Α44 – ΚΡΥΠΤΟΓΡΑΦΙΑ
ΣΗΜΕΙΩΣΕΙΣ #8

ΘΕΟΔΟΥΛΟΣ ΓΑΡΕΦΑΛΑΚΗΣ

1. Γενικη περιγραφη της μεθοδου

Η μέθοδος παραγοντοποίησης του Fermat επιχειρεί να γράψει τον αριθμό n που θέλουμε να παραγοντοποιήσουμε ως διαφορά τετραγώνων, δηλαδή να υπολογίσει $x, y, \in \mathbb{Z}$ τέτοιους ώστε $n = x^2 - y^2$. Αυτό όμως είναι γενικά δύσκολο και γι' αυτό η μέθοδος δεν είναι αποτελεσματική για μεγάλους αριθμούς. Οδηγεί όμως στην παρατήρηση ότι αν βρούμε ακεραίους x, y τέτοιους ώστε

$$x^2 \equiv y^2 \pmod{n}, \tag{1}$$

τότε το n διαιρεί το $(x + y)(x - y)$. Άρα, ο μ.κ.δ. $(x + y, n)$ θα μας δώσει μια παραγοντοποίηση του n (όχι αναγκαστικά σε πρώτους παράγοντες), εκτός αν $(x + y, n) = 1$ ή n. Με άλλα λόγια η μέθοδος αποτυγχάνει αν

$$x \equiv \pm y \pmod{n}.$$

Αν οι αριθμοί x, y, οι οποίοι ικανοποιούν την (1) επιλεγούν τυχαία, μπορεί κανείς να δείξει ότι η πιθανότητα αποτυχίας είναι το πολύ $1/2$.

Πώς μπορούμε να βρούμε ακεραίους x, y που να ικανοποιούν την (1); Ας σταθεροποιήσουμε το σύνολο

$$S = \{2, 3, 5, \ldots, p_t\}$$

των πρώτων t πρώτων αριθμών. Θα λέμε ότι ένας αριθμός παραγοντοποιήται πάνω στο S αν όλοι οι πρώτοι παράγοντες του ανοίκουν στο S. Έστω ότι έχουμε βρει αρκετά ζεύγη αριθμών $(a_i, b_i), i = 1, 2, \ldots, s$ τέτοια ώστε οι a_i και b_i παραγοντοποιούνται πάνω στο S, ας πούμε

$$a_i = \prod_{j=1}^{t} p_j^{e_{ij}} \quad i = 1, 2, \ldots, s$$

$$b_i = \prod_{j=1}^{t} p_j^{d_{ij}} \quad i = 1, 2, \ldots, s$$

και επιπλέον

$$a_i \equiv b_i \pmod{n} \quad i = 1, 2, \ldots, s.$$

Τότε έχουμε

$$\prod_{j=1}^{t} p_j^{e_{ij}} \equiv \prod_{j=1}^{t} p_j^{d_{ij}} \quad i = 1, 2, \ldots, s. \tag{2}$$

Στη συνέχεια θέλουμε να βρούμε ποιές σχέσεις να πολλαπλασιάσουμε κατά μέλη, έτσι ώστε να πάρουμε μια ισοτιμία στην οποία όλοι οι εκθέτες να είναι άρτιοι. Αυτό θα μας δώσει τα δύο τετράγωνα που ψάχνουμε.

Με τις παραπάνω σκέψεις βλέπουμε η μέθοδος μας έχει δύο στάδια: πρώτον να κατασκευάσουμε τις σχέσεις, και δεύτερον να βρούμε ποιες σχέσεις να πολλαπλασιάσουμε.

2. Συδοιασμος σχεσεων

Ας ξεκινήσουμε την περιγραφή μας από το δεύτερο στάδιο. Ας υποθέσουμε ότι έχουμε κατασκευάσει αρκετές σχέσεις της μορφής (2). Αυτό που μας ενδιαφέρει πραγματικά για το συνδοιασμό των σχέσεων δεν είναι η τιμή του κάθε εκθέτη e_{ij} και d_{ij} αλλά αν είναι άρτιος ή περιττός. Με άλλα λόγια μας ενδιαφέρει η τιμή του κάθε εκθέτη mod 2. Ας θεωρήσουμε το διάνυσμα $(x_1, \ldots, x_s) \in \mathbb{F}_2^s$. Το διάνυσμα που προκύπτει αν πολλαπλασιάσουμε κατά μέλη τις σχέσεις με δείκτη i όπου $x_i = 1$, η σχέση που προκύπτει είναι

$$\prod_{j=1}^{t} p_j^{\sum_{i=1}^{s} x_i e_{ij}} \equiv \prod_{j=1}^{t} p_j^{\sum_{i=1}^{s} x_i d_{ij}} \pmod{n}. \tag{3}$$

Η τελευταία σχέση είναι σχέση τετραγώνων αν και μόνο αν

$$\sum_{i=1}^{s} (e_{ij} - d_{ij}) x_i \equiv 0 \pmod{2} \quad j = 1, 2, \ldots, s. \tag{4}$$

Μένει να βρούμε ένα διάνυσμα $\mathbf{x} = (x_1, \ldots, x_s)$ που να ικανοποιεί την (4), δηλαδή μια λύση του γραμμικού συστήματος

$$\mathbb{A}\mathbf{x} = \mathbf{o},$$

όπου $\mathbb{A}$ είναι ένας $t \times s$ πίνακας με στοιχεία στο $\mathbb{F}_2$ και το (i, j) στοιχείο του είναι το $e_{ij} - d_{ij}$, το $\mathbf{x}$ είναι το διάνυσμα των αγνώστων, διάστασης s και το $\mathbf{o}$ είναι το μηδενικό διάνυσμα διάστασης t.

Μπορούμε να λύσουμε το σύστημα με μεθόδους της γραμμικής άλγεβρας. Αν το σύστημα δεν έχει λύση σημαίνει ότι δεν έχουμε αρκετές σχέσεις, οπότε κατασκευάζουμε περισσότερες με τη μέθοδο που θα περιγράψουμε στην επόμενη ενότητα.

3. Κατασκευη σχεσεων

Περνάμε στην περιγραφή της κατασκευής των σχέσεων (2). Αν

$$a \equiv b \pmod{n}$$

τότε

$$b = a - \lambda n$$

για κάποιο $\lambda \in \mathbb{Z}$. Ψάχνουμε να βρούμε αρκετά ζεύγη (a, b) που επιπλέον να παραγοντοποιούνται πάνω στο S. Κάθε τέτοιο ζεύγος μας δίνει και μία σχέση. Σταθεροποιούμε το λ και εξετάζουμε «μικρές» τιμές του a, ας πούμε $2 \leq a \leq B$ ώστε με μεγάλη πιθανότητα το a να παραγοντοποιήται πάνω στο S. Για κάθε τιμή του a που εξετάζουμε, πρέπει να ελέγξουμε αν το $b = a - \lambda n$ παραγοντοποιήται πάνω στο S και αν παραγοντοποιήται να βρούμε αυτή την παραγοντοποίηση.

Μια πρώτη σκέψη είναι για κάθε τιμή του a να παραγοντοποιούμε το b. Αυτό όμως φαίνεται πολύ χρονοβόρο, καθώς οι περισσότεροι από τους αριθμούς που καλούμαστε να παραγοντοποιήσουμε *δεν* παραγοντοποιούνται πάνω στο S. Μια καλύτερη μέθοδος είναι η παρακάτω.

Γράφουμε τους αριθμούς $a = 2, 3, \ldots, B$ σε μια λίστα. Δίπλα σε κάθε αριθμό a γράφουμε το $\log(a-\lambda n)$ (δεν είναι απαραίτητη καλή προσέγγιση του λογάριθμου (το ακέραιο μέρος είναι αρκετό). Για ένα πρώτο $p \in S$, οι αριθμοί $a - \lambda n$ με $2 \leq a \leq B$ που διαιρούνται με το p ικανοποιούν την

$$a \equiv \lambda n \pmod{p}.$$

Διατρέχουμε τη λίστα και σημειώνουμε δίπλα στα a που ικανοποιούν την παραπάνω ισοτιμία το p^e, όπου p^e είναι η μεγαλύτερη δύναμη του p που διαιρεί το $a - \lambda n$ και αφαιρούμε από την τιμή του λογάριθμου το $e \log p$. Το κάνουμε αυτό για κάθε πρώτο $p \in S$. Αφού έχουμε τελειώσει με τους πρώτους στο S κοιτάμε ποιά στοιχεία στη λίστα έχουν τιμή λογάριθμου κοντά στο 0. Αυτά τα στοιχεία παραγοντοποιούνται πάνω στο S και η ανάλυση τους έχει ήδη βρεθεί. Ας πάρουμε ένα τέτοιο στοιχείο a. Το $b = a - \lambda n$ αναλύεται πάνω στο S και ελπίζουμε ότι το ίδιο ισχύει και για το a. Το ελέγχουμε παραγοντοποιώντας το a, ας πούμε με δοκιμαστικές διαιρέσεις. Αν όντως συμβαίνει, όπως περιμένουμε, τότε παίρνουμε μια σχέση της μορφής (2).

Παράδειγμα 3.1. Έστω $n = 1271$. Επιλέγουμε το $S = \{2, 3, 5\}$ και σταθεροποιούμε το $\lambda = -2$. Θα εξετάσουμε τους αριθμούς $2 \leq a \leq 10$. Στην πρώτη φάση, για τη δημιουργία των σχέσεων γράφουμε τη λίστα μας

a	2	3	4	5	6	7	8	9	10
log	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8	7.8

Οι αριθμοί $a - \lambda n = a + 2 \cdot 1271 = a + 2542$ με $2 \leq a \leq 10$ που διαιρούνται με το 2 είναι αυτοί που ικανοποιούν την

$$a + 2542 \equiv 0 \pmod{2},$$

δηλαδή για a άρτιο. Συμπληρώνουμε τον πίνακα.

a	2	3	4	5	6	7	8	9	10
log	5	7.8	7.1	7.8	6.4	7.8	7.1	7.8	5.7
	2^4		2		2^2		2		2^3

Προχωρούμε στον επόμενο πρώτο, το 3. Οι αριθμοί $a + 2542$ που διαιρούνται με το 3 ικανοποιούν την

$$a + 2542 \equiv 0 \pmod{3} \iff a \equiv 2 \pmod{3},$$

δηλαδή αντιστοιχούν στους αριθμους $a = 2, 5, 8$ στη λίστα μας.

a	2	3	4	5	6	7	8	9	10
log	3.9	7.8	7.1	5.6	6.4	7.8	6	7.8	5.7
	$2^4 \cdot 3$		2	3^2	2^2		$2 \cdot 3$		2^3

Εξετάζουμε και τον τελευταίο πρώτο στο S, το 5. Οι αριθμοί $a + 2542$ που διαιρούνται με το 5 ικανοποιούν την

$$a + 2542 \equiv 0 \pmod{5} \iff a \equiv 3 \pmod{5},$$

δηλαδή αντιστοιχούν στους αριθμούς $a = 3, 8$ στη λίστα μας.

a	2	3	4	5	6	7	8	9	10
log	3.9	6.2	7.1	5.6	6.4	7.8	4.4	7.8	5.7
	$2^4 \cdot 3$	5	2	3^2	2^2		$2 \cdot 3 \cdot 5$		2^3

Με την επιλογή των παραμέτρων μας, βλέπουμε ότι, αφού εξετάσαμε όλους τους πρώτους στο S, κανένας λογάριθμος δεν έχει φτάσει κοντά στο 0. Δεν έχουμε πάρει καμμία σχέση. Πραγματικά, κανείς μπορεί να ελέγξει ότι κανένας από τους αριθμούς $a + 2542$ για $a = 2, 3, \ldots, 10$ δεν αναλύεται πλήρως πάνω στο S. Μπορούμε να διορθώσουμε το πρόβλημα είτε μεγαλώνοντας το S, είτε εξετάζοντας περισσότερες τιμές για το a, είτε και τα δύο. Για παράδειγμα, αν παίρναμε $S = \{2, 3, 5, 17\}$ τότε έχουμε

$$a + 2542 \equiv 0 \pmod{17} \iff a \equiv 8 \pmod{17},$$

που αντιστοιχεί στο $a = 8$ στη λίστα μας.

a	2	3	4	5	6	7	8	9	10
log	3.9	6.2	7.1	5.6	6.4	7.8	0	7.8	5.7
	$2^4 \cdot 3$	5	2	3^2	2^2		$2 \cdot 3 \cdot 5^2 \cdot 17$		2^3

Βλπεπουμε ότι το $a = 8$ μας δίνει μία σχέση, την

$$2 \cdot 3 \cdot 5^2 \cdot 17 \equiv 2^3 \pmod{1271}.$$

Η αντίστοιχη ισοτιμία για τους εκθέτες είναι

$$(1, 0, 0, 0) \equiv (1, 1, 0, 1) \pmod{2},$$

ή ισοδύναμα

$$(0, 1, 0, 1) \equiv (0, 0, 0, 0) \pmod{2},$$

όπου η πρώτη συντεταγμένη αντιστοιχεί στο 2, η δεύτερη στο 3, η τρίτη στο 5 και η τέταρτη στο 17. Το $(0, 1, 0, 1)$ είναι η πρώτη γραμμή του πίνακα $\mathbb{A}$.

4. Σχολια

Όλοι οι αλγόριθμοι παραγοντοποίησης που βασίζονται στο συνδοιασμό σχέσεων για τη δημιουργία διαφοράς τετραγώνων λειτουργούν σε δύο φάσεις: της δημιουργίας των σχέσεων και της επίλυσης του γραμμικού συστήματος. Η φάση της επίλυσης του συστήματος είναι η ίδια για όλους τους αλγόριθμους. Αυτό που διαφέρει είναι ο τρόπος δημιουργίας των σχέσεων, δηλαδή η πρώτη φάση.

Υπάρχουν αλγόριθμοι που η χρονική τους πολυπλοκότητα *αποδεικνύεται* ότι είναι

$$\exp\left((c_1 + o(1))(\log n)^{1/2}(\log\log n)^{1/2}\right).$$

Οι πιο πρακτικοί αλγόριθμοι έχουν την παραπάνω πολυπλοκότητα, χωρίς όμως αυτό να μπορεί να αποδειχτεί αυστηρά. Οι σημαντικότεροι αλγόριθμοι αυτής της κατηγορίας είναι το Quadratic Sieve και το Continued Fractions Sieve. Η μέθοδος που περιγράψαμε εδώ είναι γνωστη ως Linear Sieve.

Τέλος υπάρχουν αλγόριθμοι των οποίων η πρώτη φάση βασίζεται σε αλγεβρική θεωρία αριθμών και ικάζεται ότι έχουν πολυπλοκότητα

$$\exp\left((c_2 + o(1))(\log n)^{1/3}(\log\log n)^{2/3}\right).$$