Εθνικο Μετσοβιο Πολυτεχνειο

Σχολη Εφαρμοσμενων Μαθηματικων & Φυσικων Επιστημων

Σημειώσεις Διαλέξεων

Στοιχεία Θεωρίας Αριθμών

&

Εφαρμογές στην Κρυπτογραφία

Επιμέλεια σημειώσεων:
Κωστής Γκιωνης

Διδάσκοντες:
Στάθης Ζαχος
Άρης Παγουρτζης

9 Δεκεμβρίου 2011

1 Πιθανές Επιθέσεις στο RSA

Υπενθύμιση RSA

Η Alice θέλει να στείλει ένα μήνυμα στον Bob χρησιμοποιώντας το σύστημα κρυπτογράφησης RSA. Η διαδικασία που ακολουθεί είναι η ακόλουθη:

1. Ο Bob επιλέγει δυο διαφορετικούς πρώτους αριθμούς $p, q \in \mathbb{N}$ με τυχαίο τρόπο και υπολογίζει τον αριθμό $n = p \cdot q$.
2. Διαλέγει $e \in \mathbb{N} : gcd(e, \phi(n)) = 1$, όπου $\phi(n)$ είναι η συνάρτηση Euler και επειδή οι p, q είναι πρώτοι ισχύει πως $\phi(n) = (p-1)(q-1)$.
3. Τέλος αξιοποιώντας τον Επεκταμένο Αλγόριθμο του Ευκλείδη υπολογίζει τον αριθμό d που είναι τέτοιος ώστε $e \cdot d \equiv 1 \,(\mod \phi(n))$. Κοινώς βρίσκει τον αντίστροφο του e στο $\phi(n)$.
4. Στέλνει στην Alice το δημόσιο κλειδί του, δηλαδή τους αριθμούς n και e ($Pu_k(n, e)$). Το ιδιωτικό κλειδί του Bob είναι $Pr_k(p, q, d)$
5. Η Alice κρυπτογραφεί το μήνυμα m υπολογίζοντας το αντίστοιχο κρυπτομήνυμα C με τη συνάρτηση κρυπτογράφησης

$$E(m) = m^e \mod n$$

και το στέλνει στον Bob.

6. Ο Bob αποκρυπτογραφεί το C με τη συνάρτηση αποκρυπτογράφησης

$$D(C) = C^d \mod n$$

και διαβάζει το μύνημα m.

Άσκηση 1. *Υποθέτουμε πως μία εταιρεία (τα παλιά χρόνια) είχε ζητήσει από τους υπαλλήλους της να επιλέγουν δημόσια κλειδιά $Pu_k(n, 3)$ ώστε η κρυπτογράφηση να είναι γρήγορη και να μην σπαταλούνται οι διαθέσιμοι υπολογιστικοί εταιρικοί πόροι. Οι υπάλληλοι λοιπόν με τον περιορισμό το $gcd(3, \phi(n)) = 1$ επιλέγουν με τυχαίο τρόπο p, q ώστε να σχηματίσουν τα δημόσια κλειδιά τους.*

Η Alice, που είναι υπάλληλος της εταιρείας, επιθυμεί να στείλει στους τρεις συναδέλφους της Bob, Charlie και Diane ένα ιδιαίτερα σημαντικό μήνυμα, έστω m.

Η Eve, που δουλεύει σε μια ανταγωνιστική εταιρεία, καταφέρνει και υποκλέπτει τα κρυπτομήνυμα C_i με $i \in \{1, 2, 3\}$. Θα μπορέσει η Eve να αποκρυπτογραφήσει το μήνυμα;

Λύση. Ναι, θα μπορέσει. Καταρχάς η Eve βρίσκει τα δημόσια κλειδιά των Bob, Charlie και Diane στον κατάλογο της εταιρείας:

Όνομα	Δημόσιο Κλειδί
Bob	$(n_1, 3)$
Charlie	$(n_2, 3)$
Diane	$(n_3, 3)$

Η Eve σχηματίζει το παρακάτω σύστημα εξισώσεων:

$$\begin{cases} C_1 &= m^3 \mod n_1 \\ C_2 &= m^3 \mod n_2 \\ C_3 &= m^3 \mod n_3 \end{cases} \Rightarrow \begin{cases} m^3 &\equiv C_1 \mod n_1 \\ m^3 &\equiv C_2 \mod n_2 \\ m^3 &\equiv C_3 \mod n_3 \end{cases}$$

Αν οι p_i, q_i έχουν επιλεγεί με τυχαίο τρόπο μπορούμε να υποθέσουμε, με μεγάλη πιθανότητα, πως οι Bob, Charlie και Diane έχουν επιλέξει διαφορετικά p_i, q_i. Συνεπώς τα n_i είναι αν δυο σχετικά πρώτοι και σύμφωνα με το Κινέζικο Θεώρημα Υπολοίπων (CRT) το παραπάνω σύστημα έχει μοναδική λύση στο $\mathbb{Z}_N$, όπου $N = n_1 \cdot n_2 \cdot n_3$. Έτσι η Eve λύνει το σύστημα και λαμβάνει:

$$u \equiv m^3 \mod N$$

Επειδή η κρυπτογράφηση γίνεται modulo n_i έπεται πως $m < n_i, \forall i \Rightarrow m^3 < N \Rightarrow u = m^3$. Έτσι η Eve λαμβάνει το μύνημα $m = \sqrt[3]{u}$.

Σημειώνουμε πως η παραπάνω αδυναμία του RSA υπάρχει όποτε ο αριθμός των παραληπτών είναι μεγαλύτερος του εκθέτη και οι παραλήπτες έχουνε εκλέξει διαφορετικά p, q.

■

Παράδειγμα

Έστω πως τα δημόσια κλειδιά των παραληπτών είναι:

Όνομα	Δημόσιο Κλειδί (n,e)
Bob	$(2419, 3)$
Charlie	$(979, 3)$
Diane	$(1219, 3)$

Αναφέρουμε πως τα ιδιωτικά κλειδιά των παραληπτών είναι:

Όνομα	Ιδιωτικό Κλειδί (p,q,d)
Bob	$(41, 59, 1547)$
Charlie	$(89, 11, 587)$
Diane	$(53, 23, 763)$

Η Alice κρυπτογραφεί το μήνυμα $m = 327$ με τη συνάρτηση κρυπτογράφησης και λαμβάνει:

$$\begin{aligned} c_1 &= 327^3 \mod 2419 = 1557 \\ c_2 &= 327^3 \mod 979 = 798 \\ c_3 &= 327^3 \mod 1219 = 1206 \end{aligned}$$

Η Eve υποκλέπτει τα μηνύματα και σχηματίζει το σύστημα:

$$\begin{aligned} u &= C_1 \mod n_1 \\ u &= C_2 \mod n_2 \\ u &= C_3 \mod n_3 \end{aligned}$$

Με το CRT βρίσκει $u = 34965783$ και στη συνέχεια υπολογίζει το $m = \sqrt[3]{u} = 327$.

Άσκηση 2. *Η ανταλλαγή μηνυμάτων πραγματοποιείται με το κρυπτοσύστημα RSA*[1]*. Με κάποιο τρόπο έχει διέρρευσε το μυστικό κλειδί d. Το δημόσιο κλειδί είναι (n, e) (υποθέτουμε πως το e είναι μικρό). Να δείξετε πως μπορούμε να παραγοντοποιήσουμε το n, δηλαδή να βρούμε πρώτους αριθμούς p, q ώστε $n = p \cdot q$.*

Απόδειξη. Γνωρίζουμε πως στο RSA ο αριθμός d επιλέγεται έτσι ώστε να ικανοποιείται η σχέση $e \cdot d \equiv 1 \pmod{\varphi(n)}$. Σημειώνουμε πως η επιλογή του είναι μοναδική. Συνεπώς θα υπάρχει $k \in \mathbb{Z}$ ώστε:

$$e \cdot d - 1 = k\varphi(n) \tag{1}$$

Το d έχει υπολογιστεί modulo $\phi(n)$ και συνεπώς $d < \phi(n) \Rightarrow e \cdot d - 1 < e \cdot \phi(n) - 1$. Άρα λόγω της (1) ισχύει πως:

$$k \cdot \phi(n) < e \cdot \phi(n) - 1$$

Από την τελευταία ανισότητα προκύπτει πως το $k \in \{1, 2, \ldots, e-1\}$. Έτσι μέσω της εξίσωσης (1) έχουμε πως:

$$\phi(n) = \frac{e \cdot d - 1}{k}$$

και επειδή πρέπει το $\phi(n) < n$ με $O(e)$ ελέγχους προσδιορίζουμε το $\phi(n)$.

Αφού προσδιορίσαμε το $\phi(n)$ η παραγοντοποίηση του n είναι εύκολη. Γνωρίζουμε τις σχέσεις:

$$\phi(n) = (p-1)(q-1) \tag{2}$$
$$n = p \cdot q \tag{3}$$

Από την (2) λαμβάνουμε:

$$\begin{aligned} \phi(n) &= pq - p - q + 1 \Rightarrow \\ \phi(n) &\overset{(3)}{=} n - p - q + 1 \Rightarrow \\ p + q &= n + 1 - \phi(n) \Rightarrow \end{aligned}$$

[1] Η άσκηση αυτή μοιάζει με την επίθεση κοινού γινομένου.

Επίσης έχουμε πως:

$$\begin{aligned} p - q &= \sqrt{p^2 - 2pq + q^2} \\ &= \sqrt{(p+q)^2 - 4pq} \\ &\overset{(3)}{=} \sqrt{(p+q)^2 - 4n} \end{aligned}$$

Έτσι καταλήγουμε στο γραμμικό σύστημα δυο εξισώσεων με δύο αγνώστους:

$$\begin{aligned} p + q &= n + 1 - \phi(n) \\ p - q &= \sqrt{(n + 1 - \phi(n))^2 - 4n} \end{aligned}$$

από όπου προσδιορίζουμε τα p, q πολύ εύκολα.

■

Παράδειγμα για $e = 3$

Έστω πως το ιδιοτικό κλειδί είναι:

$$(p, q, d) = (41, 59, 1547)$$

το δημόσιο είναι:

$$(n, e) = (2419, 3)$$

και έχει διαρεύσει το d. Θα παραγοντοποιήσουμε το n γνωρίζοντας τα:

$$(n, e, d) = (2419, 3, 1547)$$

Αφού το $k \in 1, 2$ ελέγχουμε:

1. Αν $= 1$ τότε:

$$\phi(n) = \frac{e \cdot d - 1}{k} \quad = \frac{3 \cdot 1547}{1} = \frac{4640}{1} = 4640 > n \quad \text{(αδύνατο)}$$

2. Αν $= 2$ τότε:

$$\phi(n) = \frac{4640}{2} = 2320 < n \quad \checkmark$$

Έτσι:

$$p + q = n + 1 - \phi(n) = 2419 + 1 - 2320 = 100$$
$$p - q = \sqrt{(n + 1 - \phi(n))^2 - 4n} = \sqrt{100^2 - 4 \cdot 2419} = 18$$

και από τις παραπάνω βρίσκουμε πως $p = 41$ και $q = 59$.

■

Άσκηση 3 (chosen-ciphertext attack)**.** *Η Eve θέλει να αποκρυπτογραφήσει το κρυπτοκείμενο $c \in \mathbb{C}$ της εταιρείας NTUA. Η Eve έχει λαδώσει έναν υπάλληλο της NTUA, ο οποίος είναι διατεθειμένος να αποκρυπτογραφήσει οποιοδήποτε κρυπτοκείμενο του στείλει η Eve εκτός από αυτά που έχει αποστείλει η NTUA (Adaptive CCA, βλ. Πίνακα 1). Ο υπάλληλος θεωρεί πως με τη συμπεριφορά του δεν δημιουργεί κανένα ρήγμα στην ασφάλεια του συστήματος. Έχει δίκιο ή η Eve θα μπορέσει να αποκρυπτογραφήσει το c και να λάβει το μήνυμα m;*

Non-Adaptive CCA	Adaptive CCA
Ο Κρυπταναλυτής έχει καταφέρει να έχει στη διάθεση του κάποια ζεύγη κειμένων - κρυπτοκειμένων.	Ο Κρυπταναλυτής μπορεί να ζητά την αποκρυπτογράφηση οποιουδήποτε κειμένο εκτώς του κρυπτοκειμένου που θέλει να αποκρυπτογραφησει.

Πίνακας 1: Είδη επιθέσεων τύπου chosen-cyphertext attack.

Απόδειξη. Η Eve θα εκμεταλλευτεί τη πολλαπλασιαστική ιδιότητα του RSA. Δηλαδή αν:

$$c_1 = (m_1)^e \mod n$$
$$c_2 = (m_2)^e \mod n$$

τότε:

$$(c_1 \cdot c_2) = (m_1 \cdot m_2)^e \mod n$$

Συγκεκριμένα η Eve επιλέγει ένα $r \in \mathbb{P}$ τέτοιο ώστε να γνωρίζει το αντίστροφο του modulo n r^{-1}. Στέλνει στον συνεργάτη της το μήνυμα $r^e \cdot c$ για να της το

αποκρυπτογραφήσει. Έτσι λαμβάνει:

$$\begin{aligned}(r^e \cdot c)^d &\equiv r^{e \cdot d} \cdot c^d \pmod{n} \\ &\equiv r \cdot c^d \pmod{n} \\ &\equiv r \cdot m \pmod{n}\end{aligned}$$

Πολλαπλασιάζοντας το $(r \cdot m \pmod{n})$ με r^{-1} λαμβάνει το μήνυμα.

■

2 Τάξη γινομένου στοιχείων ομάδας

Άσκηση 4. *Έστω $(G,\cdot)$ αβελιανή ομάδα με ουδέτερο στοιχείο το e. Αν $a,b\in G$, με* $\mathrm{ord}(a)=k$ *και* $\mathrm{ord}(b)=m$. *Να δείξετε πως αν* $\gcd(k,m)=1$ *τότε* $\mathrm{ord}(a\cdot b)=km$.

Απόδειξη. Επειδή η G είναι αβελιανή ισχύει η αντιμεταθετική ιδιότητα, δηλαδή $\forall a,b\in G: a\cdot b=b\cdot a$. Από το προηγούμενο προκύπτει εύκολα πως $\forall a,b\in G$ και $n\in\mathbb{Z}: (a\cdot b)^n=a^n\cdot b^n$, ιδιότητα που δεν ισχύει γενικά σε ομάδες που δεν είναι αβελιανές. Έτσι παρατηρούμε πως $(a\cdot b)^{km}=a^{km}\cdot b^{km}=(a^k)^m\cdot(b^m)^k=e^m\cdot e^k=e$. Το οποίο σημαίνει πως

$$\mathrm{ord}(a\cdot b)\,|\,(km) \tag{1}$$

Ένα δεύτερο συμπέρασμα που αφορά τις αβελιανές ομάδες είναι πως αν τα x,y είναι αντίστροφα στοιχεία τότε έχουν την ίδια τάξη. Διότι:

$$x\cdot y=e\Rightarrow\begin{cases}(x\cdot y)^{\mathrm{ord}(x)}=e\Rightarrow y^{\mathrm{ord}(x)}=e\\(x\cdot y)^{\mathrm{ord}(y)}=e\Rightarrow x^{\mathrm{ord}(y)}=e\end{cases}\Rightarrow\begin{cases}\mathrm{ord}(y)\,|\,\mathrm{ord}(x)\\\mathrm{ord}(x)\,|\,\mathrm{ord}(y)\end{cases}\Rightarrow\mathrm{ord}(x)=\mathrm{ord}(y)$$

Αν θέσουμε $u=\mathrm{ord}(a\cdot b)$, λόγω του (2) προκύπτει πως $(a\cdot b)^u=e\Rightarrow a^u\cdot b^u=e$ το οποίο λόγω του δεύτερου συμπεράσματος σημαίνει πως:

$$\mathrm{ord}(a^u)=\mathrm{ord}(b^u) \tag{2}$$

Έστω τώρα πως $x=\mathrm{ord}(a^u)$ τότε $a^{ux}=e$ όπου x ο ελάχιστος δυνατός φυσικός ώστε $\mathrm{ord}(a)|ux$. Άρα:

$$ux=\mathrm{lcm}\,(u,\mathrm{ord}(a))=\frac{u\,\mathrm{ord}(a)}{\gcd\,(u,\mathrm{ord}(a))}\Rightarrow x=\frac{\mathrm{ord}(a)}{\gcd\,(u,\mathrm{ord}(a))}$$

Συνεπώς από σχέση (2):

$$\frac{\mathrm{ord}(a)}{\gcd\,(u,\mathrm{ord}(a))}=\frac{\mathrm{ord}(b)}{\gcd\,(u,\mathrm{ord}(b))}$$

Όμως και τα δύο κλάσματα είναι ακέραιοι (αφού είναι οι τάξεις των a^u,b^u αντίστοιχα) και επομένως είναι διαιρέτες των $k=ord(a),m=ord(b)$. Αφού όμως τα k,m είναι πρώτοι μεταξύ τους ο μόνος κοινός διαιρέτης είναι η μονάδα. Επομένως:

$$\mathsf{ord}(a) = \mathsf{gcd}\,(u, \mathsf{ord}(a)) \qquad \mathsf{ord}(b) = \mathsf{gcd}\,(u, \mathsf{ord}(b))$$

οπότε:

$$k = \mathsf{ord}(a) \mid u \qquad m = \mathsf{ord}(b) \mid u$$

και επειδή k, m πρώτοι μεταξύ τους προκύπτει

$$km \mid u = \mathsf{ord}(a \cdot b)$$

Αφού λοιπόν $u \mid km$ και $km \mid u$ προκύπτει $u = km$.

■

Α΄ Χρήσιμη Haskell για την Άσκηση 1

```haskell
---------- Euclidean Extended algorithm ----------
-- David Gray:
-- Implementing Public-Key Cryptography in Haskell
--------------------------------------------------

gcde :: Integer -> Integer -> (Integer, Integer, Integer)

gcde a b =
  let
    gcd_f (r1,x1,y1) (r2,x2,y2)
      | r2 == 0 = (r1,x1,y1)
    | otherwise =
      let
      q = r1 `div` r2
      r = r1 `mod` r2
      in
        gcd_f (r2,x2,y2) (r,x1-q*x2,y1-q*y2)
    (d,x,y) = gcd_f (a,1,0) (b,0,1)
  in
    if  d < 0
    then (-d,-x,-y)
    else (d,x,y)

------------------------- Inverse modulo n ----------
inv :: Integer -> Integer -> Integer

inv a n
  | g /= 1    = error "No inverse exists"
  | otherwise   = x `mod` n
  where (g,x,_) = gcde a n

------------------------- Relatively Prime ----------
rlp :: Integer -> Integer -> Integer

rlp a b
  |g /= 1 = 0
  |g == 1 = 1
  where (g,x,_) = gcde a b

------------------- Chinese Remainder Theorem -------
crt :: [Integer] -> [Integer] -> Integer
crt a m =
    mod (sum (zipWith (*) (zipWith (*) big_m big_n) a)) prod_m
    where
      prod_m = product m
      big_m = [div prod_m i | i <- m]
      big_n = zipWith (inv) big_m m
```

Β′ Χρήσιμη σχέση για Άσκηση 4

Πρόταση Β′.1. *Έστω $m, n \in \mathbb{Z}$, ισχύει ότι:*

$$\mathsf{lcm}(m, n) = \frac{mn}{\mathsf{gcd}\,(m, n)}$$

Απόδειξη. Σύμφωνα με το Θεμελιώδες Θεώρημα της Αριθμητική κάθε $z \in \mathbb{Z}$ μπορεί να γραφεί στη μορφή:

$$z = p_1^{e_1} \cdot p_2^{e_2} \cdots p_k^{e_k}$$

όπου p_i πρώτοι και e_i δυνάμεις. Μια άλλη ισοδύναμη έκφραση για τον z θα ήταν ως το γινόμενο όλων των πρώτων με τις δυνάμεις των πρώτων που δεν συμμετέχουν στην παραπάνω έκφραση ίσες με το 0. Έτσι:

$$z = \prod_{i=1}^{\infty} p_i^{e_i}$$

Συνεπώς έχουμε:

$$m = \prod_{i=1}^{\infty} p_i^{m_i}$$

και

$$n = \prod_{i=1}^{\infty} p_i^{n_i}$$

Επίσης ισχύουν οι παρακάτω σχέσεις:

$$\mathsf{gcd}(m, n) = \prod_{i=1}^{\infty} p_i^{min(m_i, n_i)}$$

και

$$\mathsf{lcm}(m, n) = \prod_{i=1}^{\infty} p_i^{max(m_i, n_i)}$$

Έτσι τώρα εύκολα προκύπτει η ζητούμενη σχέση:

$$\begin{aligned}
\mathsf{gcd}(m,n) \cdot \mathsf{lcm}(m,n) &= \prod_{i=1}^{\infty} p_i^{min(m_i,n_i)} \prod_{i=1}^{\infty} p_i^{max(m_i,n_i)} \\
&= \prod_{i=1}^{\infty} p_i^{min(m_i,n_i)+max(m_i,n_i)} \\
&= \prod_{i=1}^{\infty} p_i^{m_i+n_i} \\
&= \prod_{i=1}^{\infty} p_i^{m_i} + \prod_{i=1}^{\infty} p_i^{n_i} \\
&= m \cdot n
\end{aligned}$$

■