Α44 – ΚΡΥΠΤΟΓΡΑΦΙΑ
ΣΗΜΕΙΩΣΕΙΣ #12

ΘΕΟΔΟΥΛΟΣ ΓΑΡΕΦΑΛΑΚΗΣ

1. Πλεγματα

Έστω ο διανυσματικός χώρος $\mathbb{R}^d$ διάστασης d. Ο χώρος $\mathbb{R}^d$ έρχεται με ένα εσωτερικό γινόμενο $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d x_i y_i$ και τη σχετική νόρμα $|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = (\sum_{i=1}^d x_i^2)^{1/2}$. Έστω $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_d\} \subset \mathbb{R}^d$, d γραμμικώς ανεξάρτητα διανύσματα. Το σύνολο

$$\mathcal{L}(A) = \{m_1 \mathbf{a}_1 + \ldots + m_d \mathbf{a}_d \ : \ m_i \in \mathbb{Z}\}$$

ονομάζεται (πλήρες) πλέγμα διάστασης d. Το σύνολο A ονομάζεται βάση του πλέγματος. Όπως ισχύει και για το διανυσματικό χώρο $\mathbb{R}^d$, είναι φανερό ότι το $\mathcal{L}$ έχει περισσότερες από μία βάσεις. Επίσεις το σύνολο $\mathcal{L}$ είναι ομάδα με την πρόσθεση του διανυσματικού χώρου.

Ένας βολικός τρόπος για να περιγράψουμε ένα πλέγμα είναι δίνοντας μια βάση του. Συχνά αυτό το κάνουμε με ένα πίνακα του οποίου οι γραμμές είναι τα διανύσματα βάσης. Για παράδειγμα, ο πίνακας

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1d} \\ a_{21} & a_{22} & \ldots & a_{2d} \\ \vdots & \vdots & & \vdots \\ a_{d1} & a_{d2} & \ldots & a_{dd} \end{pmatrix}$$

ορίζει το πλέγμα που παράγεται από τα διανύσματα $\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{id}), \ \ i = 1, \ldots, d$. Θα συμβολίζουμε τον πίνακα της βάσης A επίσεις με A.

Έστω τώρα $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\}$ μία δεύτερη βάση του ίδιου πλέγματος, δηλαδή $\mathcal{L}(A) = \mathcal{L}(B) = \mathcal{L}$. Τότε $\mathbf{a}_i \in \mathcal{L}, \ i = 1, \ldots, d$ οπότε υπάρχουν ακέραιοι $m_{ik}, \ i, k = 1, \ldots, d$ τέτοιοι ώστε

$$\mathbf{a}_i = \sum_{k=1}^d m_{ik} \mathbf{b}_k, \ \ i = 1, \ldots, d.$$

Οι παραπάνω εξισωσεις γράφονται και

$$a_{ij} = \sum_{k=1}^d m_{ik} b_{kj}, \ \ i = 1, \ldots, d,$$

δηλαδή σε γλώσσα πινάκων

$$A = M \cdot B,$$

όπου $M = (m_{ij})_{i,j=1,\ldots,d}$ είναι $d \times d$ πίνακας με στοιχεία ακεραίους.

Με τον ίδιο ακριβώς συλλογισμό, αντιστρέφοντας τους ρόλους των $\mathbf{a}_i$ και $\mathbf{b}_i$, παίρνουμε τη σχέση

$$B = N \cdot A,$$

όπου N είναι ένας $d \times d$ πίνακας με στοιχεία ακεραίους. Έτσι έχουμε

$$A = MNA$$

και παίρνοντας ορίζουσες έχουμε

$$\det(A) = \det(M)\det(N)\det(A).$$

Αφού τα $\mathbf{a}_i$ είναι γραμμικώς ανεξάρτητα, $\det(A) \neq 0$, και έτσι έχουμε

$$\det(M)\det(N) = 1.$$

Όμως οι πίνακες M, N έχουν στοιχεία ακεραίους, οπότε οι ορίζουσες είναι ακέραιοι. Συμπεραίνουμε ότι

$$\det(M), \det(N) = \pm 1.$$

Είδαμε ότι δύο οποιεσδήποτε βάσεις A, B ενός πλέγματος συνδέονται με πίνακες με στοιχεία ακεραίους και ορίζουσα ± 1 και ότι $|\det(A)| = |\det(B)|$. Η ποσότητα $|\det(A)|$, που εξαρτάται μόνο από το πλέγμα και όχι από τη βάση, συμβολίζεται με $\det(\mathcal{L})$ ή με $\mathrm{vol}(\mathcal{L})$.

2. Αναγωγη LLL

Συχνά, από τις άπειρες βάσεις ενός δεδομένου πλέγματος, ξεχωρίζουν κάποιες με «καλές» ιδιότητες. Κανείς ορίζει αρχικά μια τέτοια ιδιότητα και στη συνέχεια προσπαθεί να βρει μια βάση που την ικανοποιεί. Η διαδικασία μετάβασης από την αρχική βάση στη βάση που ικανοποιεί την ιδιότητα ονομάζεται αναγωγή της βάσης.

Το 1982, οι Lenstra, Lenstra και Lovász όρισαν μια τέτοια ιδιότητα και έδωσαν ένα πολυωνυμικού χρόνου αλγόριθμο για την έβρεση της ανηγμένης βάσης. Η ιδιότητα, που δε θα ορίσουμε αυστηρά εδώ, ονομάζεται LLL-ιδιότητα και η τελική βάση LLL-ανηγμένη. Ο αλγόριθμος είναι γνωστός ως LLL. Αυτό που είναι σημαντικό για εμάς είναι ότι ο αλγόριθμος είναι πολυωνυμικός (δηλαδή γρήγορος) και ότι η τελική βάση $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\}$, για την οποία έχουμε υποθέσει ότι $|\mathbf{b}_1| \leq |\mathbf{b}_2| \leq \ldots \leq |\mathbf{b}_d|$ ικανοποιεί (αποδείξιμα) τις παρακάτω ιδιότητες:

(1) $|\mathbf{b}_1| \leq 2^{\frac{d-1}{2}} \min\{|\mathbf{v}| : \mathbf{v} \in \mathcal{L}\}$,
(2) $|\mathbf{b}_1| \leq 2^{\frac{d-1}{2}} (\det(\mathcal{L}))^{\frac{1}{d}}$,
(3) $|\mathbf{b}_2| \leq 2^{\frac{d}{2}} (\det(\mathcal{L}))^{\frac{1}{d-1}}$,
(4) $\det(\mathcal{L}) \leq \prod_{i=1}^{d} |\mathbf{b}_i| \leq 2^{\frac{d(d-1)}{2}} \det(\mathcal{L})$.

Ο Schnorr κατάφερε να μετατρέψει τον αλγόριθμο έτσι ώστε να οι σταθερά 2 να μπορεί να αντικατασταθεί με $1 + \epsilon$ για οποιοδήποτε $\epsilon > 0$ και ο αλγόριθμος να παραμένει πολυωνυμικού χρόνου. Φυσικά η πολυπλοκότητα εξαρτάται από το ϵ: όσο μικρότερο το ϵ τόσο περισσότερα βήματα κάνει ο αλγόριθμος.

3. Υπολογιστικα προβληματα σε πλεγματα

Κανείς μπορεί να ορίσει διάφορα ενδιαφέροντα υπολογιστικά προβλήματα πάνω σε πλέγματα. Τα δύο σημαντικότερα είναι τα Shortest Vector Problem (SVP) και Closest Vector Problem (CVP). Τα ορίζουμε:

SVP: Δεδομένου ενός πλέγματος $\mathcal{L}$ βρες το μη μηδενικό διάνυσμα του πλέγματος με τη μικρότερη νόρμα.

CVP: Δεδομένων ενός πλέγματος $\mathcal{L}$ και ενός διανύσματος $\mathbf{v} \in \mathbb{R}^d$ βρες το κοντινότερο, στο $\mathbf{v}$, διάνυσμα του πλέγματος.

Είναι φανερό ότι το SVP είναι ειδική περίπτωση του CVP. Και τα δύο προβλήματα αποδεικνύεται ότι είναι **NP**-πλήρη, δηλαδή υπολογιστικά δύσκολα. Έτσι ασχολούμαστε με τα αντίστοιχα προβλήματα προσέγγισης. Κανείς βλέπει αμέσως ότι ο αλγόριθμος LLL μας δίνει μια προσεγγιστική λύση για το SVP. Ο LLL σε συνδοιασμό με μια αναγωγή του CVP στο SVP δίνει μια μέθοδο για εύρεση προσεγγιστικής λύσης στο CVP. Για παράδειγμα, για το CVP μπορεί να αποδειχτεί το παρακάτω λήμμα.

Λήμμα 3.1. *Υπάρχει πολυωνυμικου χρόνου αλγόριθμος, ο οποίος δεδομένου ενός πλέγματος $\mathcal{L}$ διάστασης d και ενός διανύσματος $\mathbf{x} \in \mathbb{R}^d$ υπολογίζει ένα διάνυσμα $\mathbf{v} \in \mathcal{L}$ τέτοιο ώστε*

$$|\mathbf{v} - \mathbf{x}| \leq 2^{\frac{d-1}{4}} \min\{|\mathbf{w} - \mathbf{x}| \ : \ \mathbf{w} \in \mathcal{L}\}.$$

Είναι σημαντικό να προσθέσουμε ότι η εμπειρία έχει δείξει ότι ο αλγόριθμος LLL στην πράξη δίνει καλύτερα αποτέσματα απ' ότι περιγράψαμε παραπάνω. Συχνά ο αλγόριθμος βρίσκει το μικρότερο μη μηδενικό διάνυσμα του πλέγματος.

Παράδειγμα 3.2. Ας πάρουμε το πλέγμα που παράγεται από τις γραμμές του πίνακα

$$A = \begin{pmatrix} 1 & 2 & 5 \\ 1 & 3 & 3 \\ 0 & 2 & 7 \end{pmatrix}.$$

Βλέπουμε για παράδειγμα ότι το διάνυσμα $\mathbf{v} = 2\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_3 = (1, 5, 6)$ ανήκει στο πλέγμα. Η ορίζουσα του πλέγματος είναι $\det(\mathcal{L}) = |\det(A)| = 11$. Ο αλγόριθμος LLL με είσοδο τη βάση του πίνακα A δίνει αποτέλεσμα τη βάση που περιγράφεται από τις γραμμές του πίνακα

$$B = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & -2 \\ 2 & 3 & 1 \end{pmatrix}.$$

Βλέπουμε ότι το διάνυσμα $(-1, 1, 0)$ είναι ένα διάνυσμα του πλέγματος με πολύ μικρή νόρμα. Ελέγξτε αν είναι ένα από τα διανύσματα που πλέγμτος με τη μικρότερη νόρμα. Ακόμη παρατηρήστε ότι

$$|\mathbf{b}_1| \cdot |\mathbf{b}_2| \cdot |\mathbf{b}_3| = \sqrt{2}\sqrt{5}\sqrt{14} \approx 11.83,$$

που είναι πολύ κοντά στο $\det(\mathcal{L}) = 11$.

4. Εφαρμογη σε κρυπτοσυστηματα σακιδιου

Ας δουμε τώρα πώς τα παραπάνω μπορουν να χρησιμοποιηθούν για την κρυπτανάλυση του κρυπτοσυστήματος των Merkle-Hellman. Σε ό,τι ακολουθεί θα υοθέσουμε ότι η το M που επιλέγει ο κατασκευαστής του συστήματος είναι γνωστό σε όλους. Αυτό δεν είναι πολύ περιοριστική υπόθεση, κάνει όμως την περιγραφή της κρυπτανάλυσης αρκετά καθαρότερη.

Θυμίζουμε ότι το ιδιωτικό κλειδί του χρήστη αποτελείται από τα $s_1, \ldots, s_n, W$ και το δημόσιο κλειδί από τα $a_1, \ldots, a_n$. Το M είναι παράμετρος γνωστή σε όλους. Θα δείξουμε πώς από τα δημόσια δεδομένα μπορεί κανείς να υπολογίσει το W, και συνεπώς να βρει τα $s_1, \ldots, s_n$, δηλαδή να ανακτήσει το ιδιωτικό κλειδί.

Παρατηρήστε ότι το πλέγμα που γεννάται από τις γραμμές του $(n+1)\times(n+1)$ πίνακα

$$A = \begin{pmatrix} M & 0 & \ldots & 0 & 0 \\ 0 & M & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & M & 0 \\ a_1 & a_2 & \ldots & a_n & 1/M \end{pmatrix}$$

περιλαμβάνει διανύσματα που γενικά έχουν τη μορφή

$$\mathbf{v}_P = (Pa_1 - k_1 M, Pa_2 - k_2 M, \ldots, Pa_n - k_n M, P/M).$$

Τα πρώτα n διανύσματα έχουν νόρμα ίση με M. Το τελευταίο διάνυσμα της βάσης έχει νόρμα ίση με

$$\left(\sum_{i=1}^{n} a_i^2 + \frac{1}{M^2} \right)^{\frac{1}{2}} \approx c\sqrt{n} M$$

όπου c είναι κάποια σταθερά, αν υποθέσουμε ότι τα a_i είναι της τάξης του M.

Από την κατασκευή του κρυπτοσυστήματος, ξέρουμε ότι $Wa_i \equiv s_i \pmod{M}$ που σημαίνει ότι υπάρχουν ακέραιοι $k_i, i = 1, \ldots n$ τέτοιοι ώστε

$$Wa_i - k_i M = s_i, \quad i = 1, \ldots, n.$$

Βλέπουμε τώρα ότι από τα διανύσματα του πλέγματος ξεχωρίζει το

$$\mathbf{v}_W = (Wa_1 - k_1 M, \ldots, Wa_n - k_n M, W/M) = (s_1, \ldots, s_n, W/M),$$

το οποίο έχει αρκετά μικρή νόρμα σε σχέση με τα αρχικά διανύσματα:

$$|\mathbf{v}_W| = \left(\sum_{i=1}^{n} s_i^2 + \frac{W^2}{M^2} \right)^{\frac{1}{2}}.$$

Από το γεγονός ότι η ακολουθία $s_i, i = 1, \ldots, n$ είναι υπεραύξουσα και $M > \sum_{i=1}^{n} s_i$ κανείς περιμένει ότι τα s_i είναι αρκετά μικρότερα από το M. Αυτό είναι φανερό ιδιαίτερα για τις αρχικές τιμές της ακολουθίας. Έτσι κανείς περιμένει ότι το διάνυσμα $\mathbf{v}_W$ έχει ιδιαίτερα μικρή νόρμα και μπορεί να βρεθεί με χρήση του LLL. Προσέξτε ότι ακόμη και αν οι τελικές τιμές της ακολουθίας s_i, ας πούμε για παράδειγμα το s_n, είναι πολύ κοντά στο M, πράγμα που θα καθιστούσε τη

νόρμα του $\mathbf{v}_W$ μεγάλη, μπορούμε να θεωρήσουμε το πλέγμα διάστασης n που παράγεται από τις γραμμές του πίνακα

$$\begin{pmatrix} M & 0 & \dots & 0 & 0 \\ 0 & M & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & M & 0 \\ a_1 & a_2 & \dots & a_{n-1} & 1/M \end{pmatrix}$$

και πάλι να βρούμε το W και να ανακτήσουμε τα s_i. Προσέξτε ότι το διάνυσμα $(0, 0, \dots, 0, 1)$ ανοίκει πάντα στο πλέγμα. Άρα περιμένουμε ο LLL να το βρει και να είναι το μικρότερο διάνυσμα της νέας βάσης. Το $\mathbf{v}_W$ περιμένουμε να είναι το δεύτερο μικρότερο διάνυσμα της νέας βάσης.

Παράδειγμα 4.1. Ας δούμε πώς μπορούμε να βρούμε το ιδιωτικό κλειδί του παραδείγματος 2.1 του προηγούμενου μαθήματος. Θυμίζουμε ότι το δημόσιο κλειδί περιλάμβανε τα $a_1 = 172, a_2 = 117, a_3 = 69, a_4 = 138, a_5 = 104, a_6 = 153, a_7 = 210$ και θεωρούμε γνωστή την παράμετρο $M = 227$. Ορίζουμε το πλέγμα που παράγεται από τον πίνακα

$$A = \begin{pmatrix} 227 & 0 & 0 & 0 & 0 & 0 \\ 0 & 227 & 0 & 0 & 0 & 0 \\ 0 & 0 & 227 & 0 & 0 & 0 \\ 0 & 0 & 0 & 227 & 0 & 0 \\ 0 & 0 & 0 & 0 & 227 & 0 \\ 172 & 117 & 69 & 138 & 104 & 1/227 \end{pmatrix}.$$

Η ανηγμένη βάση που μας δίνει ο LLL δίνεται από τις γραμμές του πίνακα

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 2 & 7 & 14 & 27 & 33/227 \\ 33 & 66 & 4 & 8 & -17 & -46/227 \\ -27 & -54 & 38 & 76 & -48 & 17/227 \\ -100 & 27 & -19 & 38 & 24 & 105/227 \\ -17 & -34 & 108 & -11 & -5 & -107/227 \end{pmatrix}.$$

Όπως βλέπουμε, το δεύτερο διάνυσμα μας δίνει τα s_1, s_2, s_3, s_4, s_5 και το W που μπορούμε να το χρησιμοποιήσουμε για να βρούμε και τις υπόλοιπες τιμές s_6 και s_7.