Αλγόριθμοι και Πολυπλοκότητα

Ν. Μ. Μισυρλής

Τμήμα Πληροφορικής και Τηλεπικοινωνιών,
Πανεπιστήμιο Αθηνών

Επιλογή

Το πρόβλημα της επιλογής

Εισοδος: Ένα σύνολο A με n (διαφορετικούς) αριθμούς και ένας αριθμός i, όπου $1 \leq i \leq n$.

Εξοδος: Το στοιχείο $x \in A$ το οποίο είναι μεγαλύτερο από ακριβώς $i - 1$ άλλα στοιχεία του A.

- Μπορεί να επιλυθεί σε χρόνο $O(n \log n)$.
- Ταξινομούμε τους αριθμούς και στη συνέχεια απλώς επιλέγουμε το i-οστό στοιχείο.
- Ωστόσο, υπάρχουν και ταχύτεροι αλγόριθμοι για την επίλυση του προβλήματος αυτού.

Επιλογή

Τυχαιοκρατικη Επιλογη

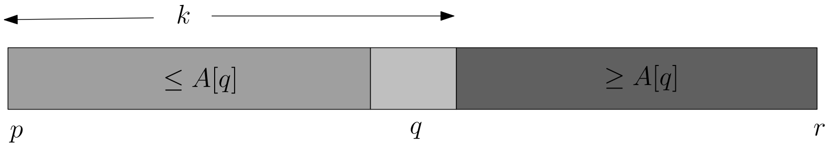
- Ο αλγόριθμος Τυχαιοκρατικη Επιλογη βασίζεται στη λογική του αλγορίθμου της ταχυταξινόμησης. Η βασική ιδέα είναι να διαμερίσουμε τη συστοιχία εισόδου αναδρομικά.
- Αντίθετα με την ταχυταξινόμηση, όμως, η οποία επεξεργάζεται αναδρομικά και τα δυο σκέλη της διαμέρισης, η Τυχαιοκρατικη Επιλογη περιορίζεται στο ένα σκέλος.
- Ενώ η ταχυταξινόμηση έχει αναμενόμενο χρόνο εκτέλεσης $\Theta(n \log n)$, ο αναμενόμενος χρόνος της Τυχαιοκρατικη Επιλογη είναι $\Theta(n)$, υπό την προϋπόθεση ότι τα στοιχεία εισόδου είναι διαφορετικά μεταξύ τους.
- Ο χρόνος εκτέλεσης χειρότερης περίπτωσης της Τυχαιοκρατικη Επιλογη είναι $\Theta(n^2)$, ακόμη και για να βρούμε το ελάχιστο.

Επιλογή

Τυχαιοκρατική Επιλογή

Τυχαιοκρατική Επιλογή (A, p, r, i)

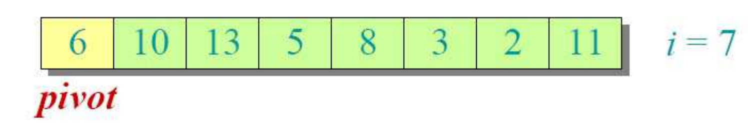
αν $p = r$
    τότε **επιστροφή** $A[p]$
$q \leftarrow$ Τυχαιοκρατική Διαμέριση (A, p, r)
$k \leftarrow q - p + 1$
αν $i = k$ ▷ το ζητούμενο στοιχείο είναι το οδηγό
    τότε **επιστροφή** $A[q]$
αλλιώς-αν $i < k$
    τότε **επιστροφή** Τυχαιοκρατική Επιλογή $(A, p, q - 1, i)$
αλλιώς **επιστροφή** Τυχαιοκρατική Επιλογή $(A, q + 1, r, i - k)$



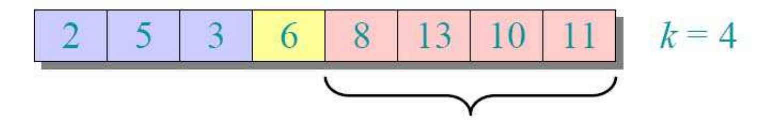
Επιλογή

Παράδειγμα

Βρείτε το μικρότερο έβδομο ($i = 7$) στοιχείο.



Διαμέριση:



Επιλογή του $7 - 4 = 3^{ου}$ στοιχείου αντίστοιχα

Επιλογή

Πολυπλοκότητα κατά μέσο όρο

$$E[T(n)] \leq \sum_{k=1}^{n} \frac{1}{n} \cdot E[T(\max(k-1, n-k))] + O(n).$$

$$\max(k-1, n-k) = \begin{cases} k-1 & \text{εάν } k > \lceil n/2 \rceil, \\ n-k & \text{εάν } k \leq \lceil n/2 \rceil. \end{cases}$$

Ανω φράγμα της μέσης τιμής $E[T(n)]$

$$E[T(n)] \leq \frac{2}{n} \sum_{k=\lfloor n/2 \rfloor}^{n-1} E[T(k)] + O(n).$$

Επιλογή

Επίλυση αναδρομικής σχέσης

- Ας υποθέσουμε ότι $E[T(n)] \leq cn$ για κάποια σταθερά $c > 0$ η οποία ικανοποιεί τις αρχικές συνθήκες της αναδρομικής σχέσης. .
- Επίσης, υποθέτουμε ότι $T(n) = O(1)$ για $n < n_0$, το n_0 θα επιλεγεί στη συνέχεια.

Επιλογή

$$
\begin{aligned}
E[T(n)] &\leq \frac{2}{n} \sum_{k=\lfloor n/2 \rfloor}^{n-1} ck + an \\
&= \frac{2c}{n} \left(\sum_{k=1}^{n-1} k - \sum_{k=1}^{\lfloor n/2 \rfloor - 1} k \right) + an \\
&= \frac{2c}{n} \left(\frac{(n-1)n}{2} - \frac{(\lfloor n/2 \rfloor - 1)\lfloor n/2 \rfloor}{2} \right) + an \\
&\leq \frac{2c}{n} \left(\frac{(n-1)n}{2} - \frac{(n/2-2)(n/2-1)}{2} \right) + an \\
&= \frac{2c}{n} \left(\frac{n^2-n}{2} - \frac{n^2/4 - 3n/2 + 2}{2} \right) + an \\
&= \frac{c}{n} \left(\frac{3n^2}{4} + \frac{n}{2} - 2 \right) + an \\
&= c \left(\frac{3n}{4} + \frac{1}{2} - \frac{2}{n} \right) + an \\
&\leq \frac{3cn}{4} + \frac{c}{2} + an \\
&= cn - \left(\frac{cn}{4} - \frac{c}{2} - an \right).
\end{aligned}
$$

Επιλογή

Επίλυση αναδρομικής σχέσης (Τελικό βήμα)

$$cn/4 - c/2 - an \geq 0$$

$$n(c/4 - a) \geq c/2$$

$$c/4 - a > 0$$

$$c > 4a$$

$$n \geq \frac{c/2}{c/4 - a} = \frac{2c}{c - 4a} = n_0.$$

Επομένως, εάν υποθέσουμε ότι $T(n) = O(1)$ για $n < n_0$, έχουμε

$$E[T(n)] = O(n).$$

Επιλογή (Selection)

Το πρόβλημα της επιλογής

Δίνεται μια ακολουθία S από n στοιχεία και ένας ακέραιος k, όπου $1 \leq k \leq n$. Να προσδιοριστεί το k μικρότερο στοιχείο στην S.

Ορισμός: Τα στοιχεία ενός συνόλου A ικανοποιούν μια γραμμική διάταξη $<$ τότε και μόνον τότε αν

1. για $a, b \in A, \quad a < b, \quad a = b \quad$ ή $\quad b < a$ και
2. για $a, b, c \in A,$ αν $a < b$ και $b < c$, τότε $a < c$

Βαθμός: Για μια ακολουθία $S = \{s_1, s_2, \ldots, s_n\}$ της οποίας τα στοιχεία είναι και στοιχεία ενός συνόλου με γραμμική διάταξη, ο <u>βαθμός</u> ενός στοιχείου s_i του S είναι το πλήθος των στοιχείων του S που προηγούνται του s_i συν 1.

Επιλογή (Selection)

Επιλογή

Δίνεται μια ακολουθία $S = \{s_1, s_2, \ldots, s_n\}$, της οποίας τα στοιχεία ικανοποιούν τη γραμμική διάταξη και ένας ακέραιος k, $1 \leq k \leq n$. Ζητείται ο προσδιορισμός του στοιχείου της S με βαθμό k. Το στοιχείο με βαθμό k θα συμβολίζεται με $s_{(k)}$.

Πολυπλοκότητα (κάτω όριο)

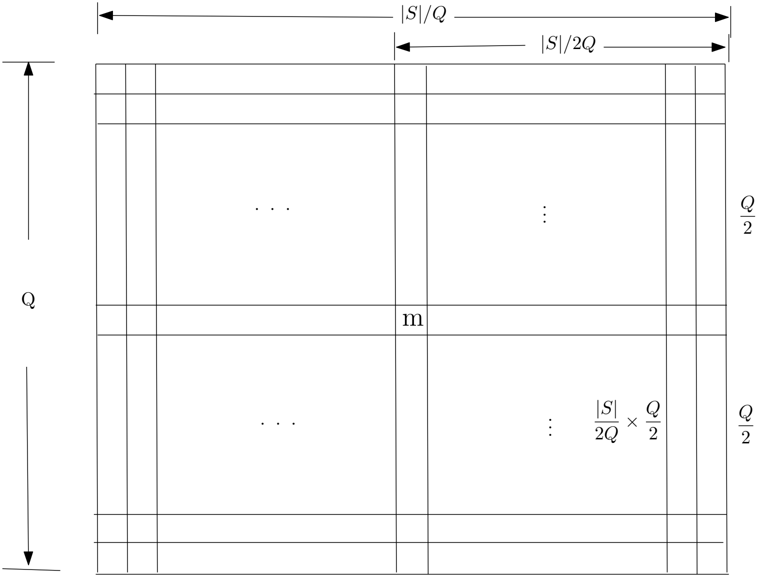
- Αν τα στοιχεία του S είναι ταξινομημένα, δηλ. $S = \{s_{(1)}, s_{(2)}, \ldots, s_{(n)}\}$, τότε το $s_{(k)}$ βρίσκεται σε ένα βήμα.
- Αν $k = 1$ ή $k = n$ και η S δεν είναι ταξινομημένη, τότε εξετάζοντας όλα τα στοιχεία της S διατηρώντας κάθε φορά το μικρότερο ($k = 1$) (ή το μεγαλύτερο $k = n$) βρίσκεται το ζητούμενο στοιχείο.
- Ο ανωτέρω αλγόριθμος της αναζήτησης δεν μπορεί να εφαρμοστεί αν $1 < k < n$.
- Άρα $T(n) = \Omega(n) \quad (\text{κάτω όριο})$

Επιλογή (Selection)

Χαρακτηριστικά Αλγορίθμου

- διαίρει και βασίλευε
- αναδρομικός αλγόριθμος
- απορρίπτει ένα πλήθος στοιχείων σε κάθε βήμα

Επιλογή (Selection)



Επιλογή (Selection)

ΕΠΙΛΟΓΗ (S,k)

Βήμα 1: **αν** $|S| \le Q$ **τότε** ταξινόμησε την S και επέστρεψε το $k^{\sigma\tau o}$ στοιχείο
αλλιώς υποδιαίρεσε την S σε $|S|/Q$ υπακολουθίες με Q στοιχεία
η καθεμία (μέχρι $Q-1$ εναπομείναντα στοιχεία)
τέλος αν

Βήμα 2: Ταξινόμησε κάθε υπακολουθία και υπολόγισε το μέσο της.

Βήμα 3: Κάλεσε την ΕΠΙΛΟΓΗ για τον υπολογισμό του m, τον μέσο των $|S|/Q$ μέσων που βρέθηκαν στο βήμα 2.

Βήμα 4: Δημιούργησε 3 υπακολουθίες $S_1, S_2,$ και S_3 με στοιχεία της S μικρότερα, ίσα και μεγαλύτερα από το m αντίστοιχα.

Βήμα 5: **αν** $|S_1| \ge k$ **τότε** {το $k^{\sigma\tau o}$ στοιχείο της S πρέπει να είναι στο S_1}
κάλεσε την ΕΠΙΛΟΓΗ για να βρεις το $k^{\sigma\tau o}$ στοιχείο του S_1
αλλιώς αν $|S_1| + |S_2| \ge k$ **τότε** επέστρεψε m {διότι το k στοιχείο $\in S_2$}
αλλιώς κάλεσε την ΕΠΙΛΟΓΗ για να βρεις το
$(k - |S_1| - |S_2|)^{\sigma\tau o}$ στοιχείο του S_3 {διότι το k στοιχείο $\in S_3$}.
τέλος αν
τέλος αν

Επιλογή (Selection)

Ανάλυση Πολυπλοκότητας

Βήμα 1: ταξινόμηση της S όταν $|S| < Q$ απαιτεί σταθερό χρόνο $O(1)$. Διαφορετικά, υποδιαίρεση της S απαιτεί $c_1 n$ χρόνο.

Βήμα 2: Η ταξινόμηση κάθε $|S|/Q$ υπακολουθίας (έχει Q στοιχεία) απαιτεί σταθερό χρόνο. Για όλες $c_2 n$ χρόνο.

Βήμα 3: $t(n/Q)$, $t(n)$: χρόνος εκτέλεσης της ΕΠΙΛΟΓΗ

Βήμα 4: $c_3 n$

Βήμα 5: $(|S|/2Q)x(Q/2) = |S|/4$ στοιχεία της S είναι ίσα ή μεγαλύτερα με m. Συνεπώς, $|S_1| \leq 3|S|/4$. Όμοια $|S_3| \leq 3|S|/4$. Συνεπώς μια αναδρομική κλήση της ΕΠΙΛΟΓΗ απαιτεί $t(3n/4)$ χρόνο.

Συνολικά:

$$t(n) = c_4 n + t(n/Q) + t(3n/4)$$

όπου $c_4 = c_1 + c_2 + c_3$.

Επιλογή (Selection)

Προσδιορισμός του Q

Αν το Q επιλεγεί τέτοιο ώστε

$$n/Q + 3n/4 < n$$

τότε οι δυο αναδρομικές κλήσεις στον αλγόριθμο εφαρμόζονται σε φθίνουσες ακολουθίες. Οποιαδήποτε τιμή του $Q \geq 5$ ικανοποιεί την ανωτέρω ανισότητα.
Για $Q = 5$ έχουμε

$$t(n) = c_4 n + t(n/5) + t(3n/4)$$

Υποθέτοντας $t(n) \leq c_5 n$ έχουμε $t(n) \leq c_4 n + c_5(19n/20)$
και για $c_5 = 20c_4$

$$t(n) \leq c_5(n/20) + c_5(19n/20) = c_5 n$$

$$t(n) = O(n) \quad \text{βέλτιστος αλγόριθμος}$$

Επιλογή (Selection)

EPILOGH

Δίνεται η ακολουθία

S={1,14,3,4,2,6,7,4,5,8,2,1,5,12,4,3,2,8,9,10,5,2,12}

και ένας ακεραίος k, $1 < k < 23$. Ζητείται να προσδιοριστεί το στοιχείο της ακολουθίας με βαθμό 19, δηλ. $k = 19$.

Λύση: Επιλέγουμε το Q ώστε να πληρείται η σχέση $\frac{n}{Q} + \frac{3n}{4} < n \Leftrightarrow Q > 4$. Έστω $Q = 5$.

Βήμα 1: Επειδή $|S| = 23 > 5 = Q$ θα υποδιαιρέσουμε την S σε $|S|/Q$ υποακολουθίες, δηλαδή σε 4 υποακολουθίες με 5 στοιχεία και σε 1 υποακολουθία με 3 στοιχεία.

1	14	3	4	2

6	7	4	5	8

2	1	5	12	4

3	2	8	9	10

5	2	12

Επιλογή (Selection)

EPILOGH

Βήμα 2: α) Ταξινόμηση κάθε υποακολουθίας
β) Βρίσκουμε το μέσο m_i κάθε υποακολουθία

1	2	3	4	14

$m_1 = 3$

4	5	6	7	8

$m_2 = 6$

1	2	4	5	12

$m_3 = 4$

2	3	8	9	10

$m_4 = 8$

2	5	12

$m_5 = 5$

Βήμα 3: Κλήση της ΕΠΙΛΟΓΗ για τον υπολογισμό του μέσου των μέσων

B_1 :

3	6	4	8	5
m_1	m_2	m_3	m_4	m_5

$\rightarrow$ B_{2b} : $\boxed{m = 5}$
Υπολογισμός μέσου

Βήμα 4: Δημιουργία 3 ανακολουθιών S_1, S_2, S_3 με στοιχεία $<, =, >$ 5 αντίστοιχα

1 3 4 2 4 2 1 4 3 2 2	5 5 5	14 6 7 8 12 8 9 10 12
$\longleftarrow S_1 \longrightarrow$	$\leftarrow S_2 \rightarrow$	$\longleftarrow S_3 \longrightarrow$

Επιλογή (Selection)

ΕΠΙΛΟΓΗ

Παρατηρούμε ότι $|S_1| = 11, |S_2| = 3, |S_3| = 9$

Βήμα 5: Παρατηρούμε ότι $k = 19 > |S_1| + |S_2| = 14$ όποτε καλούμε την ΕΠΙΛΟΓΗ στην S_3 για να βρούμε το $K_{new}^{(1)} = k - |S_1| - |S_2| \Rightarrow K_{new}^{(1)} = 19 - 14 \Rightarrow K_{new}^{(1)} = 5$, δηλ το 5^o στοιχείο της ακολουθίας S_3

Β1: Για $Q = 5$ είναι $|S_3| = 9 > 5 = Q$ όποτε $|S_3|/Q = \frac{9}{5} = 1.8$ δηλ. θα υποδιαιρέσουμε την S_3 σε 1 υποακολουθία με 5 στοιχεία και σε 1 υποακολουθία με 4 στοιχεία

14	6	7	8	12

8	9	10	12

Β2: α) Ταξινόμηση κάθε υποακολουθίας β) Εύρεση μέσου της

6	7	8	12	14

$m'_1 = 8$

8	9	10	12

$m'_2 = 9$

Β3: Το μέσο των 8, 9 είναι το $\boxed{m' = 8}$

Επιλογή (Selection)

ΕΠΙΛΟΓΗ

B4: Δημιουργία 3 υποακολουθιών S_1', S_2', S_3' με στοιχεία $<, =, >$ 8 αντίστοιχα

$$\left| \begin{matrix} 6 & 7 \\ \leftarrow S_1' & \rightarrow \end{matrix} \right| \begin{matrix} 8 & 8 \\ \leftarrow S_2' & \rightarrow \end{matrix} \left| \begin{matrix} 14 & 12 & 9 & 10 & 12 \\ \longleftarrow & & S_3' & & \longrightarrow \end{matrix} \right|$$

B5: Παρατηρούμε ότι $|S_1'| = 2, |S_2'| = 2, |S_3'| = 5$
Παρατηρούμε ότι $K_{new}^{(1)} = 5 > |S_1'| + |S_2'| = 4$ οπότε καλούμε την ΕΠΙΛΟΓΗ στην S_3' για να βρούμε το $K_{new}^{(2)} = K_{new}^{(1)} - |S_1'| - |S_2'| \Rightarrow$ $K_{new}^{(2)} = 5 - 4 \Rightarrow K_{new}^{(2)} = 1$ δηλ. το 1^o στοιχείο της ακολουθίας S_3'

B1΄: Είναι $|S_3'| = 5 = Q$ όποτε ταξινομούμε την S_3' και επιστρέφουμε το 1^o στοιχείο. Είναι

14	12	9	10	12

S_3'

9	10	12	12	14

S_3' ταξινομημένη .

Το 1^o στοιχείο της S_3' μετά την ταξινόμηση της είναι το 9 και κατά συνέπεια αυτό είναι το ζητούμενο στοιχείο