Α 44 – ΚΡΥΠΤΟΓΡΑΦΙΑ
ΣΗΜΕΙΩΣΕΙΣ #5

ΘΕΟΔΟΥΛΟΣ ΓΑΡΕΦΑΛΑΚΗΣ

1. Το πρόβλημα του διακριτού λογαρίθμου

Στο μάθημα αυτό θα δούμε κάποιους αλγόριθμους για υπολογισμό διακριτών λογάριθμων. Θυμίζουμε ότι στο πρόβλημα του διακριτού λογάριθμου (ΔΛ), μας δίνεται μια ομάδα G, $g \in G$ και κάποιο $y \in \langle g \rangle$ και θέλουμε να υπολογίσουμε το μικρότερο μη αρνητικό ακέραιο x τέτοιο ώστε $y = g^x$. Αρχικά θα περιγράψουμε δύο αλγόριθμους που μπορούν να εφαρμοστούν σε οποιαδή ομάδα G. Τέτοιους αλγόριθμους τους ονομάζουμε γενικούς. Είναι φανερό ότι το x μπορεί να βρεθεί αν υπολογίσουμε με τη σείρα τα g^k για $k = 0, 1, 2, \ldots, n-1$, και σε κάθε βήμα να συγκρίνουμε το g^k με το y. Όταν είναι ίσα έχουμε βρει το x. Ο αλγόριθμος αυτός ονομάζεται συχνά και τετριμμένος και χρειάζεται $O(n)$ πράξεις στη G. Στις επόμενες παραγράφους θα περιγράψουμε αλγόριθμους που βελτιώνουν το φράγμα αυτό.

2. Ο αλγόριθμος Baby-step/Giant-step

Ο αλγόριθμος αυτός είναι του D. Shanks. Έστω ότι μας δίνεται το $g \in G$ τάξης n (γνωστής) και το $y \in \langle g \rangle = \{g^k \mid 0 \le k \le n-1\}$. Θέλουμε να υπολογίσουμε το $0 \le x \le n-1$ τέτοιο ώστε $y = g^x$.

Έστω $1 < q < n$ ένας ακέραιος. Τότε γράφοντας την εξίσωση της διαίρεσης με υπόλοιπο του x (που δεν ξέρουμε) με το q έχουμε

$$x = q \cdot i + j, \quad \text{με} \;\; 0 \le i < \frac{n}{q} \;\; \text{και} \;\; 0 \le j < q.$$

Φυσικά τα i και j δεν τα ξέρουμε. Όμως ξέρουμε ότι είναι μονοσήμαντα ορισμένα. Δηλαδή αν υπολογίσουμε τα i και j έχουμε υπολογίσει το x. Αυτό θα κάνουμε. Γράφουμε

$$\begin{aligned} g^x = y &= g^{qi+j} \quad \Longleftrightarrow \\ yg^{-j} &= g^{gi}. \end{aligned}$$

Η τελευταία εξίσωση υποδεικνύει τον ακόλουθο αλγόριθμο.

(1) Υπολόγισε τα $u = g^{-1}$ και $w = g^q$.
(2) Για $j = 0, 1, \ldots, \lfloor n/q \rfloor$ υπολόγισε το yu^j και αποθήκευσε τα (yu^j, j).

(3) Για $i = 0, 1, \ldots, q-1$ υπολόγισε το w^i και για κάθε μία τιμή που υπολογίζεις, κοίτα αν το w^i είναι το πρώτο μέλος κάποιου ζεύγους που έχεις υπολογίσει στο βήμα (2).

(4) Οταν βρείς το i_0 τέτοιο ώστε $w^{i_0} = yu^{j_0}$, απάντησε $x = qi_0 + j_0$.

Η ορθότητα του αλγόριθμου είναι φανερή από όσα είπαμε παραπάνω. Πόσες πράξεις στη G κάνει ο αλγόριθμος; Έχουμε $O(\log q)$ πράξεις στο βήμα (1), q πράξεις στο βήμα (2) και το πολύ $\lfloor n/q \rfloor$ πράξεις στο βήμα (3). Συνολικά έχουμε $O(q + n/q)$ πράξεις. Θυμηθείτε ότι έχουμε ακόμη την ελευθερία να επιλέξουμε το q. Αν δούμε το $q+n/q$ σαν συνάρτηση του q, ελαχιστοποιήται για $q = \sqrt{n}$. Βέβαια το q πρέπει να είναι ακέραιος, οπότε επιλέγουμε το q να είναι ένας ακέραιος κοντά στο $\sqrt{n}$, για παράδειγμα το $\lfloor \sqrt{n} \rfloor$. Τότε ο αριθμός των βημάτων του αλγορίθμου είναι $O(\sqrt{n})$.

3. Ο αλγοριθμος των Pohlig και Hellman

Ο στόχος του αλγορίθμου των Pohlig και Hellman είναι διαφορετικός από αυτόν του Baby-step/Giant-step. Ας υποθέσουμε ότι γνωρίζουμε την ανάλυση του n σε πρώτους παράγοντες και είναι $n = p^t q^s$, όπου p και q είναι πρώτοι. Τότε αν καταφέρουμε να υπολογίσουμε το $x \mod p^t$ και q^s, δηλαδή αν μπορέσουμε να βρούμε $0 \le a < p^t$ και $0 \le b < q^s$ τέτοια ώστε

$$\begin{aligned} x &\equiv a \pmod{p^t} \\ x &\equiv b \pmod{q^s} \end{aligned}$$

τότε μπορούμε να βρούμε το x συνδοιάζοντας τα a και b με το Κινέζικο Θεώρημα Υπολοίπων. Άρα μένει να δείξουμε πώς μπορεί να υπολογιστεί, ας πούμε, το a. Προφανώς τα παραπάνω γενικεύονται στην περίπρωση που το n έχει περισσότερους πρώτους παράγοντες.

Αν γράψουμε το a στη βάση p, έχουμε

$$a = a_0 + a_1 p + \cdots + a_{t-1} p^{t-1}, \quad 0 \le a_i < p, \text{ για } i = 0, 1, \ldots t-1.$$

Θέλουμε να υπολογίσουμε τα $a_0, a_1, \ldots, a_{t-1}$. Έχουμε

$$y = g^x \Rightarrow y_0 = y^{n/p} = \left(g^{n/p}\right)^x = g_0^x \tag{1}$$

όπου $g_0 = g^{n/p}$. Βλέπουμε τώρα ότι η τάξη του g_0 είναι p και ο δ.λ. του y_0 ως προς το g_0 είναι a_0 διότι

$$g_0^x = g_0^{a_0 + a_1 p + \cdots + a_{t-1} p^{t-1} + A p^t} = g_0^{a_0}.$$

Μπορώ να υπολογίσω το a_0 με τον αλγόριθμο του Shanks σε $O(\sqrt{p})$ βήματα. Θέλω να συνεχίσω υπολογίζοντας το a_1. Υπολογίζω

$$y_1 = y^{n/p^2} = g_1^x = g_1^{a_0 + a_1 p + \cdots + a_{t-1} p^{t-1}} = g_1^{a_0 + a_1 p},$$

διότι το $g_1 = g^{n/p^2}$ έχει τάξη p^2. Άρα

$$h_1 = y_1 g_1^{-a_0} = g_1^{a_1 p} = (g_1^p)^{a_1}\,,$$

και το a_1 υπολογίζεται ως ο δ.λ. του h_1 ως προς βάση g_1^p. Και πάλι η τάξη του g_1^p είναι p και το a_1 μπορεί να υπολογιστεί με τον αλγόριθμο του Shanks σε $O(\sqrt{p})$ βήματα.

Έστω τώρα ότι έχω υπολογίσει τα $a_0, a_1, \ldots, a_{i-1}$ $(1 \le i < t)$. Θα δείξουμε πώς υπολογίζεται το a_i. Έχουμε

$$y_i = y^{n/p^{i+1}} = g_i^{a_0+\cdots+a_{i-1}p^{i-1}+a_ip^i+\cdots+a_{t-1}p^{t-1}} = g_i^{a_0+\cdots+a_{i-1}p^{i-1}+a_ip^i}$$

διότι το $g_i = g^{n/p^{i+1}}$ έχει τάξη p^{i+1}. Επομένως,

$$h_i = y_i g^{-a_0-a_1p-\cdots-a_{i-1}p^{i-1}} = \left(g_i^{p^i}\right)^{a_i}.$$

Άρα το a_i μπορεί να υπολογιστεί ως ο δ.λ. του h_i ως προς τη βάση $g_i^{p^i}$. Καθώς το $g_i^{p^i}$ έχει τάξη p, το a_i υπολογίζεται σε $O(\sqrt{p})$ βήματα.

Η πολυπλοκότητα του αλγορίθμου δεν είναι δύσκολο να βρεθεί. Για τον υπολογισμό του $x \mod p^t$, δηλαδή του a, κάναμε $O(t\sqrt{p})$. Και αυτό το κάνουμε για κάθε πρώτο που διαιρεί το n. Αν γενικά έχουμε $n = p_1^{t_1}\cdots p_r^{t_r}$ τότε το συνολικό κόστος είναι

$$O\left(\sum_{i=1}^{r} t_i\sqrt{p_i}\right).$$

Δεν είναι δύσκολο να δει κανείς ότι $t_i \le \log n/\log p_i$ οπότε $t_i = O(\log n)$. Ακόμη έχουμε ότι $r = O(\log n)$, δηλαδή δεν είναι δυνατό να έχουμε πάρα πολλούς πρώτους παράγοντες, οπότε τελικά

$$\sum_{i=1}^{r} t_i\sqrt{p_i} = O\left(\sqrt{p}\log^2 n\right),$$

όπου p είναι ο μέγιστος πρώτος παράγοντας του n.

Το δίδαγμα είναι ότι ο υπολογισμός ενός δ.λ. σε μια κυκλική ομάδα τάξης n μπορεί πάντα να αναχθεί στον υπολογισμό δ.λ. σε ομάδες πρώτης τάξης. Αν θέλουμε το πρόβλημα να είναι δύσκολο, πρέπει να σιγουρευτούμε ότι το n διαιρείται από κάποιο μεγάλο πρώτο. Για παράδειγμα, αν ο μόνος διαθέσιμος αλγόριθμος είναι αυτός των Pohlig και Hellman, και θέλουμε ο υπολογισμός δ.λ. να απαιτεί περίπου 2^{100} πράξεις στην ομάδα, τότε πρέπει να επιλέξουμε μια ομάδα τάξης n, όπου το n διαιρείται από κάποιο πρώτο μεγέθους περίπου 2^{200} (δηλαδή να έχει πρώτο διαιρέτη με 200 περίπου bits).

Παράδειγμα 3.1. Ας πούμε για παράδειγμα, ότι $p = 2^t + 1$ είναι πρώτος και θέλουμε να υπολογίζουμε διακριτούς λογάριθμους στην ομάδα $\mathbb{F}_p^\times$. Η τάξη της ομάδας είναι $n = p - 1 = 2^t$, που μπορεί να παραγοντοποιηθεί πολυ εύκολα.

Χρησιμοποιόντας τον αλγόριθμο των Pohlig και Hellman μπορούμε να λύσουμε το πρόβλημα σε χρόνο $O(\sqrt{2}\log^2 n) = O(\log^2 p)$. Δηλαδή το πρόβλημα λύνεται σε πολυωνυμικό χρόνο. Αν θέλουμε να βασίσουμε ένα σύστημα ElGamal στην ομάδα $\mathbb{F}_p^\times$ πρέπει να επιλέξουμε το p έτσι ώστε το $p-1$ να έχει μεγάλο πρώτο διαιρέτη.

Παράδειγμα 3.2. Ας δούμε και ένα παράδειγμα υπολογισμού διακριτού λογάριθμου με τη μέθοδο Pohlig-Hellman. Ας είναι η ομάδα μας η $\mathbb{F}_{29}^\times$ και μας δίνονται τα $y = 10$ και $g = 3$. Βλέπουμε ότι η τάξη της ομάδας είναι $n = 29 - 1 = 28 = 2^2 \cdot 7$ και η τάξη του g είναι 28, δηλαδή $\langle g \rangle = \mathbb{F}_{29}^\times$. Θέλουμε να βρούμε $0 \le x \le 28$ τέτοιο ώστε $y = g^x$ δηλαδή $10 \equiv 3^x \pmod{29}$. Το x υπολογίζεται $\mod 28$.

Σύμφωνα με τον αλγόριθμο, θα υπολογίσω το $x \mod 2^2$ και $\mod 7$. Δηλαδή, θα υπολογίσω a και b τέτοια ώστε

$$\begin{aligned} x &\equiv a \pmod{2^2} \\ x &\equiv b \pmod{7}. \end{aligned}$$

Αρχικά υπολογίζω το a. Το γράφω στη βάση 2, $a = a_0 + a_1 2$, με $0 \le a_0, a_1 \le 1$. Για το a_0 υπολογίζω:

$$y_0 = y^{n/2} = 10^{14} = 28 \mod 29,$$

$$g_0 = g^{n/2} = 3^{14} = 28 \mod 29.$$

Και αφού, όπως είδαμε, $y_0 = g_0^{a_0}$ δηλαδή $28 \equiv 28^{a_0} \pmod{29}$ βλέπουμε ότι $a_0 = 1$. Στη συνέχεια, υπολογίζω:

$$y_1 = y^{n/4} = 10^7 = 17 \mod 29,$$

$$g_1 = g^{n/4} = 3^7 = 12 \mod 29.$$

Άρα γράφω

$$h_1 = y_1 g_1^{-a_0} = 17 \cdot 12^{-1} = 17 \cdot 12^{4-1} = 28 \mod 29.$$

Στην παραπάνω γραμμή χρησιμοποίησα το ότι το $g_1 = 12$ έχει τάξη 4, επομένως $g_1^{-1} = g_1^{4-1} = 12^{4-1} \mod 29$. Δεδομένου ότι $h_1 = g_1^{2a_1} = (g_1^2)^{a_1}$, βρίσκω $g_1^2 = 28 \mod 29$ και έχω

$$28 \equiv 28^{a_1} \pmod{29},$$

οπότε βρίσκω $a_1 = 1$.

Έτσι έχω υπολογίσει $a = 1 + 1 \cdot 2 = 3$.

Προχωρώ στον υπολογισμό του b. Καθώς στην ανάλυση $28 = 2^2 \cdot 7$ το 7 εμφανίζεται με εκθέτη 1, αρκει να υπολογίσω ένα μόνο ψηφίο, το ίδιο το b. Γράφω

$$y_0 = y^{n/7} = 10^4 = 24 \mod 29,$$

$$g_0 = g^{n/7} = 3^4 = 23 \mod 29.$$

Έτσι έχουμε $y_0 = g_0^b$, δηλαδή

$$24 \equiv 12^b \pmod{29}.$$

Υπάρχουν 7 δυνατές επιλογές για το b (οι $b = 0, 1, \ldots, 6$), τις οποίες μπορώ να εξετάσω μια προς μία. Φυσικά θα μπορούσα να εφαρμόσω τον αλγόριθμο Baby-step/Giant-step. Σε κάθε περίπτωση, βρίσκω $b = 6$.

Έτσι έχω να λύσω το σύστημα

$$\begin{aligned} x &\equiv 3 \pmod{4} \\ x &\equiv 6 \pmod{7}. \end{aligned}$$

Με τον αλγόριθμο για το Κινέζικο Θεώρημα Υπολοίπων, βρίσκω $x = 27$. Πραγματικά, μπορεί κανείς εύκολα να επαληθεύσει ότι $3^{27} = 10 \mod 29$.