به نام خدا

مسئله selection : پیدا کردن k امین کوچکترین عدد در بین n عدد

یعنی عددی که دقیقاً $k-1$ عدد از آن کوچکتر است.

حالت های خاص مسئله:

$k=1$ ← عدد کمینه Minimum ← $O(n)$ زمان

$k=n$ ← عدد بیشینه Maximum ← $O(n)$ زمان

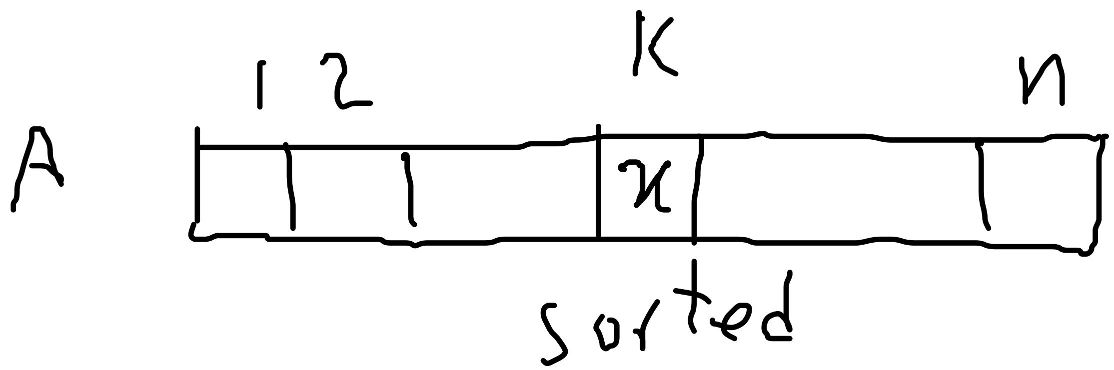
زمان حل مسئله در حالت کلی؟

راه‌حل‌های مسأله selection :

① از مرتب‌سازی استفاده کنیم :

sort می‌کنیم و عددی که در محل k ام است جواب است

$(A[k])$



زمان مرتب‌سازی $\Theta(n^2)$ یا $O(n\lg n)$

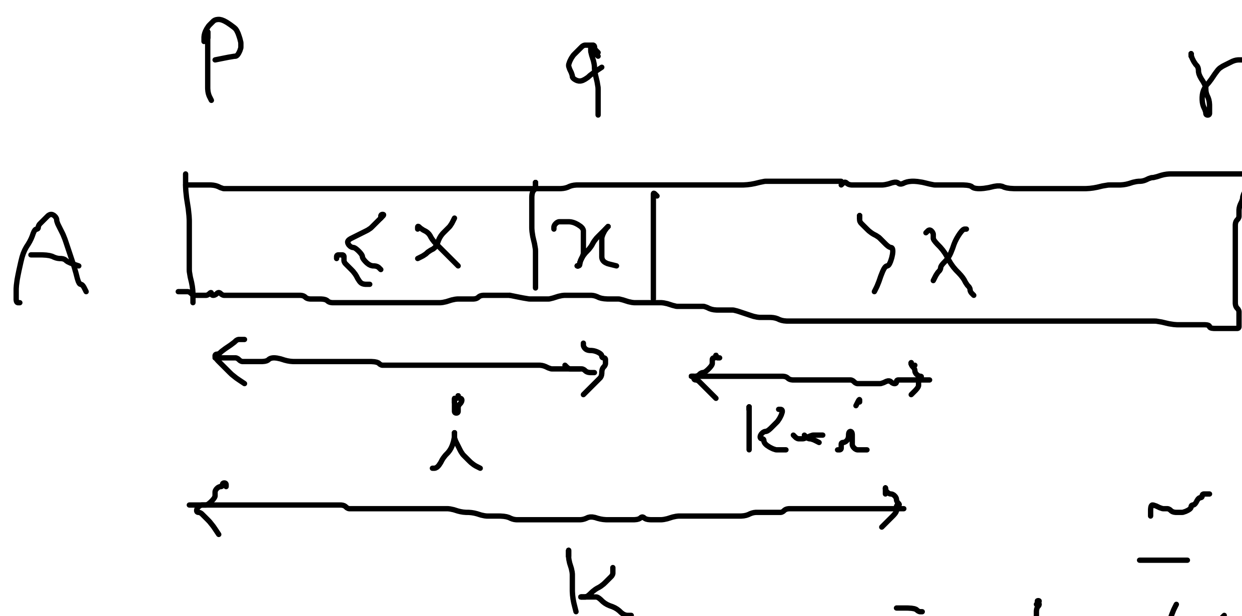
(زمان الگوریتم‌های مرتب‌سازی که تا الان بررسی شده $\Omega(n\lg n)$ است)

زمان این روش = زمان مرتب‌سازی $O(n\lg n)$

روش دوم: از partition استفاده کنیم

روش بازگشتی

روش تقسیم و غلبه



$i = q - p + 1$

شاخص x در آرایه partition شده i ام است.

اگر $K = i$ آنگاه جواب $x = A[q]$ است.

اگر $K < i$ آنگاه جواب عدد K ام در نیمه اول آرایه است.

اگر $K > i$ آنگاه جواب عدد $(K-i)$ ام در نیمه دوم آرایه است.

cibj $T(n)$

selection (A, p, r, k)

$O(1)$ 1. if $p \geq r$ then return A[p]

$\theta(n)$ 2. $q \leftarrow$ partition (A, p, r)

$O(1)$ 3. $i \leftarrow q - p + 1$

$O(1)$ 4. if $i = k$ then return A[q]

$O(1)$ 5. else if $k < i$ then

6. $T(i-1)$ return selection(A, p, q-1, k)

$O(1)$ 7. else if $k > i$ then

8. $T(n-i)$ return selection(A, q+1, r, k-i)

در بهترین حالت: شرط دستور خط ۴ برقرار است

و فراخوانی بازگشتی نداریم:

بهترین حالت

$$T(n) = O(1) + \Theta(n) + O(1) + O(1) \in \Theta(n)$$

پس $\underbrace{T(n)}_{\text{زمان الگوریتم}} \in \Omega(n)$

بدترین حالت: وقتی است که شرط خط 4 برقرار نیست

و از شرط خط 5 و 7 ، شرط برقرار است که منجر

به فراخوانی الگوریتم روی نیمه بزرگتر آرایه می‌شود:

بدترین حالت

$$T(n) = O(1) + \Theta(n) + O(1) + O(1) + O(1) + \max\{T(i-1), T(n-i)\}$$

و بدترین حالت هم این است که

$$\max\{i-1, n-i\} = n-1$$

یعنی منیمه بزرگتر آرایه $n-1$ عدد داشته باشد

$$T(n) = T(n-1) + \Theta(n)$$

بدترین حالت

$$= \Theta(n^2)$$

پس $T(n) \in O(n^2)$ زمان

الگوریتم

الگوریتم selection دارای زمان $\Omega(n)$ و

$O(n^2)$ است.

حالت متوسط هم می‌توان محاسبه کرد.

اگر شاخص عددی نزدیک میانه باشد به طوری که
تقریباً نصف اعداد از آن کوچکتر و نصف دیگر بزرگتر
باشند خواهیم داشت:

بهترین حالت

$$T(n) = 1T\left(\frac{n}{2}\right) + \Theta(n)$$

با فرض میانه
بودن شاخص

$$= \Theta(n)$$

حالت سوم
قضیه اصلی

پیدا کردن میانه: یعنی پیدا کردن k امین عدد

که $k = \frac{n}{2}$ است!

پس به نظر میرسد که برای اینکه بتوان selection

را در زمان خوبی حل کرد به میانه نیاز داریم

که برای بدست آوردن میانه باید خود

مساله selection را حل کنیم. ← loop?

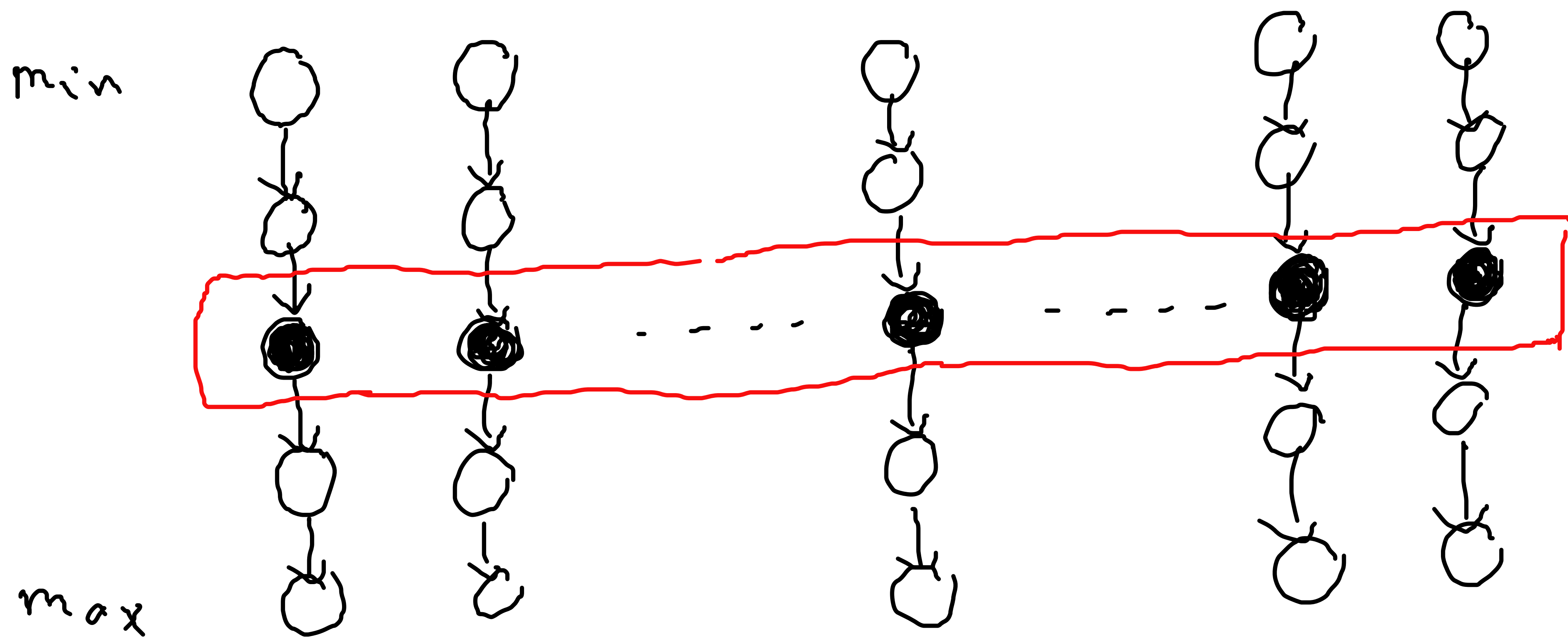
① selection (A, p, r, k) ورودی n عدد

② selection $(A, p, r, \frac{n}{2})$ ورودی n عدد

مخصوص میانه

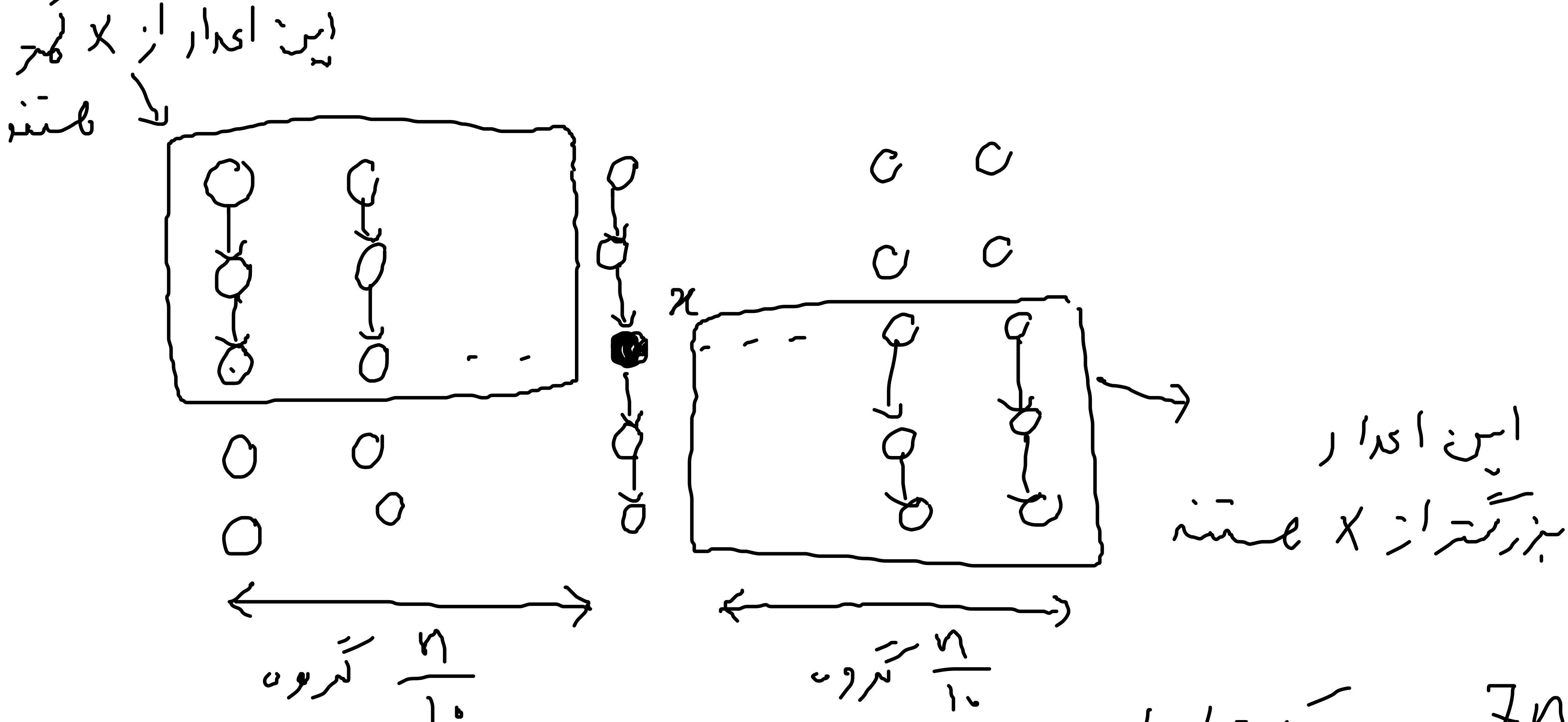
برای بهبود زمان الگوریتم selection ،
شاخص را یک عدد «نزدیک میانه» انتخاب
می‌کنیم. نحوه انجام کار:

ابتدا اعداد را به گروه‌های 5 تایی تقسیم می‌کنیم $\lceil \frac{n}{5} \rceil$ گروه
هر گروه را مرتب می‌کنیم



الگوریتم را برای این میانه‌های گروه‌ها فراخوانی می‌کنیم

میانه این میانه ها را اگر x فرض کنیم و آن را شاخص بگیریم
چه تعداد عدد از آن کمتر یا بیشتر هستند؟
گروه هایی که میانه آنها از x کمترات سمت چپ و آنهایی که میانه
آنها از x بیشتر است سمت راست قرار دارند. تعداد گروه ها $\frac{n}{5}$



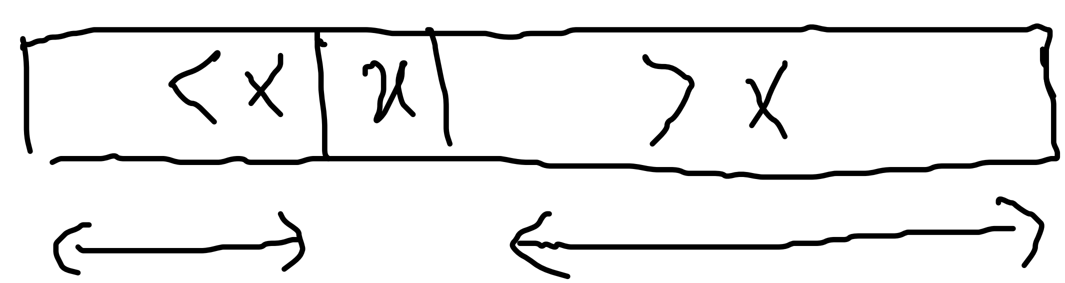
حداقل اعداد کوچکتر از x $= 3 \times \frac{n}{10}$ $\longrightarrow$ حداکثر تعداد اعداد بزرگتر از x $= \frac{7n}{10}$

همینطور برای اعداد بزرگتر از x :

حداقل تعداد اعداد بزرگتر از x $= 3 \times \frac{n}{10}$ $\longrightarrow$ حداکثر مقدار اعداد کوچکتر از x $= \frac{7n}{10}$

* بنابراین حداکثر سایز بقیه آرایه بزرگتر از x یا بقیه آرایه کوچکتر از x، $\frac{7n}{10}$ است.

$\frac{7n}{10}$ حداکثر



حداکثر $\frac{7n}{10}$

پس: با فرض استفاده از میانه میانه ها به عنوان شاخص داریم

در بدترین حالت $T(n) = T\left(\frac{n}{5}\right) + T\left(\frac{7n}{10}\right) + \theta(n)$

میانه میانه ها

فراخوانی بازگشتی

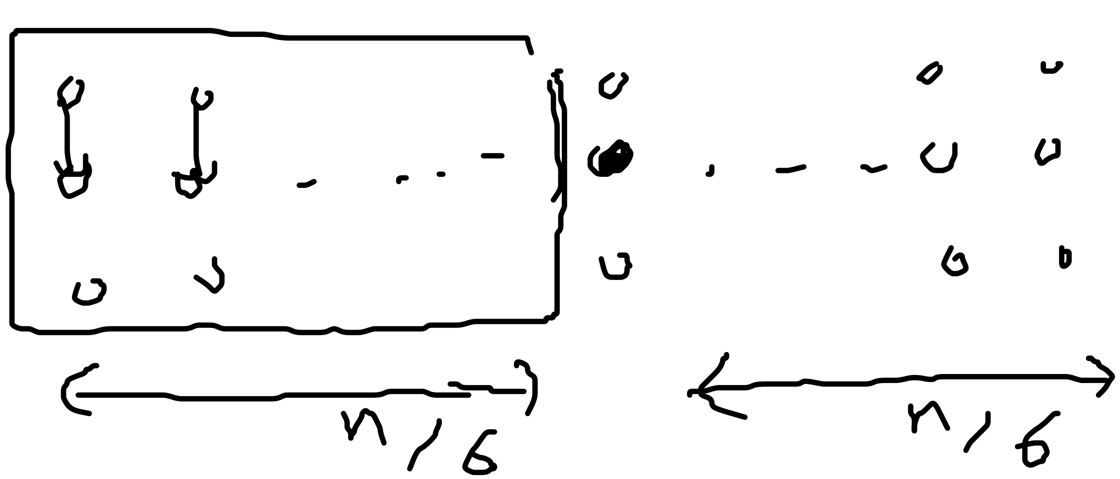
$$= T\left(\frac{2n}{10}\right) + T\left(\frac{7n}{10}\right) + \theta(n)$$

$$= \theta(n)$$

اگر گروه ها را ۳ تایی بگیریم :

$$T(n) = T\left(\frac{n}{3}\right) + T\left(\frac{2n}{3}\right) + \theta(n)$$

$$= ?$$



$\boxed{2 \times \frac{n}{6}}$ عدد کوچکتر از x

حداکثر $\frac{2n}{3}$ بزرگتر از x

improved-selection (A, p, r, k)

1. if $p \geq r$ then return A[p]

2. تقسیم اعداد به $\lceil \frac{n}{5} \rceil$ گروه ۵ و مرتب سازی هر گروه

تعداد $\frac{n}{5}$ و $k = \frac{n}{10}$

3. x = improved-selection(مجموعه میانه‌های گروه‌ها)

4. q = partition(A, p, r) با شاخص x

5. $i = q - p + 1$

6. if $i = k$ then return A[q]

7. if $k < i$ then

8. return improved-selection(A, p, q-1, k)

9. if $k > i$ then

10. return improve-selection(A, q+1, r, k-i)

بهترین حالت :

بهترین حالت

$$T(n) = T\left(\frac{n}{5}\right) + \Theta(n)$$

$$= \Theta(n)$$

بدترین حالت :

بدترین حالت

$$T(n) = T\left(\frac{n}{5}\right) + T\left(\frac{7n}{10}\right) + \Theta(n)$$

$$= \Theta(n)$$