Α 44 – ΚΡΥΠΤΟΓΡΑΦΙΑ
ΣΗΜΕΙΩΣΕΙΣ #6

ΘΕΟΔΟΥΛΟΣ ΓΑΡΕΦΑΛΑΚΗΣ

1. Διακριτοι λογαριθμοι σε πεπερασμενα σωματα

Όπως είδαμε σε περασμένο μάθημα, σε μία ομάδα τάξης N μπορούμε να υπολογίζουμε διακριτούς λογάριθμους σε χρόνο $O(\sqrt{N})$. Αν για παράδειγμα έχουμε ένα πεπερασμένο σώμα $\mathbb{F}_q$, όπου $q = p^n$, p πρώτος, τότε το παραπάνω φράγμα γίνεται $O(\sqrt{q}) = O(p^{n/2})$. Παρακάτω θα δούμε μια ομάδα αλγορίθμων, των λεγόμενων Index Calculus, που βελτιώνουν το φράγμα σε

$$\exp\left((c + o(1))\sqrt{\log q \log\log q}\right),$$

για κάποια σταθερά c που εξαρτάται από διάφορους παράγοντες, όπως η σχέση των p και n. Για να δει κανείς τη διαφορά ανάμεσα στις δύο πολυπλοκότητες, αρκεί να δει ότι το φράγμα $O(\sqrt{q})$ είναι ίσο με $O(\exp(0.5 \log q))$.

2. Μαθηματικα προαπαιτουμενα

2.1. **Διακριτοί λογάριθμοι.** Έστω μια πεπερασμένη αβελιανή ομάδα G, $g \in G$ και $h, g_1, \ldots, g_t \in \langle g \rangle$. Το διακριτό λογάριθμο οποιουδήποτε στοιχείου y ως προς τη βάση g θα τον συμβολίζουμε με $\log y$.

Βλέπουμε τώρα ότι οι βασικοί κανόνες που ισχύουν για τους "κανονικούς" λογάριθμους ισχύουν και εδώ. Ειδικότερα, αν

$$h = g_1^{e_1} \cdots g_t^{e_t},$$

τότε

$$g^{\log h} = g^{e_1 \log g_1} \cdots g^{e_t \log g_t} = g^{e_1 \log g_1 + \cdots + e_t \log g_t},$$

άρα

$$\log h \equiv e_1 \log g_1 + \cdots + e_t \log g_t \pmod{|\langle g \rangle|}.$$

2.2. **Παραγοντοποίηση πολυωνύμων.** Για τον αλγόριθμο Index Calculus σε επεκτάσεις $\mathbb{F}_{p^n}$ όπου το p είναι μικρό (στην πράξη $p = 2$), και το n πολύ μεγάλο (στην πράξη κανείς παίρνει $n \geq 1000$), είναι σημαντικό να ξέρουμε ότι μπορούμε να παραγοντοποιούμε πολυώνυμα πάνω από το $\mathbb{F}_p$ εύκολα. Ένα πολυώνυμο $h \in \mathbb{F}_p[X]$ βαθμού m παραγοντοποιήται σε χρόνο $O(m^3 \log p)$. Τα ανάγωγα πολυώνυμα στο $\mathbb{F}_p[X]$ θα τα λέμε και πρώτους του $\mathbb{F}_p[X]$.

3. Index Calculus

Θα περιγράψουμε τη μέθοδο Index Calculus για σώματα της μορφής $\mathbb{F}_{p^n}$ όπου το p είναι μικρό και το n μεγάλο. Στην ανάλυση του αλγορίθμου αυτό σημαίνει ότι το p είναι σταθερό και το n τείνει στο άπειρο. Γνωρίζουμε ότι $\mathbb{F}_{p^n} \approx \mathbb{F}_p/\langle f\rangle$ όπου το f είναι ανάγωγο πολυώνυμο πάνω στο $\mathbb{F}_p$ βαθμού n. Τα στοιχεία λοιπόν του $\mathbb{F}_{p^n}$ μπορούν να αναπαρασταθούν ως πολυώνυμα πάνω στο $\mathbb{F}_p$ βαθμού έως και $n-1$ και η αριθμητική γίνεται $\mod f$.

Και τώρα ο αλγόριθμος. Δεδομένα είναι δύο πολυώνυμα $y, g \in \mathbb{F}_p[X]$ βαθμού το πολύ $n-1$ και ζητάμε το $\log y$ (ως προς βάση g). Ο αλγόριθμος έχει δύο φάσεις. Αρχικά επιλέγουμε ένα σύνολο S που περιέχει ανάγωγα πολυώνυμα. Το S είναι η πιο βασική παράμετρος του αλγόριθμου.

Φάση 1

(1) Επέλεξε ακέραιο $s \in [1, q-1]$ με ομοιόμορφη κατανομή και υπολόγισε το πολυώνυμο $h \equiv g^s \pmod f$, $\deg h < n$.
(2) Αν το h δεν παραγοντοποιήται πλήρως πάνω στο S, πέταξε το. Αν όλοι οι πρώτοι παράγοντες του h είναι στο S, γράψε

$$h = \prod_{v \in S} v^{e_v(h)},$$

και κατάγραψε τη σχέση

$$s \equiv \sum_{v \in S} e_v(h) \log_g v \pmod{p^n - 1}.$$

Επανάλαβε τα παραπάνω βήματα μέχρι να έχεις περίπου $4n|S| \log p$ σχέσεις. Τότε λύσε το γραμμικό σύστημα για να υπολογίσεις τους $\log v$ για όλα τα $v \in S$.

Φάση 2

(1) Επέλεξε ακέραιο $s \in [1, q-1]$ με ομοιόμοφρη κατανομή και υπολόγισε το πολυώνυμο $h \equiv yg^s \pmod f$, $\deg h < n$.
(2) Αν το h δεν παραγοντοποιήται πάνω στο S, πέταξε το. Αν όλοι οι πρώτοι παράγοντες του h είναι στο S, γράψε

$$h = \prod_{v \in S} v^{e_v(h)},$$

και υπολόγισε το $\log y$ ως

$$\log_g y \equiv -s + \sum_{v \in S} e_v(h) \log_g v \pmod{p^n - 1}.$$

Η ορθότητα του αλγόριθμου είναι φανερή από όσα είπαμε στην παράγραφο 2.1.

4. Ανάλυση του αλγοριθμου

Κατ' αρχάς, για να αναλύσει κανείς τον αλγόριθμο πρέπει να ορίσει ακριβώς το S. Το S τυπικά επιλέγεται να περιέχει όλα τα ανάγωγα πολυώνυμα βαθμού έως και m, όπου το m είναι μια παράμετρος που μπορούμε να επιλέξουμε κατά βούληση.

Μπορεί κανείς να δείξει ότι έχοντας δημιουργήσει $4n|S|\log p$ σχέσεις, η πιθανότητα το σύστημα να έχει πλήρη τάξη (δηλαδή να έχουμε $|S|$ γραμμικά ανεξάρτητες εξισώσεις) είναι τουλάχιστον $1/2$.

Η πολυπλοκότητα της πρώτης φάσης είναι φανερά μεγαλύτερη από αυτή της δεύτερης φάσης και είναι

$$4n|S|\log p\frac{1}{P(n-1,m)}+(4n|S|\log p)^2,$$

όπου $P(n-1,m)$ είναι η πιθανότητα ένα πολυώνυμο βαθμού το πολυ $n-1$ να έχει όλους τους ανάγωγους παράγοντες του βαθμού το πολύ m. Τέτοια πολυώνυμα ονομάζονται m-ομαλά. Ο πρώτος όρος είναι για τη δημιουργία του συστήματος και ο δεύτερος για την επίλυση του. Ο λόγος που στον εκθέτη έχουμε τετράγωνο και όχι κύβο (όπως στην κλασική απαλειφή Gauss), είναι ότι το σύστημα είναι πολύ αραιό – οι περισσότεροι συντελεστές είναι μηδέν – και για τέτοια συστήματα υπάρχουν ειδικοί αλγόριθμοι.

Κανείς μπορεί να δείξει εύκολα ότι

$$|S|=\exp\left((1+o(1))m\log p\right),$$

και με περισσότερη δυσκολία ότι

$$P(n-1,m)=\exp\left((1+o(1))\frac{n}{m}\log\frac{n}{m}\right).$$

Άρα η πολυπλοκότητα της πρώτης φάσης γίνεται

$$\exp\left((1+o(1))\frac{n}{m}\log\frac{n}{m}\right)+\exp\left((1+o(1))2\frac{n}{m}\log p\right)$$

το οποίο είναι ίσο με

$$\exp\left((1+o(1))\left(\frac{n}{m}\log\frac{n}{m}+2\frac{n}{m}\log p\right)\right).$$

Κανείς μπορεί να δείξει ότι η βέλτιστη επιλογή για το m είναι

$$m=\sqrt{\frac{n\log n}{2\log p}},$$

και τότε η πολυπλοκότητα του αλγόριθμου γίνεται

$$\exp\left(\left(\sqrt{2\log p}+o(1)\right)\sqrt{n\log n}\right).$$

Η πολυπλοκότητα της δείτερης φάσης βλέπουμε ότι είναι

$$\frac{1}{P(n-1,m)} = \exp\left(\left(\sqrt{0.5\log p} + o(1)\right)\sqrt{n\log n}\right).$$

5. Παρατηρησεις

Η πρώτη παρατήρηση είναι ότι η πρώτη φάση του αλγόριθμου έχει στόχο τον υπολογισμό των διακριτών λογάριθμων των πολυωνύμων στη βάση S. Αυτό χρειάζεται να γίνει μόνο μία φορά. Για κάθε νέο λογάριθμο που θέλουμε να υπολογίσουμε στην ίδια ομάδα, εκτελούμε μόνο τη δεύτερη φάση.

Σύμφωνα με όσα είπαμε παραπάνω, μπορούμε να υπολογίσουμε διακριτούς λογάριθμους στο $\mathbb{F}_{p^n}$ σε χρόνο περίπου ίσο με

$$\exp\left(\log p\sqrt{n\log n}\right) = p^{\sqrt{n\log n}}.$$

Συγκρίνετε αυτό με την πολυπλοκότητα του Baby-step/Giant-step που είναι

$$p^{n/2}.$$

Για παράδειγμα, για $p = 2$ και $n \sim 1000$ ο Index Calculus κάνει περίπου $2^{83.12}$ βήματα, ενώ ο Baby-Step/Giant-step κάνει 2^{500}. Ακόμη και αν ο μεγαλύτερος πρώτος διαιρέτης του $2^n - 1$ είναι $\sim 2^{600}$, ο συνδοιασμός του Pohlig-Hellman με τον Baby-step/Giant-step χρειάζεται περιπου 2^{300} βήματα.

Ένα χαρακτηριστικό του Index Calculus είναι ότι δε λαμβάνει υπόψη τις υποομάδες του $\mathbb{F}_{p^n}$. Δηλαδή η πολυπλοκότητα του δεν εξαρτάται από το μέγεθος των πρώτων παραγόντων του $p^n - 1$ που είναι η τάξη της ομάδας. Αν, για παράδειγμα, $p = 2$ και $n \sim 1000$ όπως πριν και ο μεγαλύτερος πρώτος διαιρέτης του $2^n - 1$ είναι $\sim 2^{80}$, ο Index Calculus κάνει και πάλι $\sim 2^{83.12}$ βήματα, ενώ ο συνδοιασμός του Pohlig-Hellman με τον Baby-step/Giant-step χρειάζεται μόνο $\sim 2^{40}$ βήματα.