, \ldots, S_k\}$ such that $S_i \in U$ for all i, and cost function $c : \mathcal{S} \to \mathbb{R}$
Output: A subset of $\mathcal{S}$ containing all of U with minimum cost.

$T \leftarrow \emptyset$
$\{X_S^*\}$ = solution of **MIN-SET** on input
$\{X_S\} = \min(\lambda \cdot X_S^*, 1)$
for each $S \in \mathcal{S}$ **do**
  With probability X_s, $T \leftarrow T \cup \{S\}$
end for
return T

Theorem 11.1.2 *If we set $\lambda = \Theta(\log n)$, Randomized Rounding gives a collection T with expected cost at $O(\log n) \cdot LP$, and T is a valid set cover with high probability (at least $1 - 1/n$).*

In other words, randomized rounding is an $O(\log n)$-approximation algorithm for SET COVER.

We will begin by proving two helpful lemmas.

Lemma 11.1.3 $\mathbf{E}[c(T)] \leq \lambda \cdot LP$

Proof: $\mathbf{E}[c(T)] = \mathbf{E}\left[\sum_{S\in\mathcal{S}} c(S)X_S\right]$
Substitution

$\mathbf{E}\left[\sum_{S\in\mathcal{S}} c(S)X_S\right] = \sum_{S\in\mathcal{S}} c(S)\mathbf{E}[X_s]$
Moving $c(S)$ outside of the expectation, by linearity of expectations.

$\sum_{S\in\mathcal{S}} c(S)\mathbf{E}[X_s] = \sum_{S\in\mathcal{S}} c(S) \cdot \min(\lambda \cdot X_S^*, 1)$
Substituting $\min(\lambda \cdot X_S^*, 1))$ for $\mathbf{E}[X_s]$.

$\sum\limits_{S\in\mathcal{S}} (c(S) \cdot \min(\lambda \cdot X_S^*, 1)) \leq \sum\limits_{S\in\mathcal{S}} (c(S) \cdot \lambda X_S^*)$
The minimum of two values is less than or equal to either of the two values.

$\sum\limits_{S\in\mathcal{S}} (c(S) \cdot \lambda X_S^*) = \lambda \cdot \sum\limits_{S\in\mathcal{S}} (c(S) \cdot X_S^*)$
Pulling the constant out of the sum

$\lambda \cdot \sum\limits_{S\in\mathcal{S}} (c(S) \cdot X_S^*) = \lambda \cdot LP$
Substitution.

Thus $\mathbf{E}[c(T)] \leq \lambda \cdot LP$. ■

Lemma 11.1.4 *Let $u \in U$. Then $\mathbf{Pr}[u \text{ uncovered}] \leq e^{-\lambda}$*

Proof:

$\mathbf{Pr}[u \text{ uncovered}] = \mathbf{Pr}[X_S = 0 \quad \forall S \in \mathcal{S} : u \in S]$
u is uncovered only when none of the sets that contain it are in T.

Then $\mathbf{Pr}[u \text{ uncovered}] = \prod\limits_{S:e\in S} \mathbf{Pr}[X_S = 0]$, due to independence.

There are two cases

1. $\exists S : u \in$S such that $X_S^* \geq 1/\lambda$.
   Then $X_S^* \cdot \lambda \geq 1$
   so $\mathbf{Pr}[X_S = 0] = 0$, and hence
   $\mathbf{Pr}[u \text{ uncovered}] = 0 \leq e^{-\lambda}$

2. Otherwise
   $\prod\limits_{S:u\in S} \Pr[X_S = 0] = \prod\limits_{S:e\in S} (1\text{-}\lambda X_S^*)$
   $\prod\limits_{S:u\in S} (1\text{-}\lambda X_S^*) \leq \prod\limits_{S:u\in S} (e^{-\lambda X_S^*})$

$\prod_{S:u\in S}(e^{-\lambda X_S^*}) = e^{-\lambda \sum_{S:u\in S} X_S^*}$
$e^{-\lambda \sum_{S:u\in S} X_S^*} \leq e^{-\lambda}$ (by the LP constraint for u
Hence $\Pr[u \text{ uncovered}] \leq e^{-\lambda}$.

■

Proof of Theorem 11.1.2:

Set $\lambda = C \cdot \ln(n)$.

Lemma 11.1.3 implies that $\mathrm{E}[c(T)] \leq \mathrm{O}(\log(\mathrm{n}) \cdot \mathrm{LP})$

From Lemma 11.1.4, for every $u \in U$ we know that $\Pr[u \text{ uncovered}] \leq e^{-C \cdot ln(n)} = 1/n^C$. So by a simple union bound, $\Pr[T \text{ is not a set cover }] \leq 1/n^{C-1}$
Thus Randomized rounding returns a Set Cover with high probability by setting $C = 2$. ■

11.2 Minimizing Congestion

note: we did not finish this proof this class, so it is incomplete.
input: Graph $G = (V, E), k$ pairs $s_i, t_i : i \in [k]$
Feasible Solution: k paths $\{P_i, ..., P_k\}$ such that $\forall i, P_i$ is an $s_i - t_i$ path

Definition 11.2.1 *cong(e)* $= |\{i : e \in P_i\}|$

Objective: Minimize $\max_{e\in E}(cong(e))$

Definition 11.2.2 $\mathcal{P}_i$ = *set of all* $s_i - t_i$ *paths*

Consider linear program (**MIN-CONG**) below.

$$\begin{aligned}
\text{minimize:} \quad & W && \textbf{(MIN-CONG)} \\
\text{subject to:} \quad & \sum_{P\in\mathcal{P}_i} (X_P) = 1 \quad \forall i \in [k] && (11.2.3) \\
& \sum_{i=1}^{k} \sum_{P\in\mathcal{P}_i} X_P \leq W \quad \forall e \in E && (11.2.4) \\
& 0 \leq X_P \leq 1 \quad \forall P && (11.2.5)
\end{aligned}$$

Note: it is not clear how to solve **MIN-CONG**, since there are an exponential number of variables. However there does exist an equivalent linear program that can be solved (based on an edge formulation of flows), so we can assume that we can solve (**MIN-CONG**).

Consider the following randomized rounding algorithm:

Algorithm 2 Randomized Rounding MINIMIZING CONGESTION Algorithm

Input: Graph $G = (V, E), k$ pairs $s_i, t_i : i \in [k]$
Output: A Set of paths $P_1, \ldots, P_k$ so that $P_i \in \mathcal{P}_i$.

Solve **MIN-CONG** to get fractional solution (X^*, W^*)
for each $i \in [k]$ **do**
The values $\{X_P^*\}_{P \in \mathcal{P}_i}$ form a distribution over paths in $\mathcal{P}_i$ (since $\sum_{P \in \mathcal{P}_i} X_P^* = 1$)
Choose a path $P \in \mathcal{P}_i$ randomly from this distribution (each path P has probability equal to X_P^*)
end for

Theorem 11.2.3 *With high probability, the max congestion of the paths output by the rounding algorithm is at most $O(\log(n)) \cdot OPT$.*

Before proving this, lets first define some variables.

Definition 11.2.4 Y_e *= Number of chosen paths using* $e : e \in E$

Definition 11.2.5 Y_e^i *= 1 if path used for i contains e, otherwise 0*

Definition 11.2.6 Z_P^i *= 1 if P was chosen for i, otherwise 0*

By these definitions, we can see the following:

$$Y_e = \sum_{i=1}^{k} Y_e^i$$

The sum of all paths containing e is the number of paths using e.

$$Y_e^i = \sum_{\substack{P \in \mathcal{P}_i \\ e \in P}} Z_P^i$$

The number of paths chosen for i through e is equal to 1 if the path for i contains e, otherwise 0

$$Y_e = \sum_{i=1}^{k} \sum_{\substack{P \in \mathcal{P}_i \\ e \in P}} Z_P^i$$

Using substitution

Lemma 11.2.7 $E[Y_e] \leq W^*$

Proof:

$$\mathrm{E}[Y_e] = \sum_{i=1}^{k} \sum_{\substack{P \in \mathcal{P}_i \\ e \in P}} \mathrm{E}[Z_P^i]$$

Based on the above definitions and linearity of expectations.

$$\mathrm{E}[Y_e] = \sum_{i=1}^{k} \sum_{\substack{P \in \mathcal{P}_i \\ e \in P}} X_p^* \leq W^*$$

Using the fact that $\mathrm{E}[Z_P^i] = X_P^*$, and then using the constraints of **MIN-CONG** ■