Α31 ΚΡΥΠΤΟΓΡΑΦΙΑ
Σημειώσεις

Θεόδουλος Γαρεφαλάκης

16 Μαΐου 2022

1 Το πρόβλημα του σακιδίου

Το πρόβλημα του σακιδίου είναι το εξης: Δεδομένων θετικών πραγματικών (βαρών) $a_1, \ldots, a_n$ και θετικού πραγματικού (χωρητικότητας) E, ζητήται να υπολογιστεί ο μέγιστος αριθμός αντικειμένων που των οποίων το συνολικό βάρος είναι το πολύ E. Κανείς μπορεί να ορίσει διάφορες παραλλαγές του προβλήματος. Η παραλλαγή που μας ενδιαφέρει στην κρυπογραφία είναι η εξής: Δεδομένων των βαρών $a_1, \ldots, a_n$ και της χωρητικότητας E ζητήται να αποφασιστεί αν υπάρχει υποσύνολο των αντικειμένων (βαρών) που να έχει άθροισμα βαρών ακριβώς E και αν υπάρχει να βρεθεί. Στο κρυπτοσύστημα που θα περιγράψουμε παρακάτω θα κάνουμε χρήση αυτής της παραλλαγής.

Παράδειγμα 1. Έστω ότι μας δίνονται τα βάρη $1, 2, 7, 12, 21, 28$ και η χωρητικότητα 20. Θέλουμε να αποφασίσουμε αν υπάρχει υποσύνολο βαρών με άθροισμα 20 και αν υπάρχει να το βρούμε. Βλέπουμε ότι το υποσύνολο $1, 7, 12$ έχει το ζητούμενο άθροισμα. Αν η χωρητικότητα ήταν 11 δεν θα υπήρχε τέτοιο υποσύνολο.

Το πρόβλημα μπορεί να διατυπωθεί και ως εξης: Δεδομένων θετικών πραγματικών $a_1, \ldots, a_n$, ζητήται να βρεθούν, αν υπάρχουν, αριθμοί $x_1, \ldots, x_n \in \{0, 1\}$ έτσι ώστε $\sum_{i=1}^{n} a_i x_i = E$. Είναι γνωστό ότι το πρόβλημα του σακιδίου είναι **NP**-πλήρες, που πρακτικά σημαίνει ότι για κάθε αλγόριθμο που επιλύει το πρόβλημα υπάρχουν είσοδοι οι οποίες απαιτούν από τον αλγόριθμο υπερπολυωνυμικό χρόνο (απ" όσο γνωρίζουμε).

2 Το κρυπτοσύστημα των Merkle-Hellman

Το 1978, οι Merkle και Hellman πρότειναν ένα σύστημα κρυπτογράφησης δημόσιου κλειδιού που βασίζεται στο πρόβλημα του σακιδίου. Η βασική ιδέα είναι το δημόσιο κλειδί να είναι ένα σύνολο βαρών $\{a_1, \ldots, a_n\}$. Για να κρυπτογραφήσει κανείς ένα μύνημα, το γράφει πάνω στο αλφάβητο $\{0, 1\}$ και το κρυπτογράφημα του μηνύματος $(x_1, \ldots, x_n) \in \{0, 1\}^n$ είναι το

$$E = \sum_{i=1}^{n} a_i x_i. \tag{1}$$

Εφόσον το πρόβλημα του σακιδίου είναι δύσκολο, τότε είναι δύσκολο για κάποιον επιτιθέμενο να βρει το μήνυμα δεδομένου του κρυπτογράμματος E και του δημόσιου κλειδιού. Βέβαια, το ίδιο δύσκολο είναι να αποκρυπτογραφήσει και ο νόμιμος παραλήπτης! Επιπλέον, δεν είναι καθόλου φανερό ότι η n-άδα $(x_1, \ldots, x_n)$ που ικανοποιεί την Εξ.(1) είναι μοναδική. Για να λύσουν τα δύο αυτά προβλήματα, οι Merkle και Hellman πρότειναν να χρησιμοποιούνται βάρη με ειδική δομή, συγκεκριμμένα υπεραύξουσες ακολουθίες. Επιλέγονται αριθμοί $s_1, \ldots, s_n$ που ικανοποιούν

$$s_{j+1} > \sum_{i=1}^{j} s_i, \quad 1 \leq j \leq n-1.$$

Αν η κρυπτογράφηση γίνει με τέτοια βάρη, τότε η αποκρυπτογράφηση είναι μοναδική και γίνεται πολυ γρήγορα. Για τη μοναδικότητα, βλέπουμε ότι αν υπήρχαν δύο μηνύματα $(x_1, \ldots, x_n)$ και $(y_1, \ldots, y_n)$ που αντιστοιχούν στο ίδιο κρυπτογράφημα E, τότε

$$E = \sum_{i=1}^{n} s_i x_i = \sum_{i=1}^{n} s_i y_i$$

οπότε

$$\sum_{i=1}^{n} (x_i - y_i) s_i = 0.$$

Αν $(x_1, \ldots, x_n) \neq (y_1, \ldots, y_n)$ τότε οι δύο n-άδες διαφέρουν σε μία τουλάχιστον συντεταγμένη. Ας είναι j ο μεγαλύτερος δείκτης για τον οποίο $x_i \neq y_i$, $1 \leq i \leq n$. Τότε

$$(y_j - x_j)s_j = \sum_{i=1}^{j-1}(x_i - y_i)s_i$$

οπότε

$$s_j = |y_j - x_j||s_j| = \left|\sum_{i=1}^{j-1}(x_i - y_i)s_i\right| \leq \sum_{i=1}^{j-1}|x_i - y_i|s_i \leq \sum_{i=1}^{j-1} s_i < s_j$$

που είναι αντίφαση. Άρα για υπεραύξουσες ακολουθίες η αποκρυπτογράφηση είναι μοναδική.

Ένας αλγόριθμος αποκρυπτογράφησης είναι ο παρακάτω. Τα δεδομένα είναι η ακολουθία $s_1, \ldots, s_n$ και το κρυπτογράφημα E.

```
i = n
while i > 1 and E > 0 do
  if E >= s[i]
    x[i] = 1
    E = E - s[i]
  else
    x[i] = 0
  end
  i = i - 1
end
```

Με την επιλογή υπεραυξουσών ακολουθιών βαρών έχουμε λύσει το πρόβλημα της αποκρυπτογράφησης για τον νόμιμο παραλήπτη: η αποκρυπτογράφηση είναι μοναδική και γίνεται εύκολα. Παραμένει βέβαια το πρόβλημα ότι με τον ίδιο απλό τρόπο μπορεί να αποκρυπτογραφήσει ο καθένας. Αυτό ήταν αναμενόμενο, καθώς δεν έχουμε πει τίποτα για ιδιωτικό κλειδί, κάποια πληροφορία που έχει μόνο ο νόμιμος παραλήπτης και που τον βοηθά να αποκρυπτογραφεί. Η ιδέα των Merkle-Hellman ήταν να επιλέξουν δύο ακεραίους W, N, με $N > \sum_{i=1}^{n} s_i$ και $(W, N) = 1$. Μετά υπολογίζουν αριθμούς $a_1, \ldots, a_n$ από τις ισοτιμίες

$$a_i \equiv W^{-1}s_i \pmod{N},\ i = 1, \ldots, n \tag{2}$$

όπου $1 \leq a_i < N$, $i = 1, \ldots n$. Τώρα έχουμε όλα τα στοιχεία για να περιγράψουμε το σύστημα.

Ιδιωτικό κλειδί: Τα W, N και τα βάρη $s_1, \ldots, s_n$.

Δημόσιο κλειδί: Τα $a_1, \ldots, a_n$.

Αλγόριθμος κρυπτογράφησης
Το κρυπτογράφημα του $(x_1, \ldots, x_n) \in \{0, 1\}^n$ είναι το $C = \sum_{i=1}^{n} a_i x_i$.

Αλγόριθμος αποκρυπτογράφησης

1. Υπολόγισε το $E \equiv WC \pmod{N}$, με $1 \leq E < N$.
2. Υπολόγισε τα $x_1, \ldots, x_n$ που ικανοποιούν την $\sum_{i=1}^{n} s_i x_i = E$.

Η ορθότητα του αλγόριθμου αποκρυπτογράφησης φαίνεται από τα εξης: Οι εξισώσεις (2) συνεπάγονται

$$W\sum_{i=1}^{n} a_i x_i \equiv \sum_{i=1}^{n} s_i x_i \pmod{N}$$

άρα

$$WC \equiv \sum_{i=1}^{n} s_i x_i \pmod{N}.$$

Καθώς το N έχει επιλεγεί ώστε $N > \sum_{i=1}^{n} s_i \geq \sum_{i=1}^{n} s_i x_i$, θα έχουμε

$$E = \sum_{i=1}^{n} s_i x_i.$$

Παράδειγμα 2. Ας κατασκευάσουμε ένα σύστημα σακιδίου. Για το ιδιωτικό κλειδί, ας πάρουμε την υπεραύξουσα ακολουθία $s_1 = 1, s_2 = 2, s_3 = 7, s_4 = 14, s_5 = 27, s_6 = 55, s_7 = 120$. Επιλέγουμε $N = 227$ και $W = 33$. Το 227 είναι πρώτος, άρα $(N, W) = 1$. Υπολογίζουμε το δημόσιο κλειδί λύνοντας τις ισοτιμίες

$$\begin{aligned}
33 \cdot a_1 \equiv 1 \pmod{227} &\iff a_1 \equiv 172 \pmod{227},\\
33 \cdot a_2 \equiv 2 \pmod{227} &\iff a_2 \equiv 117 \pmod{227},\\
33 \cdot a_3 \equiv 7 \pmod{227} &\iff a_3 \equiv 69 \pmod{227},\\
33 \cdot a_4 \equiv 14 \pmod{227} &\iff a_4 \equiv 138 \pmod{227},\\
33 \cdot a_5 \equiv 27 \pmod{227} &\iff a_5 \equiv 104 \pmod{227},\\
33 \cdot a_6 \equiv 55 \pmod{227} &\iff a_6 \equiv 153 \pmod{227},\\
33 \cdot a_7 \equiv 120 \pmod{227} &\iff a_7 \equiv 210 \pmod{227}.
\end{aligned}$$

Δημοσιοποιούμε τα $a_1 = 172, a_2 = 117, a_3 = 69, a_4 = 138, a_5 = 104, a_6 = 153, a_7 = 210$. Για να κρυπτογραφήσουμε το μήνυμα $(x_1, x_2, x_3, x_4, x_5, x_6, x_7) = (1, 1, 0, 1, 0, 1, 1)$ υπολογίζουμε το

$$C = \sum_{i=1}^{6} a_i x_i = 172 + 117 + 138 + 153 + 210 = 790.$$

Το κρυπτογράφημα του μηνύματος $(1, 1, 0, 1, 0, 1, 1)$ είναι το 790. Για να αποκρυπτογραφήσουμε το 790 υπολογίζουμε το

$$E \equiv WC \equiv 33 \cdot 790 \pmod{227} \implies E = 192.$$

Βρίσκουμε τώρα διαδοχικά τα ψηφία του αρχικού μηνύματος. $192 > 120$ άρα $x_7 = 1$ και συνεχίζουμε με το $192 - 120 = 72$. $72 > 55$ άρα $x_6 = 1$. $72 - 55 = 17 < 27$ άρα $x_5 = 0$. $17 > 14$ άρα $x_4 = 1$. $17 - 14 = 3 < 7$ οπότε $x_3 = 0$. $7 > 2$ άρα $x_2 = 1$ και $7 - 2 = 5 > 1$ άρα $x_1 = 1$.

Όσα περιγράψαμε έως τώρα, έχουμε δείξει ότι το σύστημα των Merkle-Hellman είναι λειτουργικό, δηλαδή η κρυπτογράφηση και η αποκρυπτογράφηση γίνοται εύκολα (και σωστά). Δεν έχουμε πει τίποτα για την ασφάλεια του συστήματος. Αρχικά πιστευόταν ότι το σύστημα ήταν ασφαλές, καθώς ο επιτιθέμενος ήταν αντιμέτωπος με ένα πρόβλημα που όλοι πιστεύουν ότι είναι δύσκολο. Όπως αποδείχτηκε λίγα χρόνια αργότερα, το σύστημα αυτό, όπως και όλες οι παραλλαγές του, είναι εντελώς ανασφαλή. Ο λόγος είναι ότι ο επιτιθέμενος δεν έρχεται αντιμέτωπος με οποιοαδήποτε είσοδο του προβλήματος του σακιδίου, αλλά με μια πολύ ειδική είσοδο, όπου τα βάρη προέρχονται από μια υπεραύξουσα ακολουθία.