Randomized rounding is an extremely effective means of designing approximation algorithms for NP-hard problems. The process consists of three steps.

1. Write down an integer programming formulation of the problem.
2. Relax it to a (fractional) convex program that can be solved efficiently (e.g. an LP or SDP).
3. Round the fractional solution to an integral solution and argue that that integral solution isn’t too much worse.

Today we’ll see this strategy applied to the min-congestion routing problem.

**Minimum congestion routing.** Let $D = (V, A)$ be a directed graph, and let $\{(s_i, t_i)\}_{i=1}^k$ be a collection of pairs of vertices. The goal is to choose one (directed) path $p_i$ from $s_i$ to $t_i$ for every $i = 1, 2, \ldots, k$ while minimizing the maximum congestion of an edge $e \in A$. The congestion of an edge $e$ is the number of paths that use $e$, i.e.

$$ \text{congestion}(e) = \#\{i : e \in p_i\}.$$ 

We will write $\text{congestion}(p_1, \ldots, p_k)$ for the maximum congestion of any edge under the choice of paths $p_1, \ldots, p_k$. And we will let $C^*$ be the optimal congestion over all possible valid solutions.

It is not difficult to construct examples where, e.g. the greedy strategy (given paths $p_1, \ldots, p_i$ choose path $p_{i+1}$ so as to minimize the current maximum congestion) can perform very badly compared to the optimum routing [exercise!] So let’s follow the three steps above in order to do better.

**Step 1) The integer program.**

First, we let $\mathcal{P}_i$ be the set of all paths in $G$ from $s_i$ to $t_i$. Note that the size of $\mathcal{P}_i$ might be exponential in $n = |V|$. For every $p \in \mathcal{P}_i$, we will have a variable $f_{p,i} \in \{0, 1\}$ representing whether we decide to route the $s_i$-$t_i$ path along $p$. Now our program is going to find the minimum of $C$ (representing the congestion) subject to the following constraints.

The first set of constraints expresses that we have to choose (at least) one path for every $i \in [k]$.

$$ \sum_{p \in \mathcal{P}_i} f_{p,i} \geq 1 \quad \text{for every } i = 1, 2, \ldots, k. $$

The next set of constraints represents the fact that the congestion is at most $C$.

$$ \sum_{i=1}^k \sum_{p : e \in p} f_{p,i} \leq C \quad \text{for every edge } e \in A. $$

The optimal value of the integer program is precisely $C^*$. 

1
Step 2) The relaxation.

Now we relax this integer program to a linear program by optimizing over \( f_{p,i} \in [0, 1] \) instead of \( f_{p,i} \in \{0, 1\} \). We can think of this linear program as finding a \textit{fractional} flow of value 1 from \( s_i \) to \( t_i \) for every \( i \), i.e. we are allowing the \( s_i\text{-}t_i \) flow to be split along multiple paths. Even though our LP is of exponential size (exponentially many variables, but only polynomially many constraints), it is well-known (and not difficult to see) that it has an efficient separation oracle, and thus we can still find the optimum \( C_{LP} \) efficiently. Since the LP is a relaxation, we have \( C_{LP} \le C^* \).

Step 3) Rounding the LP.

First of all, we can assume that for every \( i \in [k] \), we have \( \sum_{p \in P_i} f_{p,i} = 1 \) (since the fractional congestion can only get better by decreasing the flow value to 1). Thus we can interpret the LP variables \( f_{p,i} \) for \( p \in P_i \) as a probability distribution over paths in \( P_i \). This gives a natural rounding strategy:

Indepedently, for every \( i \in [k] \), choose route \( p \in P_i \) with probability \( f_{p,i} \).

Obviously this gives a valid routing. For each \( i \), let \( p_i \) be the randomly chosen \( s_i\text{-}t_i \) path. For \( e \in A \), let \( X_e = \# \{ i : e \in p_i \} \) be the congestion of edge \( e \). We would like to argue that, with high probability, \( \max_{e \in A} X_e \) is not too much bigger than \( C_{LP} \), and consequently not much bigger than \( C^* \).

Many times, when we have to bound a maximum, the best route is to first show that \( \mathbb{E}[X_e] \) is not too big, and then argue that \( X_e \) is highly concentrated around its mean (via a Chernoff bound); if this is true, we can take a union bound over all of the edges \( e \in A \).

So let’s calculate the expectation. Let \( Y_i \in \{0, 1\} \) be 1 exactly when \( e \in p_i \), then

\[
\mathbb{E}[X_e] = \mathbb{E} \left[ \sum_{i=1}^{k} Y_i \right] = \sum_{i=1}^{k} \sum_{p : e \in p} f_{p,i} \le C_{LP} \le C^*,
\]

So in expectation we are doing great. Now we have two regimes.

Optimal solution has high congestion. If \( C^* \gg \log n \), then we will argue that for every \( \epsilon > 0 \), \( \max_{e \in A} X_e \le (1 + \epsilon)C^* \) with high probability. Unfortunately, this does not follow immediately from the Hoeffding-Azuma bound that we proved previously. Recall that bound:

\[
\mathbb{P} \left[ \sum_{i=1}^{k} Y_i \ge \mathbb{E} \left( \sum_{i=1}^{k} Y_i \right) + L \right] \le e^{-L^2/(2\sum_{i=1}^{k} \|Y_i - \mathbb{E}Y_i\|_\infty^2)}.
\]

The preceding bound is not going to be very good when \( \|Y_i - \mathbb{E}Y_i\|_\infty \) is far from the typical value of \( |Y_i - \mathbb{E}Y_i| \). This is the case if, for instance, \( \mathbb{P}(Y_i = 1) \) is close to 0 as it may be in our case if the LP decides to route the \( s_i\text{-}t_i \) flow along many different paths.

Instead, let’s use some assumptions about the \( Y_i \)’s to do better. Let’s suppose that the random variables \( \{Y_i\}_{i=1}^{n} \) are mutually independent and that \( Y_i = 0 \) with probability \( 1 - q_i \) and \( Y_i = 1 \) with probability \( q_i \). We set \( \mu = \mathbb{E}[\sum_{i=1}^{k} Y_i] = \sum_{i=1}^{k} q_i \). Then we have the following.

**Theorem 0.1** (Chernoff bound). For any \( \delta > 0 \),

\[
\mathbb{P} \left[ \sum_{i=1}^{n} Y_i \ge (1 + \delta)\mu \right] \le \left( \frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right)^\mu.
\]
Proof. As before, we will use the “Laplace transform” of a random variable $X$, i.e. we will consider the moment generating function $e^{tX}$. We have, for any $t \geq 0$,

$$
\Pr \left( \sum_{i=1}^{n} Y_i \geq (1 + \delta) \mu \right) \leq \Pr \left( e^{t \sum_{i=1}^{n} Y_i} \geq e^{t(1+\delta)\mu} \right) \leq \frac{\mathbb{E}[e^{t \sum_{i=1}^{n} Y_i}]}{e^{t(1+\delta)\mu}} = \prod_{i=1}^{n} \frac{\mathbb{E}[e^{tY_i}]}{e^{t(1+\delta)\mu}},
$$

(1)

where we have used Markov’s inequality, and then independence of the $Y_i$’s. Now we write

$$
\mathbb{E}[e^{tY_i}] = (1 - q_i) + q_i e^{t} \leq e^{q_i(e^t - 1)}
$$

using the inequality $1 + x \leq e^x$ with $x = p_i(e^t - 1)$. Plugging this into (1) yields

$$
\Pr \left( \sum_{i=1}^{n} Y_i \geq (1 + \delta) \mu \right) \leq \frac{e^{\mu(e^t - 1)}}{e^{t(1+\delta)\mu}}.
$$

Finally, we can use elementary calculus to find the optimal value of $t$ for this bound; it turns out to be $t = \ln(1 + \delta)$, and plugging that in yields the claim.

Now we can use the Chernoff bound to write

$$
\Pr (X_e \geq (1 + \delta)\mathbb{E}[X_e]) = \Pr \left( \sum_{i=1}^{n} Y_i \geq (1 + \delta)\mathbb{E} \left[ \sum_{i=1}^{n} Y_i \right] \right) \leq \left( \frac{e^{\delta}}{(1 + \delta)^{1+\delta}} \right)^{C^*}
$$

If $C^* \gg \log n$, then for every $\delta > 0$, the right-hand side is asymptotically smaller than $n^{-3}$. Thus for $n$ large enough, we have $\Pr[X_e \geq (1 + \delta)C^*] \leq n^{-3}$, and hence by a union bound, $\Pr[\exists e \in A : X_e \geq (1 + \delta)C^*] \leq 1/n$, so almost surely our algorithm outputs a routing with congestion within $1 + \delta$ of optimal.

**Optimal solution has any congestion.** In this case, we set $\delta = \frac{\alpha \log n}{\log \log n}$ for some $\alpha > 0$ that we’ll choose momentarily. Applying the Chernoff bound again (and using the fact that $C^* \geq 1$), we have

$$
\Pr (X_e \geq (1 + \delta)\mathbb{E}[X_e]) \leq \frac{e^{\delta}}{(1 + \delta)^{1+\delta}} \leq n^{-3}
$$

for $\alpha$ chosen to be a large enough constant. Now a union bound shows that we are within $\Theta(\log n/(\log \log n))$ of the optimal congestion with high probability.

**Conclusion.** Remarkably, it was recently proved by Chuzhoy, Guruswami (our Venkat), Khanna, and Talwar that unless $P = NP$, the best approximation ratio that one can efficiently achieve for the min-congestion routing problem in directed graphs is $\Theta(\log n/(\log \log n))$, i.e. the above algorithm is optimal. For the undirected case, this is still the best upper bound, but for all we know there could be a much better algorithm (even a much better rounding algorithm for the above LP).