ΚΕΦΑΛΑΙΟ 7

ΔΕΝΤΡΟ ΠΕΡΙΟΧΗΣ

Σ. ΣΙΟΥΤΑΣ, ΚΑΘΗΓΗΤΗΣ, ΤΜΗΥΠ, ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ, ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

ΠΟΛΥΔΙΑΣΤΑΤΕΣ ΔΟΜΕΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΓΕΩΜΕΤΡΙΑ

Acknowledgement:
Taken slides from CEID@Upatras && AUTH

Εισαγωγή

- Δέντρο Περιοχής
- Λογαριθμικός χρόνος απάντησης σε μη γραμμικό χώρο
- Επεκτάσιμο για προβλήματα d-διάστατου χώρου
  - Πολυ-λογαριθμικός χρόνος απάντησης σε μη-γραμμικό χώρο

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Ορισμός

- Έστω σύνολο S απο σημεία του d-διάστατου χώρου :
  $S = U_0 x\ U_1\ x\ U_2 \ldots x\ U_{d-1}$
- Άρα τα στοιχεία του S είναι της μορφής $(x_0 x_1\ x_2 \ldots x_{d-1})$ όπου $x_i \in U_i$
- Έστω P υποσύνολο d συντεταγμένων , τότε $p(S,P) = \{(x_{i1}\ , x_{i2} \ldots , x_{ik})\}$ είναι η προβολή του S στις συντεταγμένες του P

Παράδειγμα

- Έστω σύνολο S αποτελείται απο σημεία στο επίπεδο
- Τα στοιχεία του S είναι της μορφής (x_0 ,x_1) όπου x_0 είναι η συντεταγμένη ως προς την πρώτη διάσταση και x_1 είναι η συντεταγμένη ως προς την δεύτερη διάσταση
- Αν P={0} τότε p(S,P)θα ειναι ενα σύνολο απο στοιχεία της μορφής (x_0) όπου x_0 είναι η τιμή της πρώτης συντεταγμένης των στοιχείων του S

Έστω S = {(1,3),(3,4),(1,2),(1,4),(5,6)} και P={0}

Τότε p(S,P) = {(1),(3),(5)} δηλαδή όλες οι διακριτές x συντεταγμένες

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Ορισμός

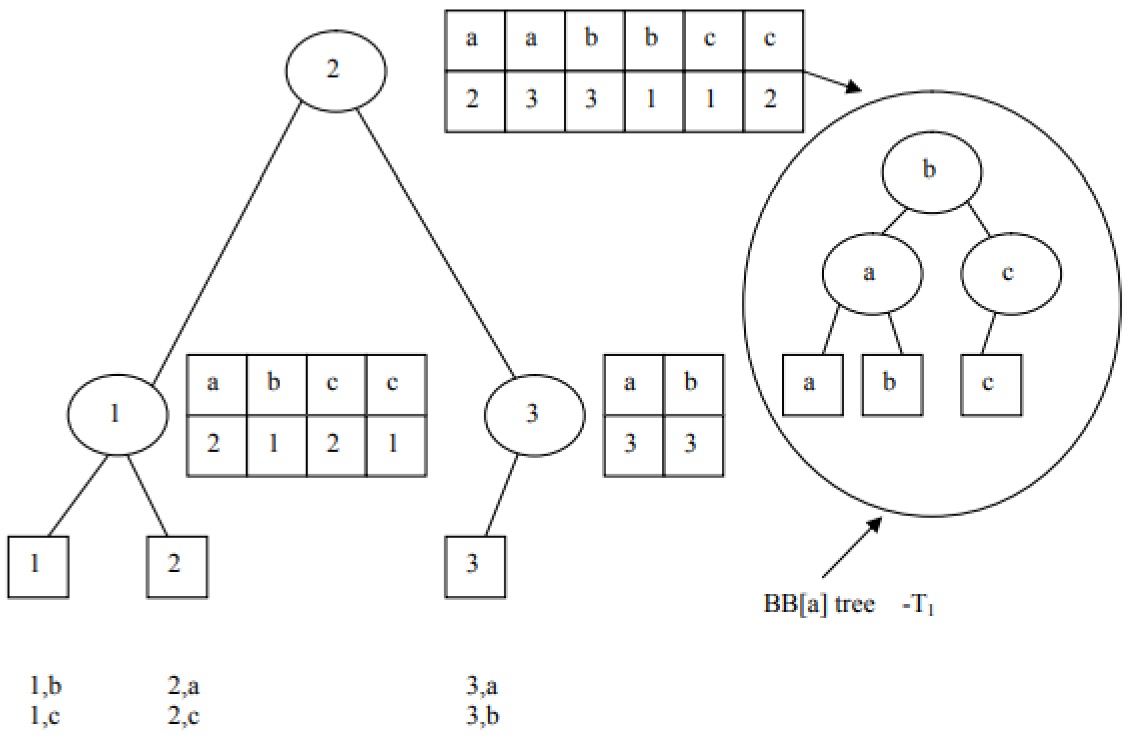
- Έστω m∈N και έστω α∈(¼ , √2/2)
- ‘Οπου m είναι μια **ρυθμιστική παράμετρος** και το α είναι μια **παράμετρος ζύγισης βάρους**
- Ένα d-διάστατο δέντρο περιοχής T για ένα πολυδιάστατο σύνολο S = U_0 x U_1 x U_2 ... x U_{d-1} ορίζεται ως εξής :
  - Εαν d=1 τότε τοT είναι ενα BB[α] δέντρο για το S
  - Εαν d>1 τότε το T αποτελειται απο ενα BB[α] πρωταρχικό δέντρο T_0 για το p_0(s)
  - Για κάθε κόμβο v του T_0 με βάθος depth(v) υπάρχει ενα βοηθητικό (d-1)-διάστατο δέντρο T_a(v)
  - Εδώ το S(v) είναι το συνολο όλων των ($x_0 x_1$ x_2 ... x_{d-1}) ∈ S ώστε το φύλλο x_0 να είναι απόγονος του v στο T_0

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Παράδειγμα

Έστω S = {(1,b),(1,c),(2,a),(2,c),(3,a),(3,b)} , m=1

Το Δέντρο Περιοχής για το παραπάνω σύνολο στοιχείων είναι το παρκάτω σχήμα :



Σχήμα 7-1

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Το $m \neq 1$ σημαίνει οτι τα βοηθητικά δέντρα ορίζονται για κάθε m κόμβους πάνω στα μονοπάτια. Στην περίπτωση αυτή σε έναν κόμβο v αντιστοιχεί ένα βοηθητικό δέντρο μόνο αν το depth(v) είναι ακεραιο πολλαπλάσιο του m.
2. Το σύνολο S του παραδείγματος είναι στις 2 διαστάσεις. Αν ο αριθμός των διαστάσεων ήταν μεγαλυτερος, τοτε σε καθε κομβο του βοηθητικού δέντρου T_1 αντιστοιχεί ενα άλλο δέντρο που αφορά την τρίτη διάσταση. Η εισαγωγή των στοιχείων στο δέντρο περιοχής γίνεται ως εξής :

   Αρχικά εισάγω κάθε στοιχείο ως προς την πρωτη συντεταγμένη στο BB[α] δέντρο T_0 που αφορα την συντεταγμένη αυτή. Σε κάθε κόμβο v του μονοπατιού στο οποίο έγινε η ένθεση, αντιστοιχεί ένα βοηθητικό δέντρο BB[α]. Το δέντρο αυτό περιέχει τα στοιχεία του S των οποίων η πρώτη συντεταγμένη είναι φύλλο του υποδέντρου του T_0 με ρίζα τον v. Σε κάθε ένα απο αυτά τα δέντρα εισάγω το στοιχείο ως προς τη δεύτερη συντεταγμένη.

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Χώρος του δέντρου περιοχής

Λήμμα

- Έστω $S_m(d,n)$ ο χώρος του d-διάστατου δέντρου περιοχής με ρυθμιστική παράμετρο m για ένα σύνολο απο n στοιχεία

$$S(d,n) = O\left(n\left(\frac{c\log n}{m}\right)^{d-1}\right) \text{, όπου } c = \left(\log\frac{1}{1-a}\right)^{-1}$$

Απόδειξη

- Το βάθος του BB[α] δέντρου με n φύλλα είναι το πολυ clogn όπου $c = -\log^{-1}(1-a)$
- Κάθε στοιχείο x∈S είναι αποθηκευμένο στο πρωταρχικό δέντρο για την πρώτη διάσταση σε (clogn)/m το πολύ βοηθητικά δέντρα για τη δεύτερη διάσταση, σε $((c\log m)/m)^2$ το πολυ βοηθητικα δεντρα για την τριτη διασταση κ.ο.κ.
- Επομένως ο συνολικός χώρος είναι :

$$O\left(n\sum_{0=<i=<d-1}\left(\frac{c\log n}{m}\right)^{i}\right) = O\left(n\left(\frac{c\log n}{m}\right)^{d-1}\right)$$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Χρόνος του δέντρου περιοχής

<u>Λήμμα</u>

- Ιδανικά d-διάστατα δέντρα περιοχής καλούνται εκείνα για τα οποία για όλους τους κόμβους v με γιους x ισχύει :
  |S(x)| =< ΓS(v)/2¬ , όπου Γz¬ δηλώνει τον μικρότερο ακέραιο ο οποίος είναι μεγαλύτερος απο τον πραγματικό αριθμό z
- Όπου m είναι η ρυθμιστική παράμετρος, τα δέντρα αυτά μπορούν να κατασκευαστούν σε χρόνο :

$$O(dn\log n + n(\frac{\log n}{m})^{d-1})$$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Χρόνος του δέντρου περιοχής

Απόδειξη

- Αρχικά διατάσσουμε το S για κάθε συντεταγμένη χωριστά. Αυτό κοστίζει χρόνο O(dnlogn)
- Έστω $T_m(d,n)$ ο χρόνος που χρειάζεται για να κατασκευάσουμε ένα ιδανικό d-διάστατο δέντρο περιοχής για ένα σύνολο S με n στοιχεία που είναι προδιατεταγμένα για κάθε συντεταγμένη. Με επαγωγή αποδεικνύουμε οτι :

$$T_m(d,n) = O\left(n\left(\frac{\log n}{m}\right)^{d-1} \right)$$

- Για d=1 : χρειαζόμαστε χρόνο μόνο O(n) για ένα BB[a] δέντρο
- Για d>1 :
  - Κατασκευάζουμε το πρωταρχικό δέντρο σε χρόνο O(n) και πρέπει να κατασκευάσουμε τα βοηθητικά δέντρα για τη δεύτερη διάσταση, τα οποία έστω οτι έχουν μέγεθος $n_1, \ldots, n_t$. Έχουμε όμως $n_1 + \ldots, + n_t$ =< (nlogn)/m διότι κάθε στοιχείο αποθηκεύεται το πολύ σε (logn)/m βοηθητικά δέντρα. Συνεπώς :

$$T_m(d,n) = O(n) + \sum_i T_m(d\text{-}1,n_i) = O(n) + O\left(\sum_i n_i \left(\frac{\log n}{m}\right)^{d-2}\right) = O\left(n\left(\frac{\log n}{m}\right)^{d-1} \right)$$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Για m=1 ο χρόνος κατασκευής είναι $O(n(\log n)^{d-1})$
2. Για m=εlogn ο χρόνος κατασκευής είναι O(dnlogn)

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Χρόνος απάντησης του δέντρου περιοχής

<u>Λήμμα</u>

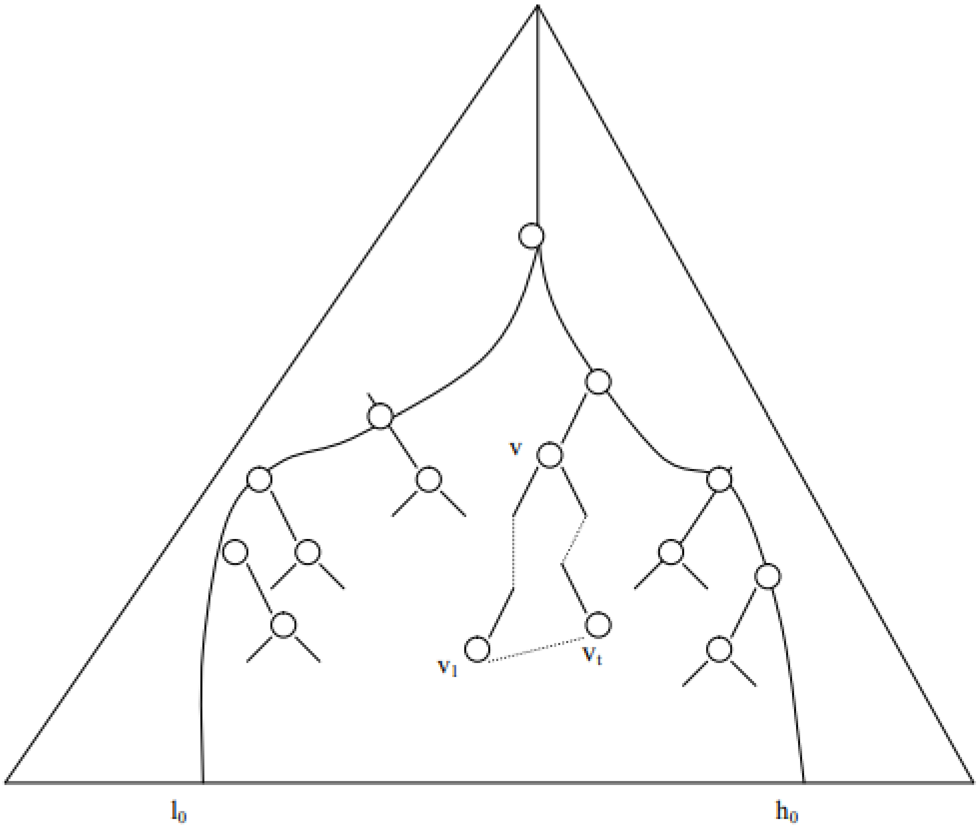
- Έστω $Q_m(d,n)$ ο χρόνος που απαιτείται για να απαντηθεί μια ερώτηση περιοχής σε ενα d-διάστατο δέντρο περιοχής με n στοιχεία.

$$Q\,(d,n) = O\left(\log n\left(c\left(\frac{2^m}{m}\right)\log n\right)^{d-1} + |A|\right)$$

<u>Απόδειξη</u>

- Για d=1 : είναι προφανές
- Για d>1 έχουμε:
  - ‘Εστω R = $[l_0,h_0] \times ... \times [l_{d-1},h_{d-1}]$ μια ορθοκανονική περιοχή ερώτησης
  - Καταρχάς ψάχνουμε για τα l_0,h_0 στο πρωταρχικό δεντρο T_0
  - Έτσι ορίζονται δύο μονοπάτια στο T_0 μήκους clogn όπως φαίνεται στο σχήμα 7-2

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ



Σχήμα 7-2

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Χρόνος απάντησης του δέντρου περιοχής

<u>Απόδειξη</u>

- Έστω M το σύνολο των κόμβων που βρίσκονται ανάμεσα στα δύο μονοπάτια και οι γονείς τους ανήκουν στα μονοπάτια αυτά. Τότε |M|=<2clogn
- Κάθε κομβος απο το M αναπαριστά ενα υποσύνολο του S του οποίου οι 0-συντεταγμένες κείνται στο διάστημα $[l_0,h_0]$
- Έστω v∈M και έστω $v_1,..,v_t$ οι επόμενοι απόγονοι του v, που έχουν βάθος ακέραιο πολλαπλάσιο του m
- Τότε t<=2 και όλα τα v έχουν βοηθητικά δέντρα
- Σε κάθε ένα απο αυτά έχουμε να λύσουμε (d-1)-διάστατα προβλήματα

Επειδή το πλήθος των v είναι το πολύ $2c((\log n)/m)2^{m-1}$:

$$Q_m(d,n) =< c\left(\frac{2^{m-1}}{m}\right)(\log n)Q_m(d-1,n)+|A|$$

$$Q_m(d,n) = O\left(\log n\left(c\left(\frac{2^{m}}{m}\right)\log n\right)^{d-1}+|A|\right)$$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

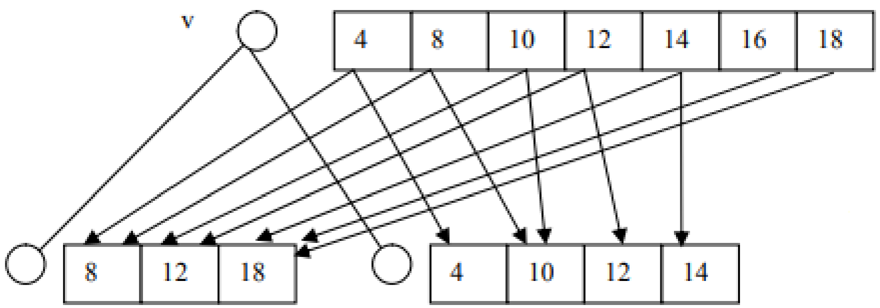
ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Για m=1 : $Q(d,n) = O((\log n)(2c\log n)^{d-1} + |A|)$
   Για d=2 : $Q_1(2,n) = O((\log n)^2 + |A|)$
2. Για m = εlogn : $Q_n(d,n) = O((\log n)(c/\varepsilon)^{d-1} n^{\varepsilon(d-1)} + |A|)$
3. Για m=1 είναι δυνατο ο παραπανω χρονος να μειωθει σε $O((\log n)^{d-1} + |A|)$ εφαρμοζοντας τη μεθοδο του fractional cascading

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

<u>Εξήγηση του fractional cascading για d=2 , m=1</u>

- Η βασικη ιδέα της μεθοδου του fractional cascading ειναι η παρακατω :
  - Αντί να ψαχνουμε τις θεσεις l_1 χωριστά για κάθε βοηθητικο δέντρο των κόμβων v∈M μπορούμε να χτίσουμε ένα πλέγμα με γέφυρες ωστε η θέση του l_1 να μπορεί να υπολογιστει μόνο μια φορά στη ρίζα και μετα ακολουθώντας τις γεφυρες κατα μήκος του μονοπατιού l_0 να βρισκομαστε παντα στη σωστη θεση
  - Καθε στοιχειο του βοηθητικου δέντρου δείχνει στο ίδιο ή άμεσο μεγαλυτερο στοιχείο των βοηθητικων δέντρων των δύο γιων του v



Σχήμα 7-3

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

<u>Δυναμοποίηση των δέντρων περιοχής</u>

- Έστω x = ($x_0 x_1$ x_2 ... x_{d-1}) ένα σημειο που θα εντεθεί ή θα αποσβεσθεί
- Κατ’αρχήν ψάχνουμε το x_0 στο πρωταρχικό δέντρο και ενθέτουμε το x
- Αυτό στοιχίζει O(logn) στο πρωταρχικό δέντρο και επιπλέον πρέπει να ενθέσουμε το x το πολυ σε (clogn)/m βοηθητικά δέντρα για την δευτερη διάσταση κ.ο.κ.
- Επομένως τα γενικα εξοδα για ένθεση χωρίς τα έξοδα επαναζυγισης είναι :

$$O((\log n)((c\log n)/m)^{d-1})$$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Επαναζυγίσεις

Για καθε δέντρο πρωταρχικό ή βοηθητικό ή βοηθητικό-βοηθητικό στο οποίο εντείνεται το x υπολογίζουμε εναν κόμβο v με ελάχιστο βαθος ο οποιος περιπιπτει εκτος ζυγισης. Τοτε αντικαθιστούμε το υποδέντρο με ριζα v με εναν ιδανικο d’-διαστατο δέντρο για το σύνολο S(v) των απογόνων του v. Εδω ισχύει :

- d’ = d εαν ο v ειναι κόμβος του πρωταρχικού δεντρου
- d’ = d-1 εαν ο v είναι κομβος του βοηθητικου δεντρου
- d’ = d-2 εαν ο v είναι κόμβος του βοηθητικου-βοηθητικου δεντρου κλπ

Αυτή η επαναζύγιση στοιχίζει : $O(d'q\log q + q\frac{(\log q)^{d'-1}}{m})$

Χειρότερη Περίπτωση

Μια ενθεση (απόσβεση) στη χειρότερη περίπτωση στοιχ $O(dn\log n + n\frac{(\log n)^{d-1}}{m})$

Η κατανεμημένη περίπτωση είναι πολύ καλύτερη.

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Λήμμα

Τα κατανεμημένα έξοδα για μια ενθεση (αποσβεση) σε ένα δ-διαστατο δέντρο περιοχης με ρυθμιστική παράμετρο m είναι :

$$O\left((m^2 + md)\left(\frac{c\log n}{m}\right)^d\right)$$

Θεωρημα

Το d-διαστατο δέντρο περιοχης με ρυθμιστικη παράμετρο m>=1 και ζυγοστάθμιση α∈(¼,1-√2/2) για ένα συνολο με n στοιχεία χρειάζεται **χώρο** :

$$O\left(n\left(\frac{c\log n}{m}\right)^{d-1}\right)$$

Υποστηριζει **ερωτήσεις ορθογώνιας περιοχης** σε χρόνο : $O\left(\log n\left(c\left(\frac{2^m}{m}\right)\log n\right) + |A|\right)$

Και εχει **κατανεμημένα δυναμικά εξοδα** : $O\left((m^2 + md)\left(\frac{c\log n}{m}\right)^d\right)$

Εδώ είναι : $c = \left(\log \frac{1}{1-a}\right)^{-1}$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

Ρυθμιστική Παράμετρος	Χώρος	Χρόνος	Κατανεμημένος Χρονος Ένθεσης/Απόσβεσης
1	$n(c\log n)^{d-1}$	$(\log n)(2c\log n)^{d-1}$	$D(c\log n)^{d}$
$\varepsilon\log n$	$n(c/\varepsilon)^{d-1}$	$(c/\varepsilon)^{d-1} n^{\varepsilon d}\log n$	$(c/\varepsilon)^{d}((\varepsilon\log n)^{2} + d\varepsilon\log n)$

ΑΝΑΛΥΣΗ ΤΟΥ ΔΕΝΤΡΟΥ ΠΕΡΙΟΧΗΣ

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Όπως αναφέραμε παραπανω με την τεχνικη του fractional cascading μπορουμε να βελτιώσουμε το χρόνο στο στατικό πρόβλημα σε $O((\log n)^{d-1} + |A|)$
2. Εφαρμόζοντας την τεχνικη του fractional cascading των Melhorn,Naher για το επίπεδο έχουμε πολυπλοκότητα χώρου O(nlogn), κατανεμημενη πολυπλοκότητα χρόνου ένθεσης, απόσβεσης O((logn)loglogn) και πολυπλοκότητα χρόνου για ερώτηση O(logn)loglogn + |A|) . Αν έχουμε μονο ένθεση ή μονο αποσβεση τοτε απο τις παραπανω πολυπλοκοτητες δεν έχουμε τον παραγοντα loglogn όπως έχουν δειξει οι Imai και Asano.
3. Επιπροσθέτως παρατίθεται μια τεχνική για το πως η κατανεμημενη πολυπλκότητα μετατρέπεται σε χειρότερο χρόνο $O(\log^d n)$ και ο χρονος απόκρισης γινεται $O(\log^d n+|A|)$