$\boxed{\text{1 Nov 2021}}$ Chernoff Bound continued!

<u>Chernoff Bound</u>: If $X_1, \ldots, X_N$ are independent $[0, 1]$-valued rand vars, and if $X = X_1 + \ldots + X_N$

$$x = \mathbb{E}[X].$$

then $\Pr\left(X > (1+\varepsilon)x\right)$ and

$$\Pr\left(X < (1-\varepsilon)x\right)$$

are both $< \exp\left(-\psi(\varepsilon) \cdot x\right)$.

where $\psi(\varepsilon) > 0$ for $\varepsilon > 0$,

$\psi(\varepsilon) > \frac{1}{3}\varepsilon^2$ for $\varepsilon < 1$.

<u>Proof:</u> We will focus on the quantity

$$\Phi_X(t) = \mathbb{E}\left[e^{tX}\right].$$

$$= \mathbb{E}\left[e^{tX_1 + tX_2 + \dots + tX_N} \right]$$

$$= \mathbb{E}\left[\prod_{i=1}^{N} e^{tX_i} \right]$$

$$= \prod_{i=1}^{N} \mathbb{E}\left[e^{tX_i} \right] = \prod_{i=1}^{N} \Phi_{X_i}(t).$$

Markov's Inequality: If Y is a rand var taking values in $[0, \infty)$

$$\Pr(Y > c \cdot \mathbb{E}Y) < \frac{1}{c}.$$

When $Y = e^{tX}$ then the event $X > (1+\varepsilon)x$ becomes

$$e^{tX} > e^{(1+\varepsilon)tx}$$

So $$\Pr(X > (1+\varepsilon)x)$$

$$= \Pr\left(e^{tX} > e^{(1+\varepsilon)tx} \right)$$

$$< \mathbb{E}\left[e^{tX} \right] \cdot e^{-(1+\varepsilon)tx}.$$

We have a battle of exponentials... choose t s.t. the $e^{-(1+\varepsilon)tx}$ term wins!

Working out $\Phi_X(t) = \prod_{i=1}^{N} E[e^{tX_i}]$.

For simplicity assume X_i is $\{0,1\}$-valued.

Letting $x_i = E[X_i]$, we have

w. prob. x_i ... $X_i = 1$, $e^{tX_i} = e^t$

w. prob. $1-x_i$ ... $X_i = \emptyset$, $e^{tX_i} = 1$

So
$$E[e^{tX_i}] = 1 - x_i + x_i e^t$$
$$= 1 + x_i(e^t - 1)$$
$$\leq e^{x_i(e^t-1)}$$

$$\prod_{i=1}^{N} E[e^{tX_i}] \leq e^{\left(\sum_{i=1}^{N} x_i\right)(e^t-1)}$$
$$= e^{x(e^t-1)}$$

$$\Pr(X > (1+\varepsilon)x) \leq e^{(e^t-1)x - (1+\varepsilon)tx}$$
$$= e^{[e^t - 1 - (1+\varepsilon)t]x}$$

Minimize $e^t - 1 - (1+\varepsilon)t$ by choosing t.

$\Longrightarrow$ set $t = \ln(1+\varepsilon)$, so
$$e^t - 1 - (1+\varepsilon)t = \varepsilon - (1+\varepsilon)\ln(1+\varepsilon).$$

So we've proven Chernoff bd with

$$\Psi(\varepsilon) = -\varepsilon + (1+\varepsilon)\ln(1+\varepsilon)$$

... for the event $X > (1+\varepsilon)x$.

For $\Pr(X < (1-\varepsilon)x)$ use $\mathbb{E}[e^{-tX}]$, go through same steps as above, find out $\Psi(\varepsilon) = \varepsilon + (1-\varepsilon)\ln(1-\varepsilon)$.

Is $\Psi(\varepsilon)$ positive? For small ε, use Taylor expansion...

$$\ln(1+\varepsilon) = \varepsilon - \frac{1}{2}\varepsilon^2 + \frac{1}{3}\varepsilon^3 - \frac{1}{4}\varepsilon^4 + \dots$$

$$(1+\varepsilon)\ln(1+\varepsilon) = \varepsilon + \frac{1}{2}\varepsilon^2 - \frac{1}{6}\varepsilon^3 + \frac{1}{12}\varepsilon^4 - \dots$$

$$\Psi(\varepsilon) = \frac{1}{2}\varepsilon^2 - \frac{1}{6}\varepsilon^3 + \frac{1}{12}\varepsilon^4 - \dots$$

$$\geq \frac{1}{3}\varepsilon^2 \quad \text{for } 0 < \varepsilon \leq 1.$$

For $\varepsilon \gg 1$, $(1+\varepsilon)\ln(1+\varepsilon)$ becomes the leading order term in $\Psi(\varepsilon)$.

Apply to randomized rounding of disjoint paths...

k random paths, $N = k$

$$X_i = \begin{cases} 1 & \text{if edge } (u,v) \text{ in path } P_i \\ 0 & \text{o.w.} \end{cases}$$

$$X = \sum_{i=1}^{N} X_i = \text{load}(u,v)$$

By solving LP, we got

$$\mathbb{E}X = r \cdot c(u,v)$$

where r is as small as possible.

In original problem formulation assume $r \geq 1$. so $r \cdot c(u,v) \geq c(u,v)$.

$$\Pr(\text{load}(u,v) > (1+\varepsilon) r\, c(u,v))$$
$$< e^{-\Psi(\varepsilon) r\, c(u,v)} \leq e^{-\Psi(\varepsilon) c(u,v)}.$$

$$\Pr(\exists \text{ an edge whose load is } > (1+\varepsilon) r\, c(u,v))$$
$$\leq m \cdot e^{-\Psi(\varepsilon) \min\{c(u,v)\}}.$$

Now choose ε such that

$$e^{-\Psi(\varepsilon)\min\{c(u,v)\}} \leq \frac{1}{2m}.$$

Then $\Pr\left(\text{rounding algo fails to achieve } (1+\varepsilon)\text{-approximation}\right)$

$$\leq m \cdot \frac{1}{2m} = \frac{1}{2}.$$

Easy to test for this.

Repeat rounding step with fresh randomness as many times as necessary (≤ 2 times in expectation) and you achieve $1+\varepsilon$ approx to optimum, **provided**

$$\Psi(\varepsilon) \cdot \min\{c(u,v)\} \geq \ln(2m).$$

Two observations:

① High capacity: if $c(u,v) > \frac{3}{\varepsilon^2}\ln(2m)$ for all (u,v) then

$$\Psi(\varepsilon) \cdot \min\{c(u,v)\}$$

$$\geq \cancel{\tfrac{1}{3}}\cancel{\varepsilon^2} \cdot \frac{\cancel{3}}{\cancel{\varepsilon^2}} \cdot \ln(2m) = \ln(2m).$$

② Low capacity: if $c(u,v) = 1$ for

some edge (u,v), then we need $\gamma(\varepsilon) \geq \ln(2m)$.

i.e.

$$(1+\varepsilon)\ln(1+\varepsilon) \geq 2\ln(2m)$$

$$1+\varepsilon \geq \frac{2\ln(2m)}{\ln(\ln(2m))}.$$

i.e. even if edges have low capacity, congestion $> O\left(\frac{\log m}{\log\log m}\right)$ is unlikely.