A study of the Golomb Ruler and Sidon Set Problems, and Determination of Large, Near-Optimal Golomb Rulers

Απόστολος Δημητρομανωλάκης

Ιούνιος 2002, Χανιά

Οργάνωση της παρουσίασης

1. Εισαγωγή στους κανόνες Golomb
2. Εισαγωγή στα σύνολα Sidon
3. Ισοδυναμία των δυο προβλημάτων
4. Κατασκευές για κανόνες Golomb
5. Αλγόριθμοι για σχεδόν βέλτιστους κανόνες
6. Αναζήτηση και αποτελέσματα

Κανόνες Golomb

Κανόνες Golomb

- Κανόνας* Golomb : ένα σύνολο (θετικών) ακεραίων αριθμών $a_1, a_2, \ldots, a_n$ τέτοιο ώστε οι θετικές διαφορές $a_i - a_j$ να είναι όλες διαφορετικές

- Ο συγκεκριμένος κανόνας μετράει τις αποστάσεις 1,2,3,4,5,7,8,9,10,11
- Βελτιστός κανόνας Golomb : Ένας κανόνας με n σημεία και το ελάχιστο δυνατό μήκος.
- Η $G(n)$ ορίζεται σαν το ελάχιστο μήκος ενός κανόνα με n σημεία. Δεν είναι γνωστή λύση κλειστού ή ανοιχτού τύπου για την $G(n)$, μόνο φράγματα.

*χάρακας

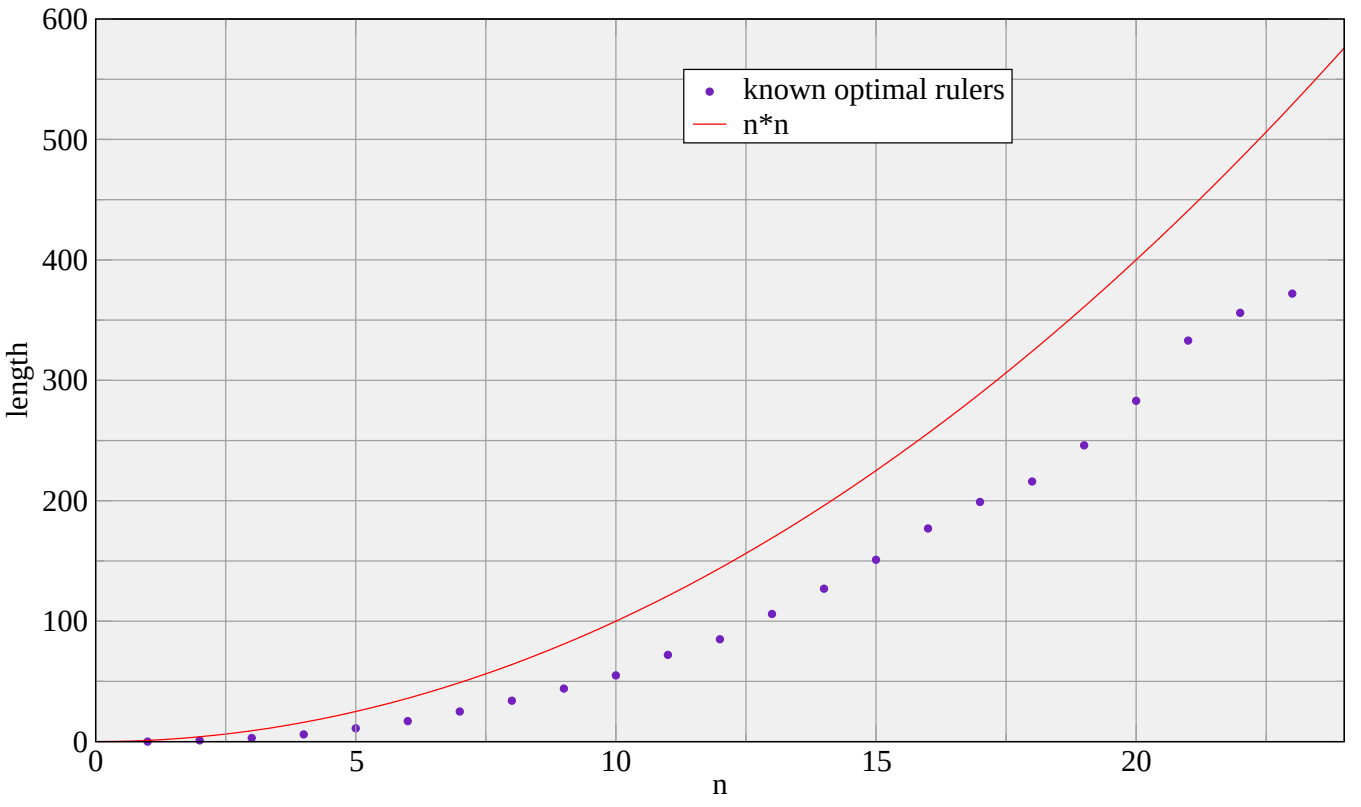
Χρήσεις των κανόνων Golomb

- Ανάθεση ραδιοσυχνοτήτων χωρίς παρεμβολές τρίτης τάξης (Babock 1953)
- Παραγωγή κωδίκων C.S.O.C. (convolutional self-orthogonal codes) (Robinson 1967)
- Γραμμική διάταξη τηλεσκοπίων στην ραδιοαστρονομία για μεγιστοποίηση των ωφέλιμων παρατηρήσεων (Blum 1974)
- Τοποθέτηση αισθητήρων στην κρυσταλλογραφία κ.α.

Γνωστοί βέλτιστοι κανόνες Golomb

- Το υπολογιστικός κόστος για να βρεθεί ένας βέλτιστος κανόνας Golomb αυξάνεται εκθετικά με το αριθμό των σημείων.
- Ως σήμερα είναι γνωστοί οι βέλτιστοι κανόνες μέχρι 23 σημεία
- Για να βρεθεί ο κανόνας 23 χρησιμοποιηθήκαν 25000 υπολογιστές για ένα διάστημα αρκετών μηνών.
- Όταν το πλήθος των σημείων ενός κανόνα είναι μεγαλύτερο απο 24 σημεία, βέλτιστοι κανόνες δεν είναι δυνατόν να βρεθούν σε λογικό χρονικό διάστημα.
- Σχεδόν βέλτιστοι κανόνες (near-optimal): κανόνες με μήκος κοντά στο βέλτιστο (ή ισως και βέλτιστο).

Μήκος γνωστών βέλτιστων κανόνων



Σύνολα Sidon

Τι είναι τα σύνολα Sidon

Ορισμός:

Ένα σύνολο Sidon είναι ενα υποσύνολο $a_1, a_2, \ldots, a_n$ του $\{1, 2, \ldots, n\}$ τετοιο ώστε τα αθροίσματα $a_i + a_j$ να είναι όλα διαφορετικά μεταξύ τους.

- $F_2(d)$: το μέγιστο πλήθος ακεραίων που μπορούν να επιλεγούν απο το $\{1,2,\ldots,d\}$ και να αποτελούν ένα σύνολο Sidon .
- Δεν είναι υπάρχει κλειστός τύπος για την συνάρτηση $F_2(d)$, μόνο ασυμπτωτικά αποτελέσματα.

Γνωστά όρια για την F_2

- Πάνω φράγματα
  - Βασικό κάτω φράγμα: $F_2(d) \leqslant \sqrt{2}\, d^{1/2}$.
  - Erdős 1941 : $F_2(d) < d^{1/2} + O(d^{1/4})$
  - Lindstrom 1969 : $F_2(d) < d^{1/2} + d^{1/4} + 1$
- Κάτω φράγματα
  - πιο δύσκολο να αποδειχθούν
  - $F_2(d) > d^{1/3}$ απο τις κατασκευές που θα δούμε αργότερα
  - Εκτος απο αυτό μόνο ασυμπτωτικά αποτελέσματα: $F_2(d) > d^{1/2} - O(d^{5/16})$ (Erdős 1944)

Ισοδυναμία των δυο προβλημάτων

Τι γινόταν ως σήμερα

- Τα σύνολα Sidon και οι κανόνες Golomb ήταν γνωστο οτι είναι ισοδύναμα προβλήματα.
- Ο διαφορετικός φορμαλισμός δεν επέτρεπε την ευκόλη επαναδιατύπωση αποτελεσμάτων μεταξυ των δυο προβλημάτων.
- Πολλοί ερευνητές και απο τις δυο πλευρές αγνοούσαν την ισοδυναμία αυτη και ξανααποδείκνυαν αποτελεσματα.
- Για παράδειγμα το οτι ασυμπτωτικά οι κανόνες Golomb εχούν μήκος n^2 (Erdős 1944) αποδείχθηκε εκ νέου το 1967 (Atkinson et al)

Απόδειξη της ισοδυναμίας

- Απο την σχέση $a_i + a_j = a_k + a_l \iff a_i - a_k = a_l - a_j$ είναι φανερο οτι ένα σύνολο Sidon είναι και κανόνας Golomb και αντιστρόφως.
- Κανόνες Golomb :
  - περιέχουν το 0 σαν στοιχείο
  - η συνάρτηση $G(n)$ ορίζεται σαν το ελάχιστο μήκος κανόνα με n σημεία
- Σύνολα Sidon :
  - το ελάχιστο στοιχείο είναι το 1
  - η συνάρτηση $F_2(n)$ ορίζεται σαν το μέγιστο πλήθος στοιχείων που μπορούν να επιλεγούν απο το $1,2,\ldots,n$.

Τι αποδείξαμε - Σχέσεις ισότητας

- Αν είναι γνωστή μια τιμη για την συνάρτηση F_2:

$$F_2(d) = n \iff \begin{array}{l} G(n) \leqslant d-1 \\ G(n+1) > d-1 \end{array}$$

- Αν είναι γνωστή μια τιμη για την $G(n)$:

$$G(n) = d \iff \begin{array}{l} F_2(d) = n-1 \\ F_2(d+1) = n \end{array}$$

Τι αποδείξαμε - Σχέσεις ανισότητας

- Υπάρχει μια αντίστροφη σχέση μεταξύ F_2 και G που μπορεί να εκφραστεί καλυτερα με το ακόλουθο θεώρημα.

  Για οποιεσδήποτε συναρτήσεις l και u, αν $l(d) < F_2(d) < u(d)$ τότε

  $$u^{-1}(n) < G(n) + 1 < l^{-1}(n)$$

- Το αντίστοιχο ισχύει και προς την άλλη κατεύθυνση: Για οποιεσδήποτε συναρτήσεις l και u, αν $l(n) < G(n) < u(n)$ τότε

  $$u^{-1}(d) \leqslant F_2(d) \leqslant l^{-1}(d)$$

Ένα καινούργιο όριο για την $G(n)$

Ο Lindstom (1969) απόδειξε οτι $F_2(d) < d^{1/2} + d^{1/4} + 1$

- Με τα θεωρήματα που αποδείξαμε μπορούμε να βρούμε ενα καλύτερο όριο για την $G(n)$.
- Η αντιστροφή συνάρτηση είναι $u^{-1}(n) = \left(\sqrt{n - \frac{3}{4}} - \frac{1}{2}\right)^4$
- Απο το θεώρημα που αποδείξαμε:

$$G(n) > u^{-1}(n) - 1 > n^2 - 2n\sqrt{n} + \sqrt{n} - 2$$

Κατασκευές για κανόνες Golomb

Κατασκευές

Κανόνες Golomb με πολυ περισσότερα απο 24 σημεία:

- Οι αλγόριθμοι εξαντλητικής αναζήτησης δεν μπορούν να βρουν καλούς χάρακες
- Υπάρχουν κατασκευές απο την θεωρία αριθμών που δίνουν καλούς χάρακες

Μια απλή κατασκευή

$$na^2 + a \quad , \quad a = \{0, 1, \ldots, n-1\}$$

- παράγει ένας χάρακα με n σημεία
- απλή και γρήγορη κατασκευή
- μέγιστο στοιχείο: $n^3 - 2n^2 + 2n \approx n^3$
- η ορθότητα αποδεικνύεται με απλά μαθηματικά

Modular κατασκευές

Οι επόμενες 3 κατασκευές έχουν καποια ιδιαιτερότητα

- modular : όλα τα ζεύγη αθροίσματων δίνουν διαφορετικα ύπολοιπα οτάν διαιρεθούν με κάποιον ακέραιο q
- διαφορετική φύση απο τους απλόυς χάρακες
- δυο παράμετροι: το πλήθος των στοιχείων n και το υπόλοιπο q (ανάλογο με το μήκος των απλών κανόνων)
- $n(n-1)$ αντι για $\frac{1}{2}n(n-1)$ διαφορετικές αποστάσεις μεταξύ των στοιχείων: $q \geqslant n(n-1)$

Η κατασκευή του Ruzsa

$$R(p,g) = pi + (p-1)g^i \mod p(p-1) \quad \text{για} \quad 1 \leqslant i \leqslant p-1$$

p : πρώτος αριθμός

g : primitive στοιχείο του Z_p^*

- $n = p-1$ στοιχεία modulo $p(p-1)$
- γρήγορη κατασκευή (3 πολλ. και 1 διαίρεση για κάθε στοιχείο)
- δίνει σχετικά καλούς κανόνες: μήκος $< n^2 + n$
- για παράδειγμα ($g = 3, p = 7$): $\{6, 10, 15, 23, 25, 26\} \mod 42$

Η κατασκευή του Bose

$$d_1, \ldots, d_q = \{a : 1 \leqslant a < q^2 \text{ και } \theta^a - \theta \in GF(q)\}$$

q : πρώτος ή δύναμη πρώτου p^n

θ : primitive στοιχείο του $GF(q^2)$

- $n = q$ στοιχεία modulo $q^2 - 1$
- σχετικά αργή: απαιτούνται πράξεις με πολυώνυμα 2ού βαθμού
- δίνει καλούς κανόνες: μήκος $< n^2 - 1$
- για παράδειγμα: $\{1, 2, 5, 11, 31, 36, 38\} \mod 48$

Η κατασκευή του Singer

$$d_0, d_1, \ldots, d_q \mod q^2 + q + 1$$

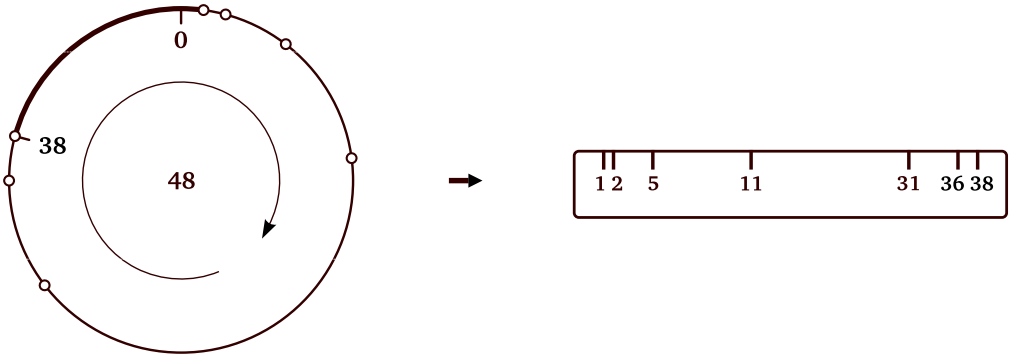
q : πρώτος ή δύναμη πρώτου p^n

- $n = q + 1$ στοιχεία modulo $q^2 + q + 1$
- απελπιστικά αργός υπολογισμός: απαιτούνται διαιρέσεις με πολυώνυμα 3ού βαθμού
- δίνει καλούς κανόνες: μήκος $< n^2 - n + 1$

Παραγωγή κανόνων Golomb

Απο μια modular κατασκευή μπορούμε να πάρουμε έναν κανονικό κανόνα:

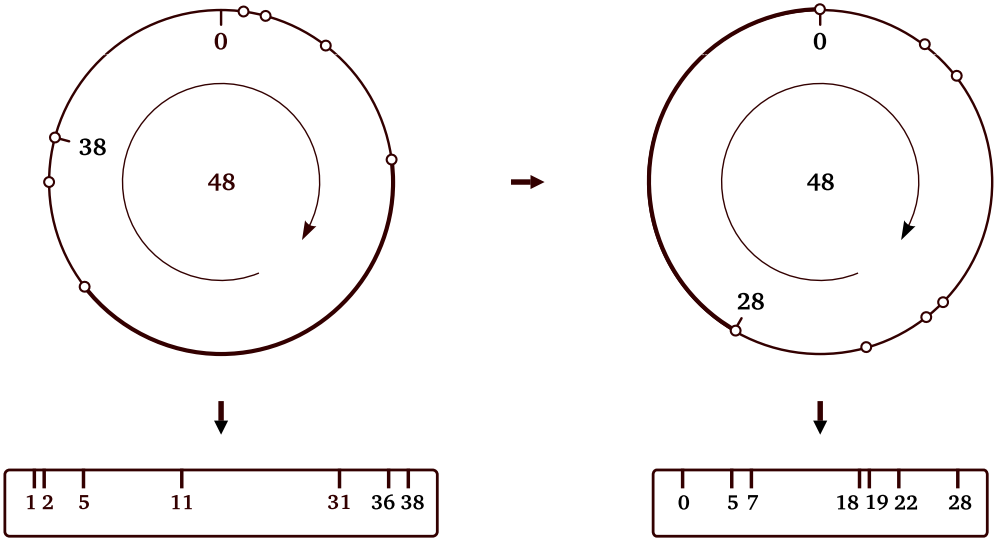
$$\{1, 2, 5, 11, 31, 36, 38\} \mod 48 \quad \text{(Bose-Chowla)}$$



- ένας κανόνας με 7 σημεία: $\{1, 2, 5, 11, 31, 36, 38\}$
- ένας κανόνας με 6 σημεία: $\{1, 2, 5, 11, 31, 36\}$

Περιστροφές

$$\{1, 2, 5, 11, 31, 36, 38\} \mod 48$$

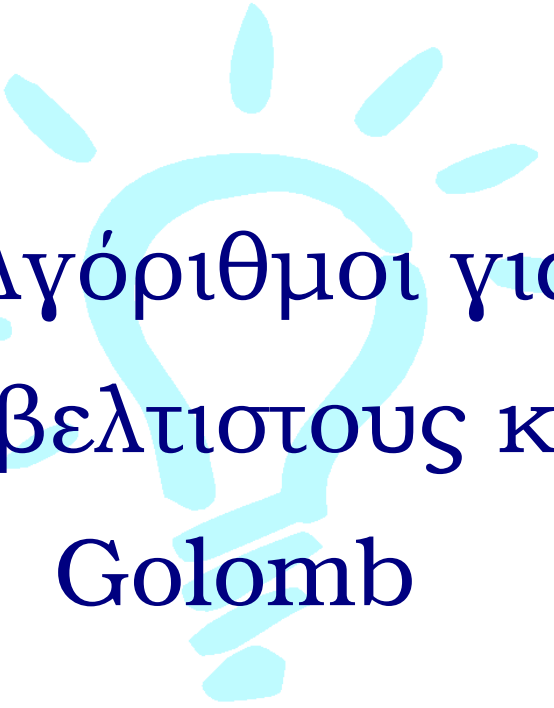


- Περιστρέφοντας την κατασκευή μπορούμε να πάρουμε έναν καλύτερο κανόνα

Πολλαπλασιάζοντας μια κατασκευή

Απο τη θεωρία αριθμών είναι γνωστό οτι αν πολλαπλασιάσουμε μια κατασκευή modulo q με έναν αριθμό g τέτοιο ώστε $(g, q) = 1$, παίρνουμε μια καινούργια modular κατασκευή.

- Το πλήθος των διαθέσιμων πολλαπλασιαστών είναι το πλήθος των ακεραίων $< q$ οι οποίοι δεν έχουν κοινό παράγοντα με το q: $(g, q) = 1$.
- $\phi(n)$: Η συνάρτηση ϕ του Euler είναι ακριβώς το πλήθος αυτό.
- Έτσι δοκιμάζοντας όλους τους πολλαπλασιαστές μπορούμε να βελτιώσουμε μια δεδομένη κατασκευή



Αλγόριθμοι για σχεδόν-βελτιστους κανόνες Golomb

Μια γνωστή εικασία

Εικασία: Υπάρχουν πάντα κανόνες με n σημάδια και μήκος κάτω απο n^2

$$G(n) < n^2$$

Στα σύνολα Sidon υπάρχει μια αντίστοιχη εικασία, που διατυπώθηκε για πρώτη φορά απο τον Erdős την δεκαετία του ′40.

Εικασία: Υπάρχουν πάντα σύνολα Sidon με μέγεθος $> \sqrt{n}$.

$$F_2(n) > \sqrt{n}$$

Οι δυο εικασίες είναι ισοδύναμες.

Το βασικό θεώρημα

- Ώς τώρα ήταν γνώστοι σχεδόν βέλτιστοι κανόνες μέχρι και 150 σημεία.
- Ο σκοπός του δεύτερου μέρους της εργασίας είναι η επέκταση αυτής της αναζήτησης σε πολύ μεγαλύτερα όρια και η απόδειξη του παρακάτω θεωρήματος:

Κανόνες Golomb με μήκος μικρότερο του n^2 υπάρχουν για όλα τα $n < 65000$. Ή ισοδύναμε στα σύνολα Sidon :

$$F_2(n) < \sqrt{n} \quad \text{για όλα τα } n \leqslant 4.225.000.000.$$

Κατασκευές που χρησιμοποιήθηκαν

- Για την απόδειξη πρέπει να χρησιμοποιήσουμε κάποιες απο τις κατασκευές που περιγράψαμε.
- Οι κατασκευές που δίνουν χάρακες με μήκος μικρότερο απο n^2, απαιτούν το n να είναι πρώτος αριθμός (ή δύναμη πρώτου).
- Για να καλύψουμε τα κενά στα οποία δεν μπορεί να εφαρμοστεί καμμία κατασκευή χρησιμοποιήσαμε την κατασκευή για τον αμέσως μεγαλύτερο δυνατό αριθμό σημείων και αφαιρέσαμε απο αυτήν κάποια σημεία.
- π.χ. για να φτιάξουμε ένα χάρακα με 10 σημεία αφαιρέσαμε 3 σημεία απο μια κατασκευή 13 σημείων.

Αλγόριθμοι

Δύο αλγόριθμοι αναπτύχθηκαν:

- RUZSA-EXTRACT$\{l,p\}$: Χρησιμοποιεί την κατασκευή του Ruzsa για τον πρώτο αριθμό p και παράγει χάρακες με $l, l+1, \ldots, p-1$ σημεία. Απαιτεί χρόνο:

$$T_1(l,p) = O(\ p^2 \log p + p^2(p-l)\ )$$

- BOSE-EXTRACT$\{l,p\}$: Χρησιμοποιεί την κατασκευή των Bose και Chowla για τον πρώτο αριθμό p και παράγει χάρακες με $l, l+1, \ldots, p$ σημεία. Απαιτεί χρόνο:

$$T_2(l,p) = O(\ p^3 \log p + p^2(p-l)\ )$$

Αναζήτηση & αποτελέσματα

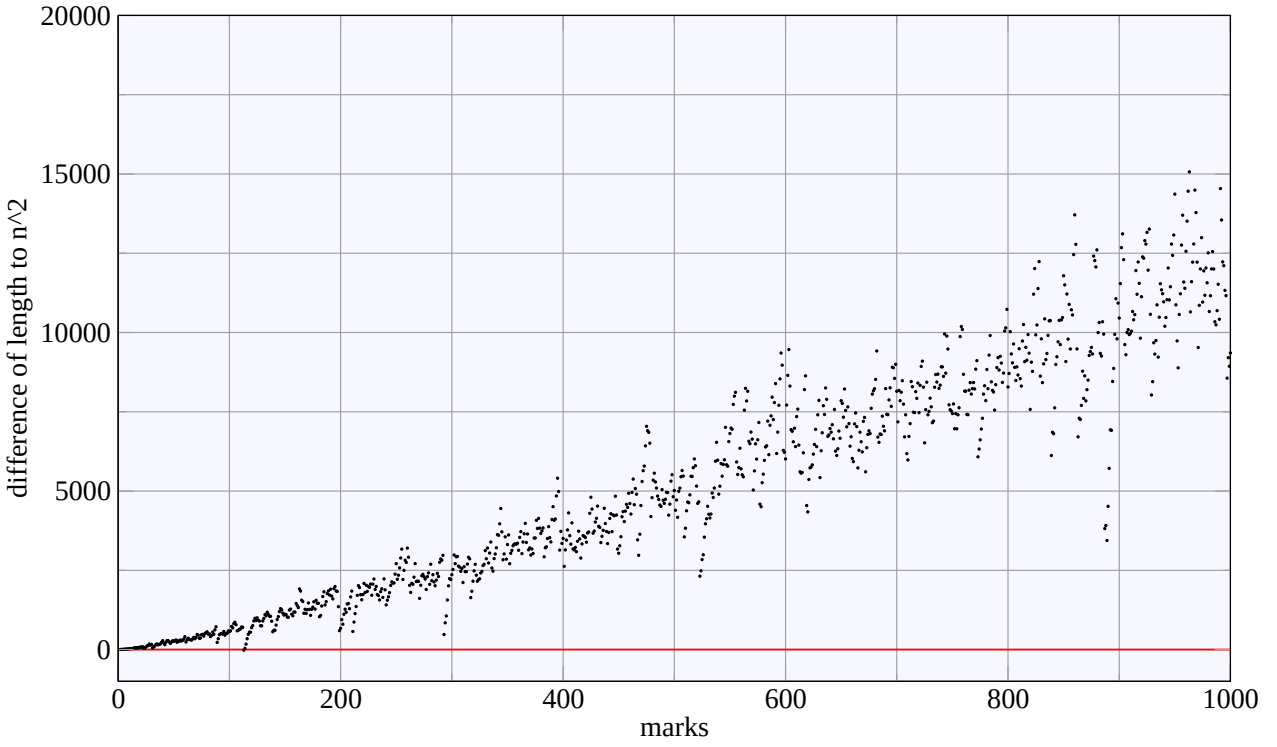
Υπολογιστικά στοιχεία

- Ο αλγόριθμος Ruzsa-Extract έτρεξε σε μια κατανεμημένη εκδοχή του για 4 μέρες σε ένα δίκτυο 10 υπολογιστών με συχνότητα CPU 1.5Ghz .
- Οι κλήσεις του αλγορίθμου χωρίζονται ως εξής:

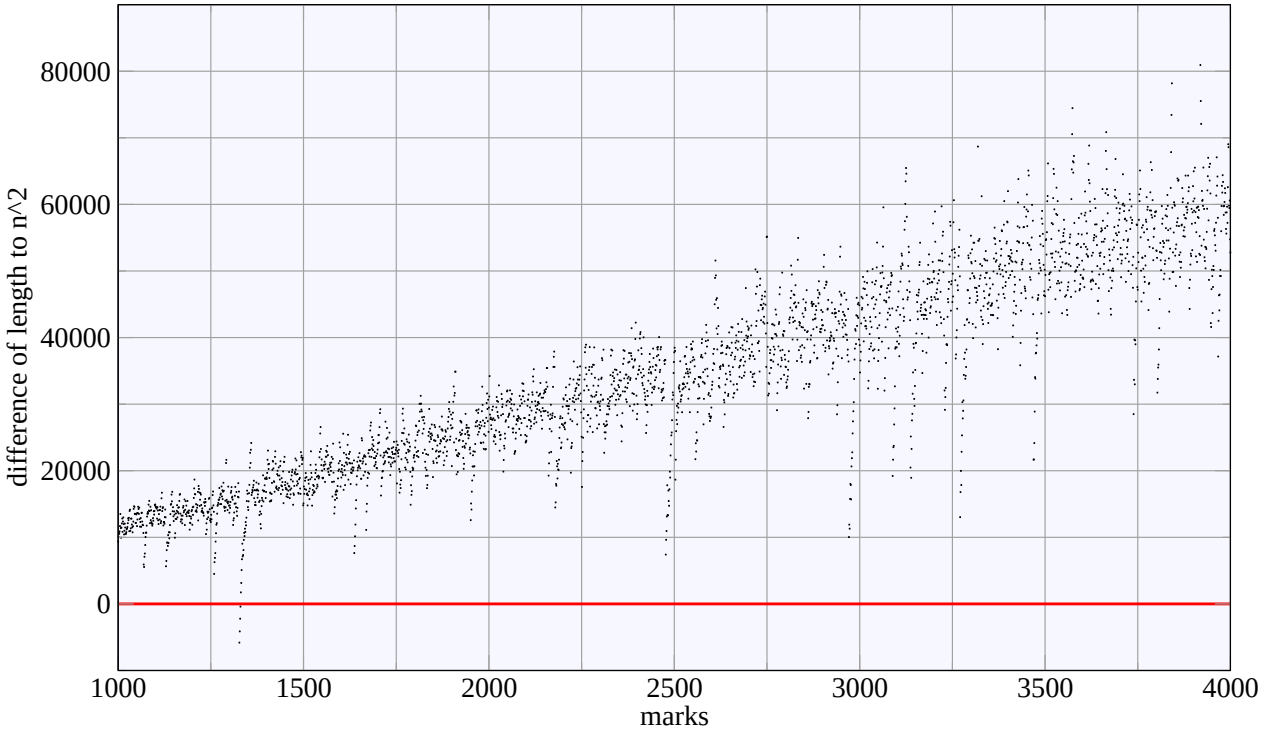
  $\textsc{Ruzsa-Extract}(\mathbf{2,3})$
  $\textsc{Ruzsa-Extract}(\mathbf{3,5})$
  $\textsc{Ruzsa-Extract}(\mathbf{5,7})$
  $\ldots$
  $\textsc{Ruzsa-Extract}(\mathbf{64997,65003})$

  Οι κλήσεις αυτές διανεμήθηκαν μέσω ενός interface client-server που υλοποιήθηκε σε TCL .
- Προβλήματα: σε κάποιες περιπτώσεις η κατασκευή δεν έδωσε κατάλληλους κανόνες. Οι περιπτώσεις αυτές αντιμετωπίστηκαν με την κατασκευή του Bose όπως θα δούμε.

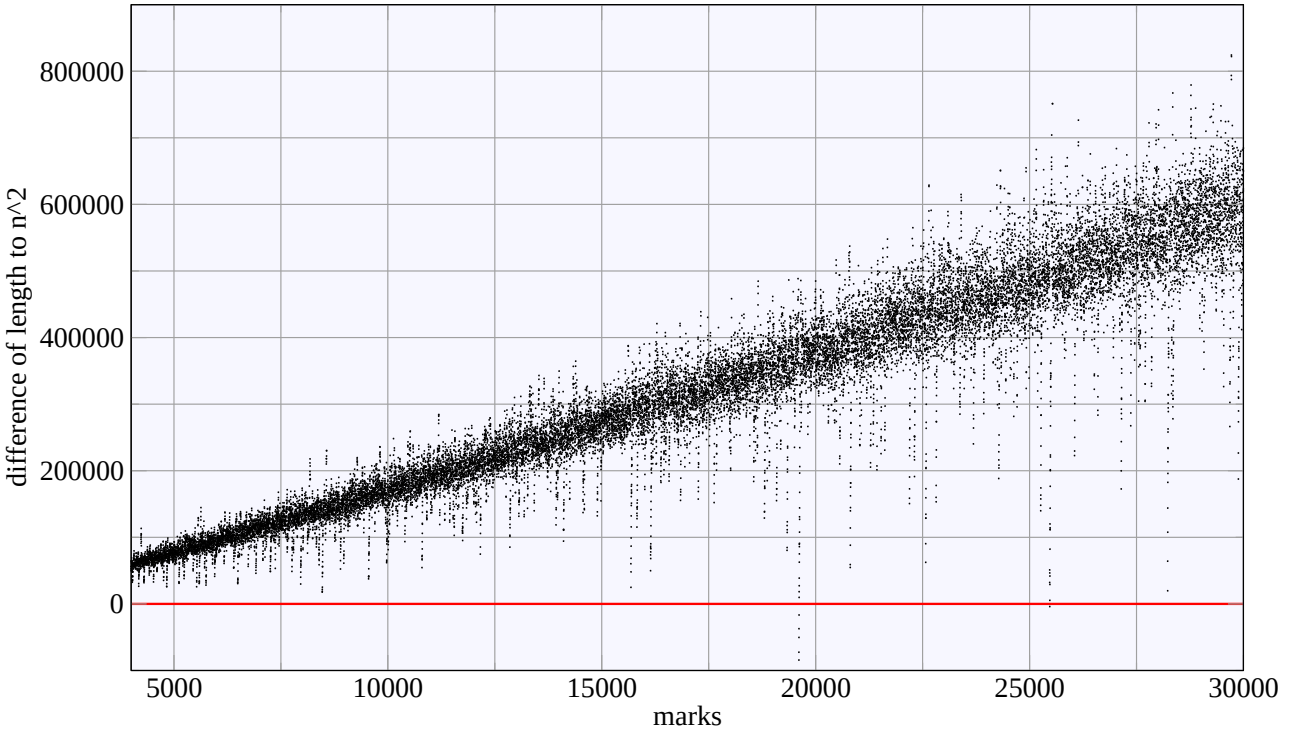
Αποτελέσματα (0-1000)



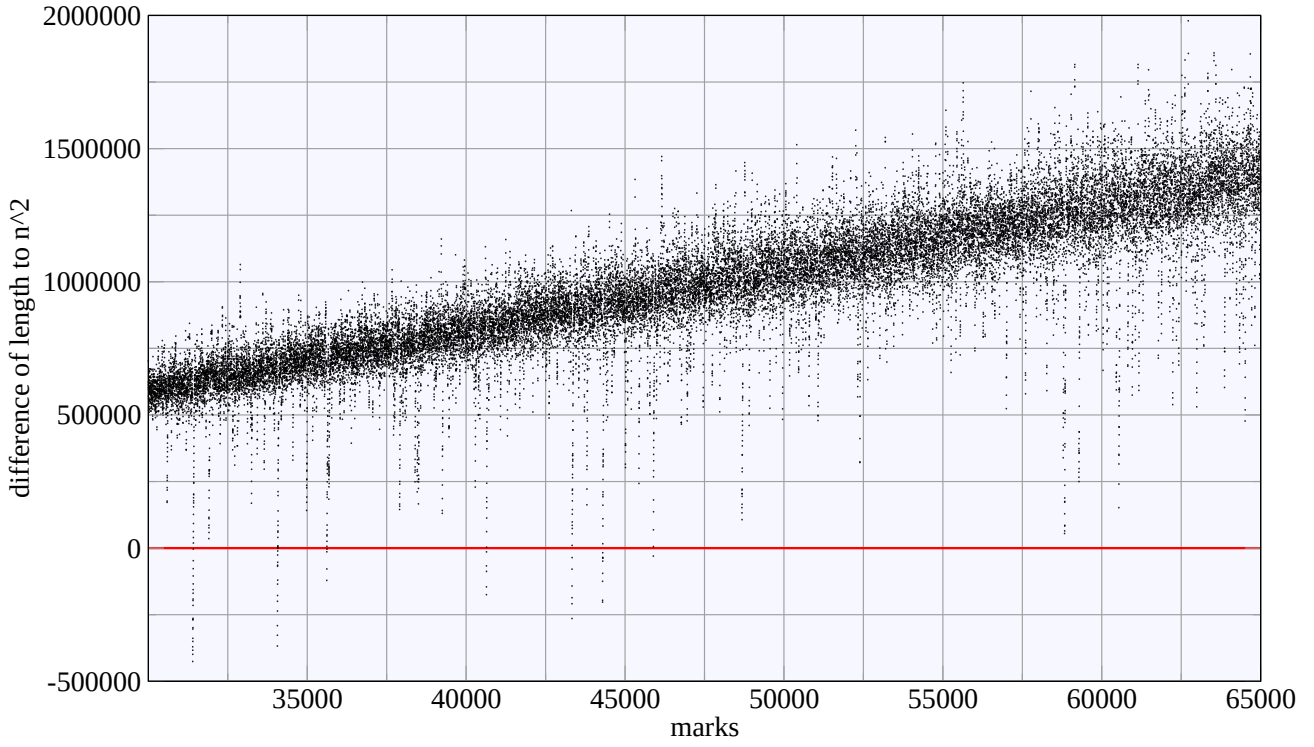
Αποτελέσματα (1000-4000)



Αποτελέσματα (4000-30000)



Αποτελέσματα (30000-65000)



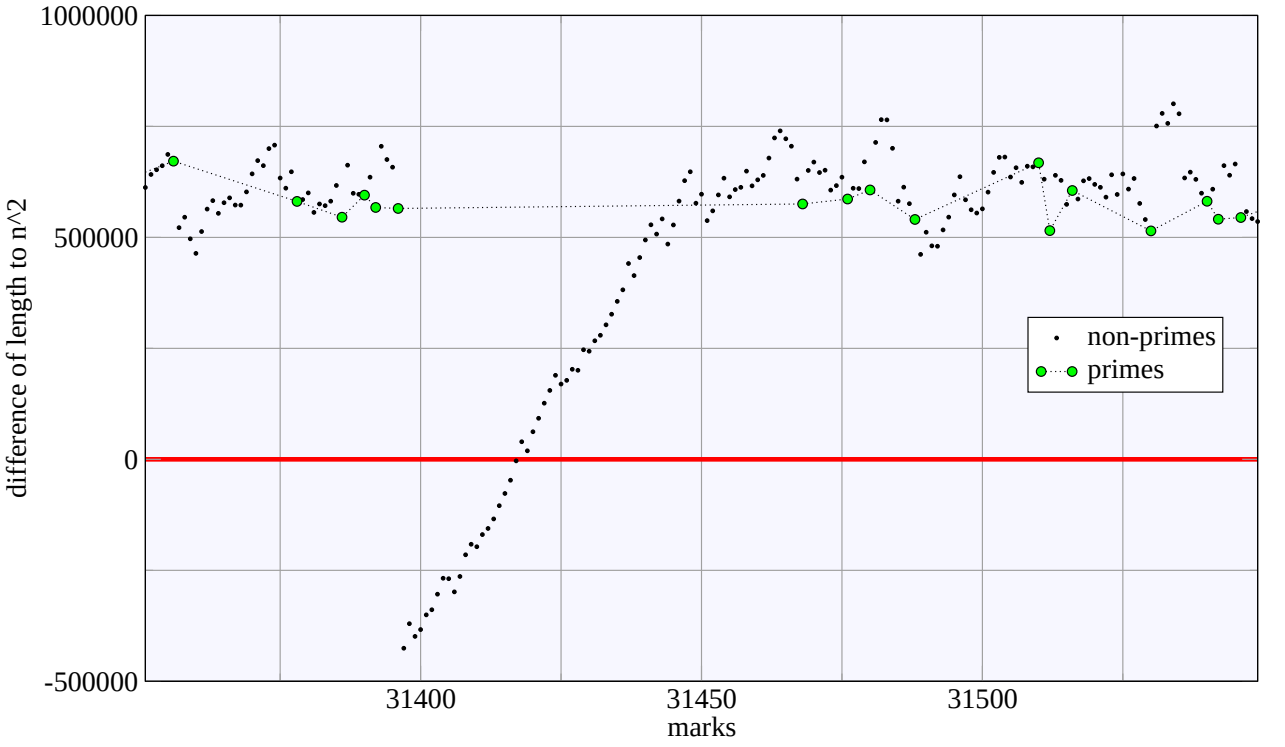
Αρνητικά αποτελέσματα

- Είχαμε αρνητικά αποτελέσματα ακριβώς στα διαστήματα οπού υπήρχε μεγάλο διάστημα χωρίς πρώτους αριθμούς. Συνολικά σε 72 περιπτώσεις απο τις 65000 απέτυχε η κατασκευή.

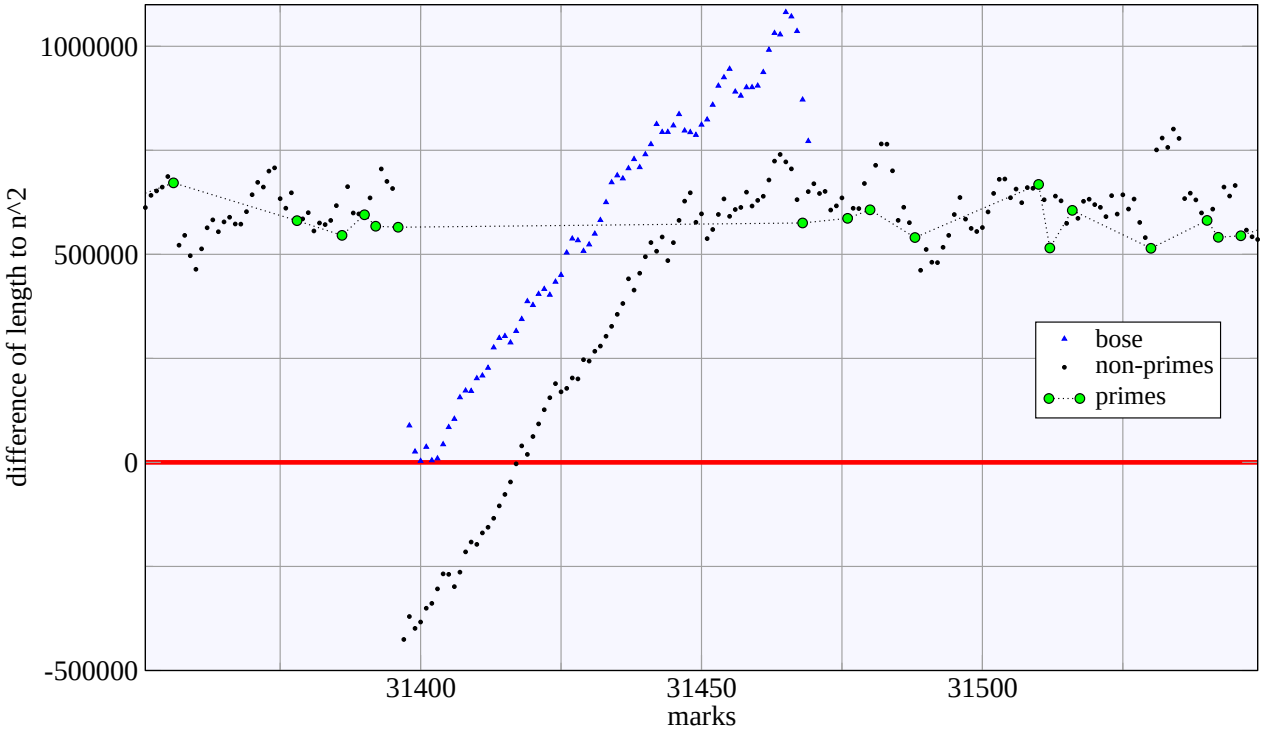
αριθμός σημείων	διάστημα μεταξύ πρώτων	μήκος διαστήματος
113	**113 – 127**	**14**
1327 – 1330	**1327 – 1361**	**34**
19609 – 19613	**19609 – 19661**	**52**
25474	**25471 – 25523**	**52**
31397 – 31417	**31397 – 31469**	**72**
34061 – 34074	**34061 – 34123**	**62**
35617 – 35623	**35617 – 35671**	**54**
40639 – 40643	**40639 – 40693**	**54**
43331 – 43336	**43331 – 43391**	**60**
44293 – 44301	**44293 – 44351**	**58**
45893	**45893 – 45943**	**50**

- Αντιμετώπιση: για να ολοκληρώσουμε την απόδειξη χρησιμοποιήσαμε τον δεύτερο αλγόριθμο (κατασκευή του Bose) σε όποιες περιπτώσεις απέτυχε ο πρώτος.

Μια άσχημη περίπτωση



Αρνητικά αποτελέσματα



Επίλογος - Μελλοντικές επεκτάσεις

- Απόδειξη ενός καινούργιου κάτω φράγματος για την συνάρτηση $G(n)$:

$$G(n) > n^2 - 2n\sqrt{n} + \sqrt{n} - 2$$

- Χρησιμοποιώντας δύο κατασκευές απο την θεωρία αριθμών αποδείξαμε ότι κανόνες Golomb με μήκος μικρότερο του n^2 υπάρχουν για όλα τα $n < 65000$. Ή ισοδύναμε στα σύνολα Sidon :

$$F_2(n) < \sqrt{n} \quad \text{για όλα τα } n \leqslant 4.225.000.000.$$

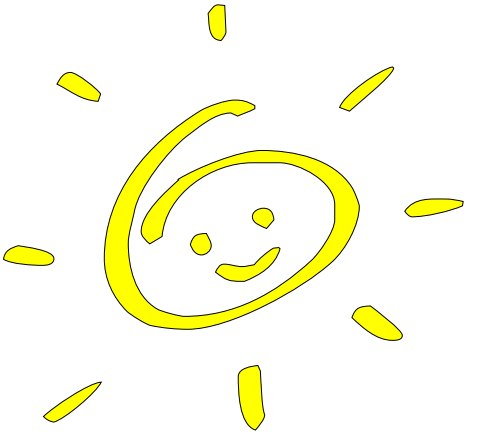
- Όλα τα αποτελέσματα και οι κώδικες υπάρχουν στην σελίδα:
  http://www.softnet.tuc.gr/~apdim/diploma.

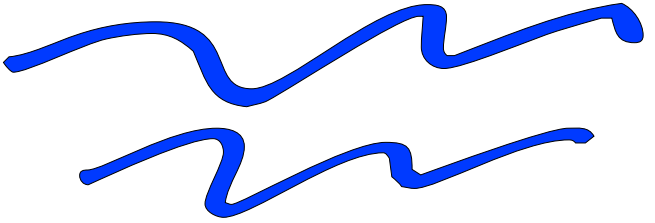
- Μελλοντικα: επέκταση της αναζήτησης σε ακόμα μεγαλύτερα όρια και αυτοματοποίηση της διαδικασίας παραγωγής κανόνων.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω:

- Τον κ. Δόλλα για την επιβλέψη και καθοδήγηση του.
- Την κοινότητα του linux για όλες τια ανάγκες σε λογισμικό αυτής της εργασίας απο την στοιχειοθεσία ως και την παρουσίαση.
- Όλους τους φίλους/ες που συμπαραστάθηκαν για να ολοκληρωθεί αυτή η εργασία.





Καλό καλοκαίρι!