# PARALLEL DENSEST SUBGRAPH

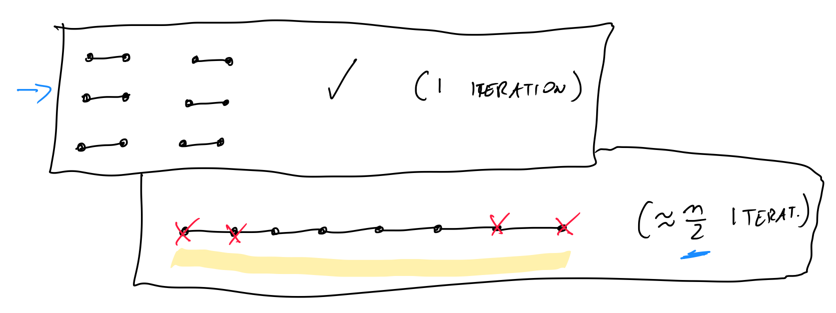
GREEDY $(G(V,E))$:

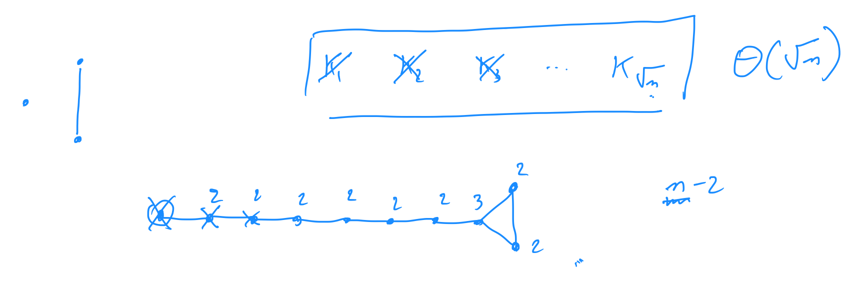
$S_0 \leftarrow V$
$i \leftarrow 0$
WHILE $S_i \neq \phi$:
  LET $v_i$ BE A NODE OF MINIMUM DEGREE IN $G[S_i]$
  $S_{i+1} \leftarrow S_i - \{v_i\}$
  $i \leftarrow i+1$
RETURN A $S_i$ MAXIMIZING $\rho(S_i) = \frac{|E(S_i)|}{|S_i|}$.

THIS ALGORITHM REQUIRES $n$ ITERATIONS.

PARALLEL_GREEDY $(G(V,E))$:
?

$S_0 \leftarrow V$
$i \leftarrow 0$
WHILE $S_i \neq \phi$:
  $A_i \leftarrow \{v \mid v \in S_i \wedge deg_{S_i}(v) = \min_{u \in S_i} deg_{S_i}(u)\}$
  $S_{i+1} \leftarrow S_i - A_i$
  $i \leftarrow i+1$
RETURN A $S_i$ MAXIMIZING $\rho(S_i) = \frac{|E(S_i)|}{|S_i|}$.




GREEDY$_{AVG}$ $(G(V,E))$:

$S_0 \leftarrow V$
$i \leftarrow 0$
WHILE $S_i \neq \phi$:
  LET $v_i$ BE A NODE SUCH THAT $deg_{S_i}(v_i) \le \operatorname{avg}_{v \in S_i} deg_{S_i}(v)$.
  $S_{i+1} \leftarrow S_i - \{v_i\}$
  $i \leftarrow i+1$
RETURN A $S_i$ MAXIMIZING $\rho(S_i) = \frac{|E(S_i)|}{|S_i|}$.

EX: PROVE THAT GREEDY$_{AVG}$ RETURNS A 2-APPRX.
(YOU JUST NEED TO CHANGE 1 LINE OF 1 LEMMA).

PARALLEL_GREEDY$_{AVG}$ $(G(V,E))$:
?

$S_0 \leftarrow V$
$i \leftarrow 0$
WHILE $S_i \neq \phi$:
  $A_i \leftarrow \{v \mid v \in S_i \wedge deg_{S_i}(v) \le \operatorname{avg}_{u \in S_i} deg_{S_i}(u)\}$
  $S_{i+1} \leftarrow S_i - A_i$
  $i \leftarrow i+1$
RETURN A $S_i$ MAXIMIZING $\rho(S_i) = \frac{|E(S_i)|}{|S_i|}$.



$2\left(1-\frac{1}{n}\right) = \frac{2n-2}{n} = \frac{2\cdot(n-2)+1\cdot 2}{n}$

GREEDY$_\varepsilon$ $(G(V,E))$:

BAHMANI
KUMAR
VASSILVITSKII
$\sim 2012$

$S_0 \leftarrow V$
$i \leftarrow 0$
WHILE $S_i \neq \phi$:
  $A_i \leftarrow \{v_i \mid v_i \in S_i \wedge deg_{S_i}(v_i) \le (1+\varepsilon) \operatorname{avg}_{u \in S_i} deg_{S_i}(u)\}$
  $S_{i+1} \leftarrow S_i - A_i$
RETURN THE "BEST" $S_i$.

WE WANT TO PROVE THAT
- GREEDY$_\varepsilon$ RETURNS A GOOD APPRX TO DS;
- " RUNS QUICKLY.

L1: GREEDY$_\varepsilon$ RETURNS A $2\cdot(1+\varepsilon)$-APPROXIMATION, $\forall \varepsilon \ge 0$.

P: LET $S^* \neq \phi$ BE AN OPTIMAL SOLUTION. (THEN $|S^*| \ge 2$, WLOG)

CLAIM 1: $\forall v \in S^*$, $deg_{S^*}(v) \ge \rho(S^*) \triangleq \frac{|E(S^*)|}{|S^*|}$.

P: $\rho(S^*) = \frac{|E(S^*)|}{|S^*|} \ge \rho(S^* - \{v\}) = \frac{|E(S^*-\{v\})|}{|S^*-\{v\}|} = \frac{|E(S^*)| - deg_{S^*}(v)}{|S^*|-1}$

BY THE OPTIMALITY OF $S^*$

$$\frac{|E(S^*)| - deg_{S^*}(v)}{|S^*|-1} \le \frac{|E(S^*)|}{|S^*|}$$

$$|E(S^*)| - deg_{S^*}(v) \le \frac{|S^*|-1}{|S^*|} \cdot |E(S^*)| = \left(1 - \frac{1}{|S^*|}\right) \cdot |E(S^*)|$$

$$\cancel{|E(S^*)|} - deg_{S^*}(v) \le \cancel{|E(S^*)|} - \frac{|E(S^*)|}{|S^*|}$$

$$\frac{|E(S^*)|}{|S^*|} \le deg_{S^*}(v). \quad \square$$

CLAIM 2: AT LEAST ONE NODE IS REMOVED IN EACH ITERATION.

P: $\min_{v \in S_i} deg_{S_i}(v) \le \operatorname{avg}_{v \in S_i} deg_{S_i}(v) \le (1+\varepsilon) \cdot \operatorname{avg}_{v \in S_i} deg_{S_i}(v)$.

THUS, $|A_i| \ge 1$ IF $|S_i| \ge 1$. $\square$

CONSIDER THE 1ST ITERATION $i$ SUCH THAT AT LEAST ONE NODE OF $S^*$ IS REMOVED FROM $S_i$.

THAT IS, LET $i$ BE THE SMALLEST INTEGER S.T.
$A_i \cap S^* \neq \phi$.

LET $v \in A_i \cap S^*$. SINCE $i$ IS THE FIRST ITERATION IN WHICH $A_i \cap S^* \neq \phi$, IT MUST BE THAT $S_i \supseteq S^*$.

$$\rho(S^*) \underset{\text{CLAIM 1}}{\le} deg_{S^*}(v) \underset{S_i \supseteq S^*}{\le} deg_{S_i}(v) \underset{\text{GREEDY CHOICE}}{\le} (1+\varepsilon) \cdot \operatorname{avg}_{u \in S_i} deg_{S_i}(u)$$

$$= (1+\varepsilon) \frac{\sum_{u \in S_i} deg_{S_i}(u)}{|S_i|}$$

$$= (1+\varepsilon) \frac{2|E(S_i)|}{|S_i|} = 2\cdot(1+\varepsilon)\cdot\rho(S_i).$$

THUS, $\rho(S_i) \ge \frac{1}{2\cdot(1+\varepsilon)} \cdot \rho(S^*)$. THUS, GREEDY$_\varepsilon$ RETURNS SOMETHING NO WORSE THAN A $2\cdot(1+\varepsilon)$-APPRX. $\square$

L2: GREEDY$_\varepsilon$ ITERATES FOR AT MOST $O\left(\frac{\log n}{\varepsilon}\right)$ TIMES, $\forall \varepsilon > 0$.

P: FIX AN ITERATION $i$.

$$2|E(S_i)| = \sum_{v \in S_i} deg_{S_i}(v) = \sum_{v_i \in A_i} deg_{S_i}(v) + \sum_{v \in S_i - A_i} deg_{S_i}(v)$$

$$\ge 0 + \sum_{v \in S_i - A_i} deg_{S_i}(v)$$

$$> \sum_{v \in S_i - A_i} \left(\operatorname{avg}_{u \in S_i} deg_{S_i}(u)\right)(1+\varepsilon)$$

$$= \sum_{v \in S_i - A_i} \left(\frac{2|E(S_i)|}{|S_i|}(1+\varepsilon)\right)$$

$$= |S_i - A_i| \frac{2|E(S_i)|}{|S_i|}(1+\varepsilon)$$

$A_i \subseteq S_i$

$$= (|S_i| - |A_i|) \frac{2|E(S_i)|}{|S_i|}(1+\varepsilon).$$

$$2|E(S_i)| > (|S_i| - |A_i|) \frac{2|E(S_i)|}{|S_i|}(1+\varepsilon)$$

$$= \left(1 - \frac{|A_i|}{|S_i|}\right) 2|E(S_i)|(1+\varepsilon)$$

$$= \left(1 - \frac{|A_i|}{|S_i|}\right) 2|E(S_i)| + \left(1 - \frac{|A_i|}{|S_i|}\right) 2|E(S_i)| \cdot \varepsilon$$

$$= 2|E(S_i)| - 2|A_i| \frac{|E(S_i)|}{|S_i|} + \left(1 - \frac{|A_i|}{|S_i|}\right) 2|E(S_i)| \cdot \varepsilon$$

$$1 > 1 - \frac{|A_i|}{|S_i|} + \left(1 - \frac{|A_i|}{|S_i|}\right)\varepsilon$$

$$\cancel{|S_i|} > \cancel{|S_i|} - |A_i| + (|S_i| - |A_i|)\varepsilon$$

$$(1+\varepsilon)|A_i| \ge \varepsilon |S_i|$$

$$|A_i| \ge \frac{\varepsilon}{1+\varepsilon}|S_i| \quad ✓$$

NOW, RECALL THAT $|A_i| = |S_i| - |S_{i+1}|$. THUS,

$$|S_i| - |S_{i+1}| \ge \frac{\varepsilon}{1+\varepsilon}|S_i|$$

$$|S_i|\left(1 - \frac{\varepsilon}{1+\varepsilon}\right) \ge |S_{i+1}|$$

THUS,

$$|S_{i+1}| \le \frac{1}{1+\varepsilon}|S_i|.$$

THEN, $|S_0| = n$.

$$|S_1| \le \frac{1}{1+\varepsilon}|S_0| = (1+\varepsilon)^{-1} \cdot n$$

$$|S_2| \le \frac{1}{1+\varepsilon}|S_1| \le (1+\varepsilon)^{-2} \cdot n$$

$$\vdots$$

$$|S_i| \le (1+\varepsilon)^{-i} \cdot n$$

$$\vdots$$

CONSIDER $i^* = \lceil \log_{1+\varepsilon} n \rceil$. EITHER THE PROCESS ENDS BEFORE ITERATION $i^*$, OR,

$$|S_{i^*}| \le (1+\varepsilon)^{-i^*} \cdot n \le (1+\varepsilon)^{-\log_{1+\varepsilon} n} \cdot n = \frac{1}{n} \cdot n = 1$$

THE NUMBER OF ITERATIONS IS THEN AT MOST $i^*$:

$$i^* = O(\log_{1+\varepsilon} n) = O\left(\frac{\log n}{\log(1+\varepsilon)}\right) = O\left(\frac{\log n}{\varepsilon}\right). \quad \square$$

$\ln(1+\varepsilon) = \varepsilon + \frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3} + \dots + \frac{\varepsilon^i}{i} + \dots$
FOR $0 < \varepsilon < 1$