Bezpieczeństwo systemów komputerowych

Szyfry asymetryczne

Marcin Peczarski

Instytut Informatyki Uniwersytetu Warszawskiego

15 listopada 2011

Na podstawie wykładu Anny Kosieradzkiej z 2009 r.

Uczestnicy gry

- Alicja ma klucz prywatny i publiczny.
- Bolek wysyła Alicji wiadomości zaszyfrowane jej kluczem publicznym.
- Alicja deszyfruje wiadomości od Bolka, używając swojego klucza prywatnego.
- Cezary, Ewa, ... próbują odszyfrować te wiadomości lub zdobyć klucz prywatny Alicji.

Najpopularniejszym szyfrem asymetrycznym jest RSA

- p, q – losowo wybrane duże liczby pierwsze
- $n = pq$ – moduł
- e – liczba względnie pierwsza z $(p-1)(q-1)$
- d – liczba wyznaczona tak, że zachodzi
  $ed \bmod (p-1)(q-1) = 1$
- (n, d) – klucz prywatny
- (n, e) – klucz publiczny
- Szyfrowanie
  $$C = M^e \bmod n$$
- Deszyfrowanie
  $$M = C^d \bmod n$$

Testy pierwszości

- Potrzebujemy efektywnie znajdować duże liczby pierwsze.
- Wielomianowy algorytm deterministyczny
  - Agrawal, Kayal, Saxena: $\tilde{O}(\log n)^{7,5}$
  - Lenstra, Pomerance: $\tilde{O}(\log n)^{6}$
- Algorytmy probabilistyczne
  - Rozumowanie *a contrario*
  - Używamy twierdzenia postaci:
    „jeśli n jest liczbą pierwszą, to $S(n)$”,
    gdzie $S(n)$ jest formułą łatwą do sprawdzenia.
  - Jeśli zachodzi $S(n)$, twierdzimy, że z dużym prawdopodobieństem n jest pierwsza.
  - Zły przykład $S(n)$:
    „liczba n jest równa 2 lub jest nieparzysta”.

Test pierwszości Fermata

- Wybieramy jako $S(n)$ małe twierdzenie Fermata: „jeśli n jest pierwsza, to $a^{n-1} \equiv 1 \pmod{n}$”.
- Losujemy a i sprawdzamy, czy powyższe zachodzi.
- Niestety istnieją liczby złożone (liczby Carmichaela), dla których zachodzi $a^{n-1} \equiv 1 \pmod{n}$ dla dowolnego a względnie pierwszego z n.

Test pierwszości Solovaya–Strassena

- Symbol Legendre'a, dla liczby pierwszej p

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & \text{jeśli } p \text{ dzieli } a, \\ 1, & \text{jeśli istnieje } b, \text{ takie że } b^2 \equiv a \pmod{p}, \\ -1, & \text{w p.p.} \end{cases}$$

- Symbol Jacobiego

$$\left(\frac{a}{n}\right) = \left(\frac{a}{p_1}\right)^{\alpha_1} \left(\frac{a}{p_2}\right)^{\alpha_2} \cdots \left(\frac{a}{p_k}\right)^{\alpha_k},$$

gdzie $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$.

- Jeśli n jest nieparzystą liczbą pierwszą, to

$$a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \pmod{n}.$$

Test pierwszości Solovaya–Strassena, cd.

- Losujemy $0 < a < n$ i jeśli

$$a^{(n-1)/2} \not\equiv \left(\frac{a}{n}\right) \pmod{n},$$

  to n jest liczbą złożoną, a a jest tego świadkiem.
- W p.p. n jest pierwsza lub złożona.
- Dla każdej liczby złożonej n prawdopodobienstwo, że losowo wybrane a jest świadkiem, wynosi co najmniej $1/2$.
- Jeśli wypróbujemy m losowych wartości a i żadna z nich nie okaże się świadkiem złożoności dla n, to z prawdopodobieństwem co najmniej $1 - 2^{-m}$ liczba n jest pierwsza.
- Potrafimy efektywnie potęgować modulo i obliczać symbol Jacobiego.

Test pierwszości Millera–Rabina

- Dla danego nieparzystego n niech $n-1 = 2^s t$, gdzie t jest nieparzyste.
- Jeśli liczba n jest pierwsza, to $a^{n-1} \equiv 1 \pmod{n}$ (małe twierdzenie Fermata).
- Jeśli liczba n jest pierwsza i $a^{2^r t} \equiv 1 \pmod{n}$ dla $0 < r \leqslant s$, to $a^{2^{r-1} t} \equiv 1 \pmod{n}$ albo $a^{2^{r-1} t} \equiv -1 \pmod{n}$.
- Losujemy $0 < a < n$ i obliczamy: $a^t \bmod n$, $a^{2t} \bmod n$, $a^{4t} \bmod n$, $\ldots$, $a^{2^{s-1} t} \bmod n$, $a^{n-1} \bmod n$.
- Jeśli ostatnia liczba na powyższej liście jest różna od 1 lub, przeglądając od końca, pierwsza napotkana liczba różna od 1 jest też różna od $n-1$, to a jest świadkiem, że n jest złożona.
- Dla każdej liczby złożonej n prawdopodobienstwo, że losowo wybrane a jest świadkiem, wynosi co najmniej $3/4$.
- Jeśli wypróbujemy m losowych wartości a i żadna z nich nie okaże się świadkiem złożoności dla n, to z prawdopodobieństwem co najmniej $1-(1/4)^m$ liczba n jest pierwsza.

Łamanie RSA

- Znajomość $\varphi(n) = (p-1)(q-1)$ jest równoważna znajomości rozkładu n na czynniki pierwsze.
- Znając p i q, łatwo obliczamy $\varphi(n)$.
- Znając n i $\varphi(n)$, rozwiązujemy układ równań:

$$\begin{aligned} \varphi(n) &= (p-1)(q-1), \\ n &= pq. \end{aligned}$$

Sposoby łamania RSA

- Faktoryzacja n
  - brutalne poszukiwanie dzielnika (ang. *brute force*)
  - metoda sita kwadratowego
  - ...
- Wykorzystanie słabości
  - zbyt bliskie p i q
  - zbyt małe d lub e
  - częściowa znajomość d
  - atak z pomiarem czasu
  - ...

Faktoryzacja za pomocą sita kwadratowego

- Każdy wie, że jeśli $n = a^2 - b^2$, to $n = (a + b)(a - b)$.
- Rozważamy kolejne a, począwszy od $\lceil \sqrt{n} \rceil$, i sprawdamy, czy $a^2 - n$ jest kwadratem.
- Jeśli znajdziemy takie całkowite b, że $b^2 = a^2 - n$, to poznamy rozkład n.
- Jak długo będziemy szukać?

Faktoryzacja za pomocą sita kwadratowego, przykład

- Niech $n = 1649$, $\lceil \sqrt{n} \rceil = 41$, mamy:
  - $41^2 - n = 32$,
  - $42^2 - n = 115$,
  - $43^2 - n = 200$,
  - $\ldots$,
  - $57^2 - n = 1600 = 40^2$.
- Trzeba szukać dość długo, pierwszy kwadrat pojawia się dla $a = 57$.
- Znaleźliśmy rozkład $1649 = (57 + 40)(57 - 40) = 97 \cdot 17$.
- Szukanie kwadratów nie jest szybsze niż próbowanie kolejnych dzielników.
- Zauważmy, że $32 \cdot 200 = 80^2$, czyli $(41 \cdot 43)^2 \equiv 80^2 \pmod{n}$.
- Co nam daje ta magiczna zależność?
- Co łączy liczby 32 i 200?

Liczby gładkie i faktoryzacja

- Liczbę nazywamy gładką (B-gładką), jeśli jej dzielniki pierwsze są małe (mniejsze niż B).
- Szukamy liczb gładkich wśród liczb $a^2 - n$, np. sitem Erastotenesa.
- Szukamy wśród nich kombinacji takich liczb, których iloczyn jest kwadratem.
- Używamy wektora wykładników, np. liczbę $504 = 2^3 3^2 7$ reprezentujemy jako wektor $(3, 2, 0, 1)$.
- Mnożeniu liczb odpowiada dodawanie wektorów.

Proste słabości RSA

- Łatwo jest rozłożyć n na czynniki, jeśli p i q są zbyt bliskie.
- Jeśli $p-1$ i $q-1$ mają zbyt małe dzielniki, to można użyć algorytmu „$p-1$" Pollarda.
- Dwie osoby używają tego samego n i różnych e_1, e_2.
  - Zwykle wtedy e_1, e_2 są względnie pierwsze.
  - Kryptoanalityk może uzyskać szyfrogramy tego samego tekstu jawnego M, czyli $C_1 = M^{e_1} \bmod n$, $C_2 = M^{e_2} \bmod n$.
  - Algorytmem Euklidesa może wyznaczyć takie r i s, że $re_1 + se_2 = 1$.
  - Można założyć, że $r < 0$, wtedy $M = (C_1^{-1})^{-r} C_2^s \bmod n$.
- Należy rozważnie używać RSA, np. szyfrowanie każdej litery osobno umożliwia złamanie za pomocą analizy częstości występowania liter.

Atak Wienera na zbyt małe d

- Przez „zbyt małe d” rozumiemy $d < \frac{1}{3}\sqrt[4]{n}$.
- Mamy $ed - k\varphi(n) = 1$, czyli

$$\left|\frac{e}{\varphi(n)} - \frac{k}{d}\right| = \frac{1}{d\varphi(n)}.$$

- Aproksymując $\varphi(n)$ przez n, można dowieść, że

$$\left|\frac{e}{n} - \frac{k}{d}\right| < \frac{1}{2d^2}.$$

- Znamy ułamek e/n, szukamy ułamka k/d.
- Takich ułamków k/d jest mało, $O(\log n)$.

Atak Wienera na zbyt małe d, cd.

- Rozwijamy e/n w ułamek łańcuchowy

$$\frac{e}{n} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}.$$

- Dowodzi się, że wszystkie k/d są reduktami tego ułamka.
- Sprawdzamy początkowe $O(\log n)$ redukty.

Atak z pomiarem czasu

- Alicja deszyfruje komunikat C otrzymany od Bolka.
- Alicja potęguje $M = C^d \bmod n$, stosując prosty algorytm

```
{d = d[k-1] d[k-2] ... d[1] d[0]}
x := C;
M := 1;
for i := 0 to k-1 do
begin
  if d[i] = 1 then M := M * x mod n;
  x := x * x mod n;
end
```

- Znamy oczywiście bit `d[0]`, szukamy bitu `d[1]`.

Atak z pomiarem czasu, cd.

- Cezary wysyła Alicji zaszyfrowane (jej kluczem publicznym) wiadomości $C_1, \ldots, C_m$.
- Cezary mierzy czasy:
  - t'_i deszyfrowania wiadomości C_i,
  - t''_i mnożenia C_i razy C_i^2.
- Jeśli bit `d[1]` jest równy 1, to ciągi (t'_i) i (t'_i) są skorelowane.
- Jeśli bit `d[1]` jest równy 0, nie widać żadnej korelacji.
- Cezary, poznawszy `d[1]`, próbuje poznać kolejne bity d.

Atak na implementację korzystającą z chińskiego twierdzenia o resztach

- Alicja podpisuje wiadomość M, czyli oblicza $C = M^d \bmod n$.
- Dla ułatwienia obliczeń Alicja korzysta z chińskiego twierdzenia o resztach:
  - $d_p = d \bmod (p-1)$, $d_q = d \bmod (q-1)$,
  - $C_p = M^{d_p} \bmod p$, $C_q = M^{d_q} \bmod q$,
  - $s_p \equiv 1 \pmod{p}$, $s_p \equiv 0 \pmod{q}$,
  - $s_q \equiv 1 \pmod{q}$, $s_q \equiv 0 \pmod{p}$,
  - $C = s_p C_p + s_q C_q$.
- Wszystko działa...

Atak na implementację korzystającą z chińskiego twierdzenia o resztach, cd.

- Przypuśćmy, że nastąpiło prekłamanie i Alicja obliczyła C_p prawidłowo, ale zamiast C_q otrzymała błędne B_q.
- W efekcie Alicja wysyła Cezaremu błędny podpis B, dla którego zachodzi

$$B^e \equiv M \pmod{p} \quad \text{i} \quad B^e \not\equiv M \pmod{q}.$$

- Wtedy Cezary oblicza

$$\mathrm{NWD}(n, B^e - M) = p.$$

Nie wolno używać tej samej pary kluczy do szyfrowania i podpisywania wiadomości

- Alicja otrzymuje zaszyfrowaną wiadomość $c = m^e \bmod n$.
- Ewa podsłuchała tę wiadomość.
- Ewa losuje liczbę $r < n$.
- Ewa oblicza:
  - $x = r^e \bmod n$,
  - $y = xc \bmod n$.
- Zauważmy, że $r = x^d \bmod n$.
- Ewa wysyła Alicji do podpisania wiadomość y.
- Alicja odsyła Ewie $u = y^d \bmod n$.
- Ewa oblicza
  $r^{-1}u \bmod n = x^{-d}y^d \bmod n = x^{-d}x^d c^d \bmod n = m$.

Krzywe eliptyczne

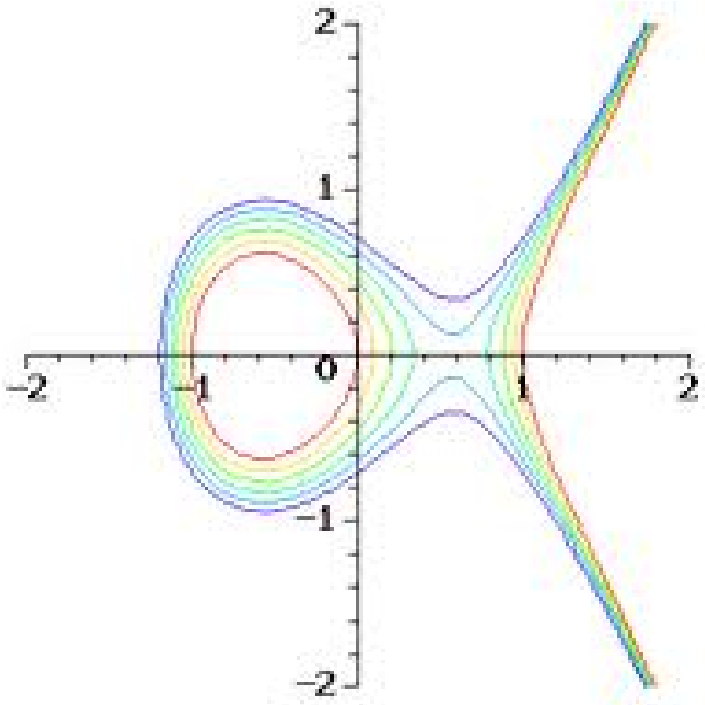
- Krzywą eliptyczną E nad ciałem K nazywamy krzywą zadaną równaniem postaci

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

o współczynnikach z ciała K.

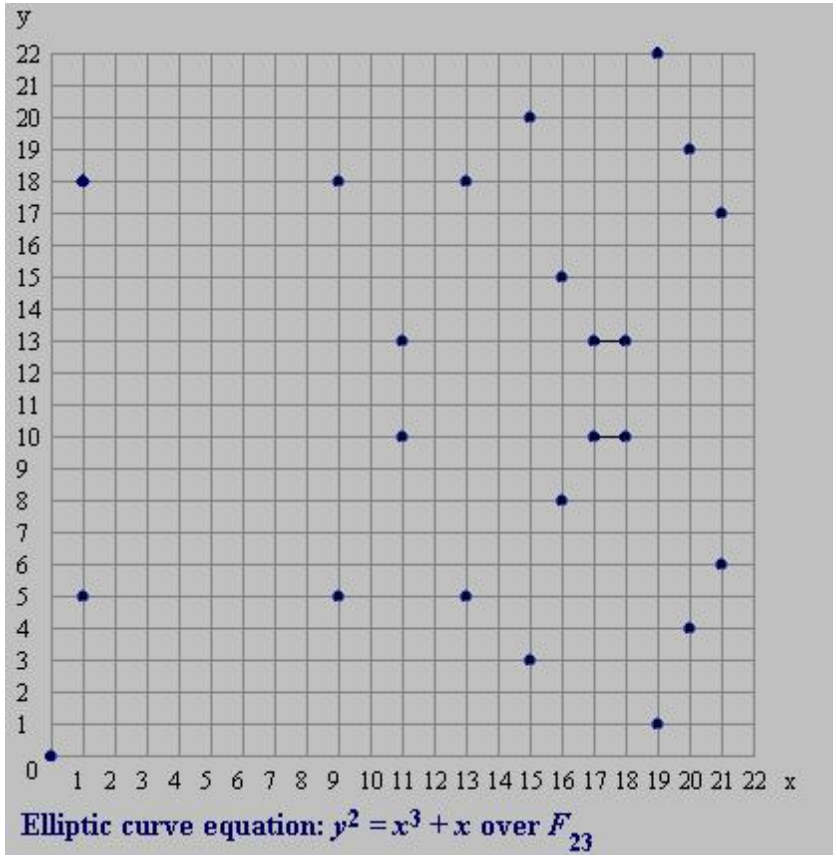
Krzywe eliptyczne – przykłady

Krzywe opisane równaniami $y^2 = x^3 - x + b/10$ dla $b = 0, 1, 2, 3, 4, 5$ nad ciałem liczb rzeczywistych



Krzywe eliptyczne – jeszcze jeden przykład

Krzywa opisana równaniem $y^2 = x^3 + x$ nad Z_{23} składa się z 23 punktów: $(0,0)$, $(1,5)$, $(1,18)$, $(9,5)$, $(9,18)$, $(11,10)$, $(11,13)$, $(13,5)$, $(13,18)$, $(15,3)$, $(15,20)$, $(16,8)$, $(16,15)$, $(17,10)$, $(17,13)$, $(18,10)$, $(18,13)$, $(19,1)$, $(19,22)$, $(20,4)$, $(20,19)$, $(21,6)$, $(21,17)$ oraz punktu w nieskończoności, oznaczanego przez 0.



Krzywe eliptyczne – dodawanie

- Punkty krzywej eliptycznej tworzą grupę addytywną.
- Elementem neutralnym dodawania jest punkt w nieskończoności.
- Dla krzywej z poprzedniego przykładu:
  - $(x, y) + 0 = (x, y)$,
  - $0 + (x, y) = (x, y)$,
  - $-(x, y) = (x, -y)$,
  - $(0, 0) + (1, 5) = (1, 18)$,
  - $(0, 0) + (9, 5) = (18, 13)$,
  - $\ldots$
- Definiuje się mnożenie punktu krzywej eliptycznej przez liczbę całkowitą.
- Dla krzywej z poprzedniego przykładu:
  - $2 \cdot (1, 5) = (0, 0)$,
  - $3 \cdot (1, 5) = (1, 18) = -(1, 5)$,
  - $4 \cdot (1, 5) = 0$,
  - $\ldots$

Schemat Diffiego–Hellmana

- Umożliwia uzgodnienie wspólnego tajnego klucza, używając otwartego kanału transmisyjnego.
- Alicja i Bolek uzgadniają wspólną dużą liczbę pierwszą p i generator g grupy Z_p^*.
- Alicja losuje a i wysyła Bolkowi g^a.
- Bolek losuje b i wysyła Alicji g^b.
- Alicja oblicza wspólny klucz $k = (g^b)^a$.
- Bolek oblicza wspólny klucz $k = (g^a)^b$.
- Bezpieczeństwo opiera się na trudności rozwiązania problemu logarytmu dyskretnego: znając p, g, g^a i g^b, nie umiemy efektywnie obliczyć a, b i g^{ab}.

Schemat Diffiego–Hellmana z wykorzystaniem krzywej eliptycznej

- Alicja i Bolek uzgadniają wspólną krzywą eliptyczną E i generator P grupy punktów na E.
- Alicja losuje a i wysyła Bolkowi aP.
- Bolek losuje b i wysyła Alicji bP.
- Alicja oblicza wspólny klucz $k = a(bP)$.
- Bolek oblicza wspólny klucz $k = b(aP)$.
- Bezpieczeństwo opiera się na trudności rozwiązania problemu logarytmu dyskretnego na krzywej eliptycznej: znając E, P, aP i bP, nie umiemy efektywnie obliczyć a, b i abP.

Zalety i wady krzywych eliptycznych

- Zaleta: krzywe eliptyczne wymagają krótszych kluczy niż RSA. 3072-bitowe RSA zapewnia podobny poziom bezpieczeństwa jak krzywa nad ciałem $GF(2^{256})$.
- Zaleta: dla jednego ciała $GF(p^m)$ istnieje wiele krzywych eliptycznych do wyboru.
- Wada: obliczenia na krzywych eliptycznych mają tajemniczą strukturę.

Gdzie jeszcze warto zajrzeć?

- W lipcu 2009 r. za pomocą 200 konsol PlayStation3 złamano system krytpograficzny używający krzywej eliptycznej $y^2 = x^3 + ax + b$ nad ciałem Z_p z generatorem P, gdzie
  $a = 44516852250937147720845982735484 24$,
  $b = 20611183968086532029029961663885 14$,
  $P = (188281465057972534892222377871375 2,$
  $341987549103317082716786189608268 8)$,
  a $p = (2^{128} - 3)/(11 \cdot 6949)$ jest 112-bitową liczbą pierwszą,
  http://lacal.epfl.ch/page81774.html.
- Dan Boneh, Twenty Years of Attacks on the RSA Cryptosystem,
  http://crypto.stanford.edu/~dabo/papers/RSA-survey.pdf.