Ferienakademie 2001:

Kryptographie und Sicherheit offener Systeme

Faktorisierung

Stefan Büttcher

＜stefan@buettcher.org＞

Definition. (RSA-Problem)

Gegeben:

- $n = pq$, ein RSA-Modul mit unbekannten Primfaktoren p und q,
- ein dazugehöriger Public Key b.

Gesucht:

- der zu b passende Private Key a: $a \times b \equiv 1 \pmod{\phi(n)}$.

Definition. (Faktorisierungsproblem)

Gegeben:

- eine (zusammengesetzte) Zahl n mit unbekannten Primfaktoren

Gesucht:

- Primzahlpotenzen $p_i^{e_i}$ und eine endliche Indexmenge I:

$$\prod_{i \,\epsilon\, I} (p^{e_i}) = n.$$

Satz. Das RSA-Problem lässt sich in polynomialer Zeit auf das Faktorisierungsproblem reduzieren:

$$\text{RSA-Problem} \leq_P \text{Faktorierungsproblem.}$$

Begründung:

- Es sei $\mathcal{A}$ ein Algorithmus, der die Primfaktorzerlegung $\mathcal{P}(n) = \{p_1^{e_1}, \ ..., \ p_k^{e_k}\}$ in $\mathcal{O}(|n|^j)$, $j \ \epsilon \ N$, berechnet.
- Wegen $a \times b \equiv 1 \ (\mathsf{mod} \ \phi(n))$ lässt sich sich der Private Key a mit dem erweiterten Euklidischen Algorithmus in $\mathcal{O}(|n|^2)$ berechnen.

$\Rightarrow$ Wenn das Faktorisierungsproblem effizient lösbar ist, dann ist RSA unsicher!

Ein erster Versuch: Trial Division

Idee: Für alle natürlichen Zahlen $p \leq \lfloor \sqrt{n} \rfloor$ teste, ob $p \mid n$.

- $p \mid n$: Teiler gefunden: *STOP*.
- $p \nmid n$: Kein Teiler: *CONTINUE*.

Aufwand: $\mathcal{O}(|n|^2 \times \sqrt{2}^{|n|})$.

- Division: $\mathcal{O}(|n|^2)$.
- Anzahl Schleifendurchläufe (worst case): $\sqrt{n} \approx \sqrt{2}^{|n|}$.

Verbesserung: Teste nur Primzahlen p.

- $\mathcal{P} := \{p \leq \lfloor \sqrt{n} \rfloor : \ p \text{ ist Primzahl}\}$. $|\mathcal{P}| \approx \frac{\sqrt{n}}{ln(\sqrt{n})}$.
- Gesamtaufwand: $\mathcal{O}(|n| \times \sqrt{2}^{|n|})$.

```
Number trialDivision(Number n) {
  Number s := ⌊√n⌋;
  boolean isPrime[] := new boolean[s];
  for (i := 2; i <= s; i++) isPrime[i] := true;
  for (i = 2; i <= ⌊√s⌋); i++)
    if (isPrime[i]) {
      j := i * i;
      while (j <= s) {
        isPrime[j] := false;
        j := j + i;
      }
    }
  for (i = 2; i <= s; i++)
    if ((isPrime[i]) && (n % i == 0)) return i;
  return -1;
}
```

Methode von Fermat

Idee: Finde Zahlen $x,\ y:\ \ n = x^2 - y^2 = (x - y) \times (x + y)$.

```
x := ⌈√n⌉;
y := 0;
do {
  while (x² - y² > n)
    y := y + 1;
  if (x² - y² < n)
    x := x + 1;
} while (x² - y² ≠ n);
```

- Entgegengesetzte Laufrichtung: $\lceil \sqrt{n} \rceil$, ($\lceil \sqrt{n} \rceil - 1$), ..., 1.
- Gut geeignet zum Finden von Faktoren in der Nähe von $\sqrt{n}$.

Pollards Rho-Algorithmus

Sei n zusammengesetzt und p ein nicht-trivaler Teiler von n. Sei ferner $f(x) := x^2 + 1$. Betrachte die Folge

$$x_0 := 1, \;\; x_i := f(x_{i-1}) \bmod n, \; i \geq 1.$$

Sei nun $y_i := x_i \bmod p$. Wegen

$$x_i \equiv f(x_{i-1}) \pmod{n}$$

folgt dann

$$y_i \equiv f(y_{i-1}) \pmod{p}.$$

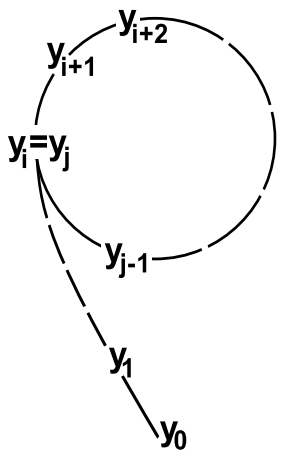
Da es nur p verschiedene Kongruenzklassen modulo p gibt, ergibt sich irgendwann eine Kollision:

$$y_i = y_j, \;\; i \neq j,$$

und somit ein Zyklus:

$$y_{i+k} = y_{j+k} \;\; \forall \, k \, \epsilon \, N.$$

Der Namensgeber:



Aus $y_i = y_j$ folgt nun:

$$x_i \equiv x_j \pmod{p} \quad \Rightarrow \quad p \mid (x_i - x_j).$$

Falls $x_i \neq x_j$, ist mit $ggT(n, x_i - x_j)$ ein echter Teiler von n gefunden.

Problem: Wie findet man die Indizes i und j?

Speichern aller Werte bei Weitem zu teuer!

Ausweg: Verfahren von Brent oder Floyd's Cycle Trick

- *Floyd's Cycle Trick*

  Konstruiere eine Folge $(x_i,\ y_i) := (f(x_i),\ f(f(y_i)))$:
  $(x_1,\ x_2),\ (x_2,\ x_4),\ (x_3,\ x_6),\ (x_4,\ x_8),\ \ldots$
  Ist ein Paar $(x,\ y)$ mit $ggT(n,\ x-y)$ gefunden: *STOP*.

- *Verfahren von Brent*

  Betrachte die Differenzen

$$x_1 - x_3,$$

$$x_3 - x_6,\ x_2 - x_7,$$

$$x_7 - x_{12},\ x_7 - x_{13},\ x_7 - x_{14},\ x_7 - x_{15}$$

  Und deren Produkte, z.B. $m := (x_3 - x_6) \times (x_2 - x_7)$. Ist ein Produkt m mit $ggT(n, m)$ gefunden: *STOP*. (Bei $ggT(n, m) = n$ evtl. Backtracking.)

Pollards (p-1) - Algorithmus

Definition. Sei B eine natürliche Zahl. Eine ganze Zahl b heiße

- *B-glatt*, falls für alle Primfaktoren $p_i \mid b$ gilt: $p_i \leq B$,
- *B-potenzglatt*, falls für alle Primpotenzen $p_i^{e_i} \mid b$ gilt: $p_i^{e_i} \leq B$.

Pollards Idee

Sei n eine zusammengesetzte Zahl und ein p Primfaktor. Sei k beliebig, so dass $(p-1) | k$. Wegen

$$a^{p-1} \equiv 1 \,(\mathsf{mod}\, p) \quad (\mathsf{kleiner\ Fermatscher\ Satz})$$

ist auch

$$\begin{aligned} & a^k \equiv 1 \,(\mathsf{mod}\, p) \quad (\mathsf{kleiner\ Fermatscher\ Satz}) \\ \Rightarrow \quad & p \mid (a^k - 1). \end{aligned}$$

Wenn $n \not| (a^k - 1)$, dann ist wie beim Rho-Verfahren $ggT(n, a^k - 1)$ ein echter Teiler von n.

Problem: Wie findet man ein geeignetes k?

Sei $(p-1)$ B-potenzglatt. Dann ist $k := B!$ Vielfaches von $p - 1$.

```
for (j := 2; j <= B; j++) {
  a := a^j mod n;
  if (j % 5 == 0) {
    g := gcd(a - 1, n);
    if (g > 1) return g;
  }
}
```

Platzbedarf: $\mathcal{O}(|n|)$. **Zeitbedarf:** $\mathcal{O}(\frac{B \times |n|}{ln(B)})$.

Fermats Idee: Finde x, y, so dass

$$x^2 - y^2 = n.$$

Variation: (von Kraitchik oder Legendre)

$$\begin{aligned} & x^2 \equiv y^2 \pmod{n} \\ \Rightarrow \quad & n \mid (x^2 - y^2) \\ \Rightarrow \quad & n \mid (x+y) \times (x-y). \end{aligned}$$

Wenn nun $n \not\mid (x-y)$, dann ist mit $g := ggT(n, x-y)$ ein nicht-trivialer Faktor gefunden.

Problem: Wie findet man Paare $(x, y) : \ x^2 \equiv y^2 \pmod{n}$?

Methode von Dixon

Sei $f(s) := s \times s \bmod n$ und $\mathcal{B}$ eine Menge von Primzahlen („Faktorbasis").

Suche Paare $(s,\ f(s))$, so dass $s \neq f(s)$ und $f(s)$ über der Faktorbasis zerlegbar ist:

$$p \text{ ist prim} \Rightarrow (p \mid f(s) \Rightarrow p \,\epsilon\, \mathcal{B}).$$

Bilde Produkte

$$S := s_0 \times s_1 \times \cdots \times s_j\,,$$
$$S' := f(s_0) \times f(s_1) \times \cdots \times f(s_j),$$

so dass alle Primfaktorpotenzen von S' geraden Exponenten haben.

Es gilt

$$s_i^2 \equiv f(s_i) \pmod n \ \ \forall\, i \,\epsilon\, I \ \Rightarrow\ S^2 \equiv S' \pmod n.$$

Die gesuchte Form ist erreicht:

$$x := S\ ,\ \ y := \sqrt{S'}.$$

Beispiel. Sei $n := 187$ ein RSA-Modul und $\mathcal{B} := \{3,\ 5,\ 31\}$ die Faktorbasis. Es seien bereits folgende Kongruenzen gefunden:

$$\begin{aligned} 23^2 &\equiv 5 \times 31 \pmod{n}, \\ 24^2 &\equiv 3 \times 5 \pmod{n}, \\ 29^2 &\equiv 3 \times 31 \pmod{n}. \end{aligned}$$

Dann folgt:

$$\begin{aligned} (23 \times 24 \times 29)^2 &\equiv (3 \times 5 \times 31)^2 \pmod{n}, \\ 113^2 &\equiv 91^2 \pmod{n}. \end{aligned}$$

$\Rightarrow ggT(113 - 91,\ 187) = 11$ ist ein Teiler von n.

Ist $|\mathcal{B}|$ die Größe der Faktorbasis, dann müssen höchstens $C := |\mathcal{B}| + 1$ Kongruenzen gefunden werden, damit die quadratische Form hergestellt werden kann.

Herstellen der quadratischen Form durch Gauß'sche Elimination.

Beispiel. Sei $n := 187$ und $\mathcal{B} := \{2, 3, 5, 7, 11, 17, 31\}$. Folgende Kongruenzen sind bereits gefunden:

$$20^2 \equiv 2 \times 17,\ 22^2 \equiv 2 \times 5 \times 11,\ 23^2 \equiv 5 \times 31,$$
$$24^2 \equiv 3 \times 5,\ 27^2 \equiv 2^3 \times 3 \times 7,\ 28^2 \equiv 2^2 \times 3^2,$$
$$29^2 \equiv 3 \times 31 \text{ und } 33^2 \equiv 2 \times 7 \times 11.$$

Es resultiert die Matrix:

2	3	5	7	11	17	31		20	22	23	24	27	28	29	33
1	0	0	0	0	1	0	|	1	0	0	0	0	0	0	0
1	0	1	0	1	0	0	|	0	1	0	0	0	0	0	0
0	0	1	0	0	0	1	|	0	0	1	0	0	0	0	0
0	1	1	0	0	0	0	|	0	0	0	1	0	0	0	0
3	1	0	1	0	0	0	|	0	0	0	0	1	0	0	0
2	2	0	0	0	0	0	|	0	0	0	0	0	1	0	0
0	1	0	0	0	0	1	|	0	0	0	0	0	0	1	0
1	0	0	1	1	0	0	|	0	0	0	0	0	0	0	1

Die Matrix wird modulo 2 reduziert und dann mit Gauß transformiert:

2	3	5	7	11	17	31		20	22	23	24	27	28	29	33
1	0	0	0	0	1	0		1	0	0	0	0	0	0	0
1	0	1	0	1	0	0		0	1	0	0	0	0	0	0
0	0	1	0	0	0	1		0	0	1	0	0	0	0	0
0	1	1	0	0	0	0		0	0	0	1	0	0	0	0
1	1	0	1	0	0	0		0	0	0	0	1	0	0	0
0	0	0	0	0	0	0		0	0	0	0	0	1	0	0
0	1	0	0	0	0	1		0	0	0	0	0	0	1	0
1	0	0	1	1	0	0		0	0	0	0	0	0	0	1

2	3	5	7	11	17	31		20	22	23	24	27	28	29	33
1	0	0	0	1	0	0		1	0	0	0	0	0	0	0
0	1	1	0	0	0	0		0	0	0	1	0	0	0	0
0	0	1	1	0	1	0		1	0	0	1	1	0	0	0
0	0	0	1	0	1	1		1	0	0	0	0	0	1	0
0	0	0	0	1	1	1		1	1	0	1	1	0	1	0
0	0	0	0	0	1	0		1	1	0	1	1	0	0	1
0	0	0	0	0	0	0		0	0	1	1	0	0	1	0
0	0	0	0	0	0	0		0	0	0	0	0	1	0	0

Es ergibt sich, dass

$$ggT(113 - 19,\ 187) = 11 \text{ und } ggT(28 - 6,\ 187) = 11.$$

Frage: Wie viele Kongruenzen müssen faktorisiert werden?

Annahmen:

- Die Zeilen $(B \times B)$-Matrix sind linear unabhängig.
- Ist ein perfektes Quadrat gefunden, gilt in 50% aller Fälle:

$$x \equiv \pm y \pmod{n} \Rightarrow gcd(x - y, n) = n.$$

Dann ist die Wahrscheinlichkeit, einen nicht-trivialen Teiler zu finden

$$P = 1 - 2^{B-C}.$$

Optimale Wahl der Faktorbasis

Welche Primzahlen sollten aufgenommen werden?

Gesucht ist nach Primzahlen p_i, so dass

$$p \mid (r \times r - n)$$
$$\Rightarrow \quad n \equiv r^2 \pmod{n},$$

also n quadratischer Rest modulo p ist.

$\Rightarrow$ das Legendre-Symbol (n/p) muss +1 sein.

$\rightarrow$ Reduktion der Faktorbasis um ungefähr die Hälfte!

Das Quadratische Sieb

Wenn n quadratischer Rest modulo p ist, also

$$n \equiv s^2 \pmod{p} \text{ oder } n \equiv (-s)^2 \pmod{p}$$

gilt, dann ist wegen

$$(s \pm k \times p)^2 = s^2 \pm 2skp + k^2p^2$$

auch

$$n \equiv r^2 \pmod{p} \ \forall\, r \equiv \pm s \pmod{p}.$$

Das heißt: Wenn es muss für jedes $p \,\epsilon\, \mathcal{B}$ nur einmal der Wert von s berechnet werden. Danach steht genau fest, welche $f(r)$ durch ein bestimmtes p teilbar sind.

$\rightarrow$ Es fallen keine erfolglosen Probedivisionen mehr an!

Beispiel. Sei $n := 391$. Aufbau einer Faktorbasis $\mathcal{B}$ mit $|\mathcal{B}| = 4$:

$$(n/2) = +1,\ 1^2 \equiv n \pmod 2,$$
$$(n/3) = +1,\ 1^2 \equiv 2^2 \equiv n \pmod 3,$$
$$(n/5) = +1,\ 1^2 \equiv 4^2 \equiv n \pmod 5,$$
$$(n/7) = -1,\ (n/11) = -1$$
$$(n/13) = +1,\ 1^2 \equiv 12^2 \equiv n \pmod{13}.$$

Aufbau des Siebintervalls:

$$16^2 \equiv -135,\ 17^2 \equiv -102,\ 18^2 \equiv -67,$$
$$19^2 \equiv -10,\ 20^2 \equiv +19,\ 21^2 \equiv +50,$$
$$22^2 \equiv +93,\ 23^2 \equiv +138,\ 24^2 \equiv +185.$$

Ohne Probedivision ist sofort erkennbar:

- $f(17)$, $f(19)$, $f(21)$ und $f(23)$ sind durch 2 teilbar,
- $f(16)$, $f(19)$, $f(21)$ und $f(24)$ sind durch 5 teilbar.

Dividieren ohne Division

Angenommen, eine Zahl m könne komplett üeber der Basis zerlegt werden:

$$\exists\, \overrightarrow{p} := (p_1, p_2, ..., p_k) : \qquad m = \prod_{i=1}^{k} p_i$$

$$\Rightarrow \quad log(m) = \sum_{i=1}^{k} log(p_i).$$

$\rightarrow$ Beim Sieben sind keine Divisionen mehr nötig!

Zusammenfassung: Quadratisches Sieb

- Aufbau einer Faktorbasis mit Primzahlen p : $(n/p) = +1$,
- Lösen der Kongruenzen $t^2 \equiv n \pmod{p}$,
- Finden von genügend vielen zerlegbaren Kongruenzen,
- Konstruktion der quadratischen Form.

Continued Fractions (CFRAC)

Selbes Prinzip wie beim Quadratischen Sieb, aber Geschwindigkeit nicht durch Siebverfahren, sondern durch spezielle Reihenentwicklung, so dass $f(r) < 2\sqrt{n}$.

Asymptotische Laufzeiten

- Continued Fractions: $\mathcal{O}\left(e^{2\times\sqrt{ln(n)\times ln(ln(n))}}\right)$,
- Quadratic Sieve: $\mathcal{O}\left(e^{(1+o(1))\times\sqrt{ln(n)\times ln(ln(n))}}\right)$,
- Elliptic Curve Method: $\mathcal{O}\left(e^{(1+o(1))\times\sqrt{2\times ln(p)\times ln(ln(p))}}\right)$,
- Number Field Sieve: $\mathcal{O}\left(e^{(1.92+o(1))\times(ln(n))^{1/3}\times(ln(ln(n)))^{2/3}}\right)$.