Napadi na RSA kriptosustav s malim tajnim eksponentom

Jović, Monika

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku

Monika Jović

Napadi na RSA kriptosustav s malim tajnim eksponentom

Diplomski rad

Osijek, 2016.

Sveučilište J.J. Strossmayera u Osijeku
Odjel za matematiku

Monika Jović

Napadi na RSA kriptosustav s malim tajnim eksponentom

Diplomski rad

Mentor: doc.dr.sc. Ivan Matić

Osijek, 2016.

Sadržaj

1. Uvod

Od početka čovječanstva postojala je želja za razmjenom informacija na način da ih mogu pročitati samo osobe kojima je ona namijenjena. Stoga su nastale metode pomoću kojih se željene informacije mogu slati na tajni način. S razvojem čovječanstva, razvile su se složenije metode tajnog prijenosa podataka. U počecima razmjene informacija na tajni način bilo je dovoljno posjedovati olovku i papir, dok je danas to nezamislivo bez primjene računala.

Danas privatnost i zaštita podataka igraju veliku ulogu u svakodnevnom korištenju medija, interneta, bankarstva i slično. Razvojem kriptografije paralelno se razvijaju i tehnike razbijanja kriptosustava. Iako se svaki kriptosustav može razbiti grubom silom, taj posao je neefikasan i neisplativ. Cilj je razviti algoritam kojim se u razumnom vremenu pronalaze tajni podaci. Cijena izvođenja algoritma, odnosno složenost algoritma, se mjeri u vremenu izvršavanja i potrebnoj memoriji.

U ovom radu se razmatra vremenska složenost algoritama. Nemoguće je točno odrediti vremensku složenost algoritma, nego se asimptotski može odrediti broj potrebnih operacija kada ulazni podaci neograničeno rastu. Za procjenu efikasnosti algoritma koristimo $\mathcal{O}$-notaciju. Vrijede sljedeće nejednakosti

$$\mathcal{O}(1) < \mathcal{O}(\log n) < \mathcal{O}(n) < \mathcal{O}(n \log n) < \mathcal{O}(n^2) < \mathcal{O}(n^3) < ... < \mathcal{O}(2^n) < \mathcal{O}(n!),$$

pri čemu je n veličina ulaza. Algoritmi kojima je vrijeme izvršenja eksponencijalno, mogu biti nerješivi u razumnom vremenu.

2. Matematička podloga

Unutar ovog poglavlja navest ćemo zaključke iz teorije brojeva, modularne aritmetike, rešetki, te kriptografije i kriptoanalize kako bi u sljedećem poglavlju definirali RSA kriptosustav i predstavili napade na RSA kriptosustav s malim tajnim eksponentom.

2.1. Teorija brojeva

U nastavku se nalaze neki rezultati iz teorije brojeva. Budući da je u ovoj grani matematike jedan od najvažnijih pojmova djeljivost, započeti ćemo s definicijom djeljivosti dva broja.

Definicija 2.1. *Neka su $a, b \in \mathbb{Z}$ te $a \neq 0$. Kažemo da a dijeli b ako postoji $d \in \mathbb{Z}$ takav da je $b = a \cdot d$ i tada pišemo $a \mid b$. Broj a nazivamo djeliteljem broja b, a broj b višekratnikom broja a. Ukoliko a ne dijeli b, onda pišemo $a \nmid b$.*

Teorem 2.1 (**Teorem o dijeljenju s ostatkom**). *Za proizvoljan cijeli broj a i prirodan broj b postoje jedinstveni cijeli brojevi q i r takvi da je $a = bq + r$, gdje je $0 \leq r < b$.*

Broj r nazivamo **ostatak pri djeljenju**, a q **kvocijent cjelobrojnog dijeljenja**. Dokaz teorema o dijeljenju s ostatkom možete pronaći u [8].

Cijeli broj c koji dijeli cijele brojeve a i b, odnosno $c \mid a$ i $c \mid a$, nazivamo **zajedničkim djeliteljem** brojeva a i b. Ukoliko je barem jedan od brojeva a i b različit od nule, tada iz konačno mnogo zajedničkih djelitelja brojeva a i b, najvećeg od njih označavamo s (a, b). Kažemo da su cijeli brojevi a i b relativno prosti ukoliko je $(a, b) = 1$.

Efikasan algoritam za određivanje najvećeg zajedničkog djelitelja dva broja a i b je Euklidov algoritam koji je dan sljedećim teoremom.

Teorem 2.2 (**Euklidov algoritam**). *Neka su $a, b \in \mathbb{Z}$ i $b > 0$. Pretpostavimo da je uzastopnom primjenom Teorema 2.1 dobiven niz jednakosti*

$$\begin{aligned} a &= bq_0 + r_1, & 0 < r_1 < b, \\ b &= r_1 q_1 + r_2, & 0 < r_2 < r_1, \\ r_1 &= r_2 q_2 + r_3, & 0 < r_3 < r_2, \\ &\vdots & \\ r_{j-2} &= r_{j-1} q_{j-1} + r_j, & 0 < r_j < r_{j-1}, \\ r_{j-1} &= r_j g_j. \end{aligned} \tag{1}$$

Tada je (a, b) jednak posljednjem nenul ostatku, odnosno r_j.

Prvu jednakost Euklidovog algoritma možemo zapisati u obliku $r_1 = a - bq_0$. Drugu jednakost analogno možemo zapisati $r_2 = b - r_1 q_1$, te uvrštavanjem prve u drugu jednadžbu dobivamo $r_2 = b(1 + q_0 q_1) - aq_1$. Analognim zaključivanjem treća jednadžba ima sljedeći

oblik $r_3 = a(1 + q_1q_2) - b(q_1 + q_2 + q_0q_1q_2)$. Nastavljanjem niza jednadžbi dolazimo do zaključka da postoje $x, y \in \mathbb{Z}$ takvi da vrijedi

$$ax + by = r_j = (a, b),$$

a dobivenu jednakost nazivamo **Bezoutov identitet**.

Rješenja jednadžbe $ax + by = (a, b)$ dobivamo pomoću proširenog Euklidovog algoritma na sljedeći način:

$$\begin{array}{lll} r_{-1} = a, & r_0 = b; & r_i = r_{i-2} - q_i r_{i-1}; \\ x_{-1} = 1, & x_0 = 0; & x_i = x_{i-2} - q_i x_{i-1}; \\ y_{-1} = 0, & y_0 = 1; & y_i = y_{i-2} - q_i y_{i-1}, \end{array}$$

te vrijedi $ax_i + by_i = r_i$ za $i = -1, 0, \ldots, j$. Primjećujemo da za $i = j$ vrijedi $ax_j + by_j = r_j = (a, b)$.

Zapisom jednakosti Euklikovog algoritma na drugačiji način dolazimo do pojma **verižnog razlomka**. Za niz jednakosti (1) zapisanih na sljedeći način

$$\begin{aligned} \frac{a}{b} &= q_0 + \frac{r_1}{b}, \\ \frac{b}{r_1} &= q_1 + \frac{r_2}{r_1}, \\ &\vdots \\ \frac{r_{j-2}}{r_{j-1}} &= q_{j-1} + \frac{r_j}{r_{j-1}}, \\ \frac{r_{j-1}}{r_j} &= q_j, \end{aligned}$$

dobivamo raspis broja $\frac{a}{b}$ u jednostavan verižni razlomak:

$$\frac{a}{b} = q_0 + \cfrac{1}{g_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots\ q_{j-1} + \cfrac{1}{q_j}}}}.$$

Praktičniji zapis verižnog razlomka $\frac{a}{b}$ je $\frac{a}{b} = [q_0; q_1, q_2, \ldots]$, dok za **$n$-tu konvergentu** verižnog razlomka $\frac{a}{b}$ koristimo izraz $\frac{a_n}{b_n} = [q_0; q_1, q_2, \ldots q_n]$, gdje su q_i **parcijalni kvocijenti**. Verižni razlomak koji ima konačno mnogo parcijalnih kvocijenata je konačan verižni razlomak.

Teorem 2.3 (**Lagrangeov teorem o verižnim razlomcima**). *Neka je $\alpha \in \mathbb{Q}$, te $a, b \in \mathbb{Z}$ takvi da je $b \leq 1$ i $|\alpha - \frac{a}{b}| < \frac{1}{2b^2}$. Tada je $\frac{a}{b}$ neka konvergenta od α.*

2.2. Modularna aritmetika

Modularna aritmetika predstavlja aritmetički sustav u kojem se brojevi vraćaju u krug nakon što dođu do određene vrijednosti modulo. Drugi naziv za ovu vrstu aritmetike je teorija kongruencija.

Definicija 2.2. *Neka su $a, b \in \mathbb{Z}$, te $n \in \mathbb{N}$. Ukoliko n dijeli $a - b$, tada kažemo da je a kongruentan b modulo n i pišemo $a \equiv b \pmod{n}$. Uočavamo da je $n \mid a$ ako i samo ako je $a \equiv 0 \pmod{n}$.*

Pretpostavimo da brojeve a i b dijelimo s n. Tada prema Euklidovom algoritmu dobivamo $a = q_1 n + r_1$ i $b = q_2 n + r_2$, gdje su $0 \leq r_1 \leq n-1$ i $0 \leq r_2 \leq n-1$. Ako brojevi a i b pri dijeljenju s n daju isti ostatak kažemo da je a kongruentan b modulo n. Za ostatak pri djeljenju a s brojem n koristimo oznaku $a \mod n$. Dakle, vrijedi da je $a \mod n = b \mod n$ ako i samo ako je $a \equiv b \pmod{n}$.

Neka je $n \neq 1 \in \mathbb{N}$. Skup $S = \{a_1, a_2, \ldots, a_n\}$ nazivamo potpuni sustav ostataka modulo n ako za svaki $b \in \mathbb{Z}$ postoji točno jedan $a_j \in S$ za koji vrijedi $b \equiv a_j \pmod{n}$. Ukoliko za $b \in \mathbb{Z}$ takav da je $(b, n) = 1$, postoji točno jedan $a_j \in S$ za koji vrijedi $b \equiv a_j \pmod{n}$, onda S nazivamo reducirani sustav ostataka modulo n. Svi reducirani sustavi ostataka modulo n imaju isti broj elemenata.

Postoji beskonačno mnogo potupnih sustava ostataka modulo n, jedan od njih je i $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ kojeg nazivamo sustav najmanjih nenegativnih ostataka modulo n.

Primjerice, skup $\{0, 1, 2, 3\}$ je potpun sustav ostataka modulo 4, dok je skup $\{1, 3\}$ reducirani sustav ostataka modulo 4.

Na skupu $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ se definiraju operacije zbrajanja, oduzimanja i množenja na analogan način kao i na skupu $\mathbb{Z}$, a rezultat je reduciran modulo n, odnosno ukoliko rezultat nije iz skupa $\{0, 1, \ldots, n-1\}$ tražimo njegov ostatak pri djeljenju s n. Navedimo osnovna svojstva operacija zbrajanja i množenja na skupu $\mathbb{Z}_n$:

1. zatvorenost obzirom na zbrajanje $\forall a, b \in \mathbb{Z}_n$ vrijedi $a + b \in \mathbb{Z}_n$
2. asocijativnost zbrajanja $\forall a, b, c \in \mathbb{Z}_n$ vrijedi $(a + b) + c = a + (b + c)$
3. postojanje neutralnog element obzirom na zbrajanje $\exists 0 \in \mathbb{Z}_n$ t.d. $\forall a \in \mathbb{Z}_n$ vrijedi $a + 0 = 0 + a = a$
4. postojanje suprotnog elementa obzirom na zbrajanje $\forall a \in \mathbb{Z}_n \ \exists (n-a) \in \mathbb{Z}_n$ t.d. vrijedi $a + (n - a) = (n - a) + a = 0$.
   Suprotni element $n - a$ jednostavnije označavamo s $-a$
5. komutativnost zbrajanja $\forall a, b \in \mathbb{Z}_n$ vrijedi $a + b = b + a$
6. zatvorenost obzirom na množenje $\forall a, b \in \mathbb{Z}_n, ab \in \mathbb{Z}_n$

7. komutativnost obzirom na množenje $\forall a, b \in \mathbb{Z}_n$ vrijedi $ab = ba$

8. asocijativnost obzirom na množenje $\forall a, b, c \in \mathbb{Z}_n$ vrijedi $(ab)c = a(bc)$

9. distributivnost množenja obzirom na zbrajanje slijeva i zdesna $\forall a, b, c \in \mathbb{Z}_n$ vrijedi $(a + b)c = (ac) + (bc)$ i $a(b + c) = (ab) + (ac)$.

Skup $\mathbb{Z}_n$ je grupa obzirom na operaciju zbrajanja jer $\mathbb{Z}_n$ zadovoljava svojstva 1.-4. Štoviše, jer vrijedi svojstvo 5., $\mathbb{Z}_n$ je i Abelova grupa. Budući da vrijede svojstva 1.-10. kažemo da je $\mathbb{Z}_n$ prsten.

$\mathbb{Z}_n^* = \{1, 2, \ldots, n-1\}$ je skup svih invertibilnih elemenata u $\mathbb{Z}_n$, te za $\mathbb{Z}_n^*$ vrijedi da je Abelova grupa obzirom na operaciju množenja.

Za $n \in \mathbb{N}$ definiramo funkciju $\varphi\colon \mathbb{N} \to \mathbb{N}$ koja vraća broj brojeva u nizu $1, 2, \ldots, n$ koji su relativno prosti s n. Funkciju φ nazivamo Eulerova funkcija. Uočavamo da je $\varphi(n)$ upravo broj elemenata reduciranog sustava ostataka modulo n.

Ako n možemo rastaviti na proste faktore $n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_k^{\alpha_k}$ tada vrijedi

$$\varphi(n) = \prod_{j=1}^{k} p_j^{\alpha_j}(p_j - 1).$$

Teorem 2.4 (Eulerov teorem). *Ako je $(a, n) = 1$, onda je $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

2.3. Rešetke

Definicija 2.3. *Neka je $\{v_1, v_2, \ldots, v_m\} \subseteq \mathbb{R}^n, m \leq n$ skup linearno nezavisnih vektora. Rešetka $\mathcal{L}$ je skup svih linearnih kombinacija vektora $v_1, v_2, \ldots, v_m$ s cjelobrojnim koeficijentima, odnosno:*

$$\mathcal{L} = \left\{ \sum_{i=1}^{m} x_i v_i | x_i \in \mathbb{Z} \right\}.$$

Vektor v_i je **vektor baze** rešetke $\mathcal{L}$, a **bazu rešetke** čini neki skup linearno nezavisnih vektora koji generiraju $\mathcal{L}$. Oznaka za bazu rešetke $\mathcal{L}$ je $\mathcal{B} = \{v_1, \ldots, v_n\}$.

Rešetka $\mathcal{L}$ se može prikazati preko matrice baze. Za danu bazu $\mathcal{B} = \{v_1, \ldots, v_n\}$, matrica baze B je $m \times n$ matrica čiji redovi čine vektori baze, odnosno

$$B = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}.$$

U nastavku ćemo koristiti oznaku $\mathcal{B}$ kao oznaku baze rešetke $\mathcal{L}$, ali i kao oznaku matrice baze rešetke $\mathcal{L}$, te će iz konteksta biti jasno na što se pri tome misli.

Neka je $\vec{v} = \sum_{i=1}^{m} x_i v_i$ vektor u rešetki $\mathcal{L}$. Tada za vektor $\vec{v}$ definiramo Euklidsku normu na sljedeći način

$$||\vec{v}|| = \left(\sum_{i=1}^{m} x_i^2 \right)^{\frac{1}{2}}.$$

Dimenzija ili rang rešetke $\mathcal{L}$ je broj vektora baze, odnosno $\dim(\mathcal{L}) = m$. Ako je $m = n$, onda za rešetku kažemo da je punog ranga.
Ako rešetka ima dimenziju $\dim(\mathcal{L}) \geq 2$, onda ona ima beskonačno mnogo baza koje imaju isti broj baznih vektora. Sve su bazne matrice u parovima povezane unimodularnom matricom. Odnosno, za dvije bazne matrice $\mathcal{B}$ i $\mathcal{B}'$ rešetke $\mathcal{L}$ postoji unimodularna matrica $\mathcal{U}$ takva da vrijedi $\mathcal{B} = \mathcal{U}\mathcal{B}'$. Unimodularna matrica je cjelobrojna kvadratna matrica čija determinanta iznosi ± 1.

Volumen ili determinanta rešetke $\mathcal{L}$ je

$$\mathrm{vol}(\mathcal{L}) = \sqrt{\det(\mathcal{B}\mathcal{B}^T)},$$

gdje je $\mathcal{B}$ neka od matrica baze rešetke $\mathcal{L}$. Ako je rešetka punog ranga onda je

$$\mathrm{vol}(\mathcal{L}) = \sqrt{\det(\mathcal{B}\mathcal{B}^T)} = |\det(\mathcal{B})|.$$

Budući da su svake dvije bazne matrice povezane unimodularnom matricom, volumen rešetke $\mathcal{L}$ ne ovisi o izboru bazne matrice.

Jedno svojstvo rešetki je da uvijek postoji **najkraći ne-nul vektor** u svakoj rešetki. Odnosno, postoji ne-nul vektor $v \in \mathcal{L}$ takav da za svaki $u \in \mathcal{L}$ vrijedi $||v|| \leq ||u||$. Za normu najkraćeg vektora koristimo oznaku $\lambda_1(\mathcal{L})$ ili λ_1. Uočimo da uvijek postoje dva najkraća vektora u rešetki $\mathcal{L}$, odnosno ako je v najkraći vektor u rešetki, onda je najkraći i vektor $-v$. Minkowskijevim teoremom je dana granica koju mora zadovoljavati norma najkraćeg vektora.

Teorem 2.5 (**Minkowskijeva granica**). *Neka je $\mathcal{L}$ m-dimenzionalna rešetka. Tada postoji vektor $v \in \mathcal{L}$ za koji vrijedi*

$$||v|| \leq \sqrt{m}\mathrm{vol}(\mathcal{L})^{1/m}.$$

2.3.1. LLL-reducirana baza

Prethodno smo spomenuli da svaka rešetka dimenzije barem 2 ima beskonačno mnogo baza. Najčešće nas zanima tzv. **reducirana baza rešetke** koja se sastoji od kratkih vektora. Postupak dobivanja reducirane baze za neku danu bazu nazivamo redukcija baze rešetke. Bazni vektori u reduciranoj bazi su poredani u rastućem poretku obzirom na normu.

Gaussova redukcija rešetke

Kada je rešetka 2-dimenzionalna, onda za konstrukciju reducirane baze koristimo Gaussov

algoritam. Ideja algoritma je smanjivati dulji vektor baze za neki višekratnik drugog vektora baze.

$$v_2 = v_2 - \lfloor \mu \rceil v_1, \qquad \mu = \frac{(v_1, v_2)}{||v_1||^2}$$

gdje je $\lfloor x \rceil = \lfloor x + 1/2 \rfloor$. Ako je v_2 i dalje veći, algoritam staje. U suprotnom, mijenjamo v_1 i v_2 te nastavljamo postupak smanjivanja.

Algorithm 1 Gaussova redukcija

```
while do
    if ||v2|| < ||v1|| then
        zamijeni v1 i v2;
    end if
    ⌊μ⌉ = (v1,v2)/||v1||^2
    if μ = 0 then
        vrati trenutne vektore v1 i v2;
        break
    else
        v2 = v2 − μv1;
    end if
end while
```

Ovim postupkom za neku bazu rešetke $\mathcal{L}$ dobivamo bazu v_1, v_2, koju nazivamo **Gauss reducirana baza**, tako da je v_1 najkraći vektor u rešetki, a $0 \leq \mu \leq 1/2$.

Algoritam se može izvršiti u kvadratnom vremenu obzirom na veličinu ulaznih parametara.

U nastavku je prikazano da se Gaussovim algoritmom dobiva najkraći vektor baze.

Dokaz. Vektori Gauss reducirane baze zadovoljavaju

$$||v_1|| \leq ||v_2|| \qquad \text{i} \qquad \frac{(v_1, v_2)}{||v_1||^2} \leq \frac{1}{2}$$

Neka je z ne-nul vektor u rešetki čiji su vektori baze v_1 i v_2

$$z = \alpha_1 v_1 + \alpha_2 v_2, \qquad \alpha_1, \alpha_2 \in \mathbb{Z}.$$

Vrijedi

$$\begin{aligned} ||z||^2 &= ||\alpha_1 v_1 + \alpha_2 v_2||^2 \\ &= \alpha_1^2 ||v_1||^2 + 2\alpha_1\alpha_2 (v_1, v_2) + \alpha_2^2 ||v_2||^2 \\ &\geq \alpha_1^2 ||v_1||^2 - 2|\alpha_1\alpha_2||(v_1, v_2)| + \alpha_2^2 ||v_2||^2 \\ &\geq (\alpha_1^2 - |\alpha_1||\alpha_2| + \alpha_2^2)||v_1||^2 \end{aligned}$$

Izraz $\alpha_1^2 - \alpha_1\alpha_2 + \alpha_2^2$ je jednak 0 ako i samo ako je $\alpha_1 = \alpha_2 = 0$. No koeficijenti α_1 i α_2 su cijeli brojevi koji nisu oba jednaki nuli, stoga vrijedi $||z||^2 \geq ||v_1||^2$. □

LLL-reducirana baza

Posebna skupina reduciranih baza je **LLL-reducirana baza**.

Neka su $v_1, \ldots, v_m$ vektori baze u rešetki $\mathcal{L}$, te neka su vektori $v_1^*, \ldots, v_m^*$ dobiveni Gram-Schmidt postupkom ortogonalizacije čiji se algoritam nalazi u nastavku.

Algorithm 2 Gram-Schmidt

```
for i = 1, ..., m do
    v_i^* = v_i;
    for j = 1, ..., i − 1 do
        μ_{i,j} = (v_i, v_j^*) / ||v_j||^2;
        v_i^* = v_i^* + μ_{i,j} v_j^*;
    end for
end for
```

Dobiveni koeficijenti $\mu_{i,j}$ ne moraju biti cijeli brojevi, stoga dobivena baza ne mora biti baza rešetke. U polinomijalnom vremenu je moguće reducirati bazu Gram-Schmidtovom metodom.

Definicija 2.4. *Neka je $\{v_1, \ldots, v_m\}$ baza za rešetku $\mathcal{L}$, te $\{v_1^*, \ldots, v_m^*\}$ odgovarajuća baza dobivena Gram-Schmidtovim postupkom. Baza $\{v_1, \ldots, v_m\}$ je LLL-reducirana ako zadovoljava:*

1. *uvjet veličine*
   $|\mu_{i,j}| \leq \frac{1}{2}, \quad 1 \leq j < i \leq m;$
2. *Lovaszov uvjet*
   $||v_i^*||^2 \geq (\frac{3}{4} - \mu_{i,i-1}^2)||v_{i-1}^*||^2, \quad 1 < i \leq m.$

LLL-reducirana baza je skoro ortogonalna (prema prvom uvjetu), te su vektori poredani prema rastućim vrijednostima normi (drugi uvjet).

Uočavamo da je prvi vektor u LLL-reduciranoj bazi najmanje norme, te se može pokazati da vrijedi $||v_1|| \leq 2^{(m-1)/2}||x||$ gdje je x ne-nul vektor rešetke $\mathcal{L}$.

Teorem 2.6. *Neka je $\{v_1, \ldots, v_m\}$ LLL-reducirana baza rešetke $\mathcal{L}$, te $\{v_1^*, \ldots, v_m^*\}$ odgovarajuća baza dobivena Gram-Schmidtovim postupkom. Vrijedi:*

1. $||v_j||^2 \leq 2^{i-1}||v_i^*||^2, \qquad 1 \leq j \leq i \leq m$
2. $\text{vol}(\mathcal{L}) \leq \prod_{i=1}^{m} ||v_i|| \leq 2^{m(m-1)/4}\text{vol}(\mathcal{L})$
3. $||v_1|| \leq 2^{(m-1)/4}(\text{vol}(\mathcal{L}))^{1/m}.$

Dokaz.

1. Prema definiciji LLL-reducirane baze slijedi

$$||v_i^*||^2 \geq (\frac{3}{4} - \mu_{i,i-1}^2)||v_{i-1}^*||^2 \geq \frac{1}{2}||v_{i-1}^*||^2, \qquad i = 1, \ldots, m.$$

Indukcijom dobivamo

$$||v_j^*||^2 \leq 2^{i-j}||v_i^*||^2, \qquad 1 \leq j \leq i \leq m.$$

Iz definicije Gram-Schmidtove baze dobivamo

$$\begin{aligned} ||v_i||^2 &= ||v_i^*||^2 + \sum_{j=1}^{i-1} \mu_{i,j}^2 ||v_j^*||^2 \leq (1 + \sum_{j=1}^{i-1} 2^{i-j-2})||v_i^*||^2 \\ &= \frac{1 + 2^{i-1}}{2}||v_i^*||^2 \leq 2^{i-1}||v_i^*||^2. \end{aligned}$$

Za $1 \leq j \leq i \leq n$ dobivamo

$$||v_j||^2 \leq 2^{j-1}||v_j^*||^2 \leq 2^{j-1+i-j}||v_i^*||^2 = 2^{i-1}||v_i^*||^2.$$

2. Budući da su vektori v_i^* ortogonalni, za tada je $\text{vol}(\mathcal{L}) = \prod_{i=1}^m ||v_i^*||$. Iz $||v_i^*|| \leq ||v_i||$ i tvrdnje 1. slijedi

$$\text{vol}(\mathcal{L}) \leq \prod_{i=1}^m ||v_i|| \leq \prod_{i=1}^m 2^{(i-1)/2}||v_i^*|| \leq 2^{m(m-1)/4} \prod_{i=1}^m ||v_i^*|| \leq 2^{m(m-1)/4}\text{vol}(\mathcal{L}).$$

3. Ako uzmemo da je $j = 1$ u 1. i napravimo produkt po svim i, tada za prethodnu nejednakost dobivamo

$$||v_1||^{2m} \leq \prod_{i=1}^m 2^{i-1}||v_i^*||^2 \leq 2^{m(m-1)/2}\text{vol}(\mathcal{L})^2.$$

□

Uočimo da za $1 \leq i \leq m$ vrijedi

$$||v_1|| \leq ||v_2|| \leq \ldots \leq ||v_i|| \leq 2^{m(m-i)/(4(m+1-i))}\text{vol}(\mathcal{L})^{1/(m+i-1)}.$$

A. K. Lenstra, H. W. Lenstra i L. Lovasz su 1982. godine dali polinomijalni algoritam za konstrukciju LLL-reducirane baze iz neke baze rešetke koji je dobio naziv LLL-algoritam. LLL-algoritam za m-dimenzionalnu rešetku s n-dimenzionalnim vektorima ima vrijeme $\mathcal{O}(nm^5B^3)$ gdje je B granica na veličinu ulaznih vektora baze.

Problem najkraćeg vektora - SVP
Problem se sastoji u pronalasku najkraćeg ne-nul vektora u rešetki. Potrebno je pronaći $v \in \mathcal{L}$ minimalne norme, odnosno $||v|| = \lambda_1(\mathcal{L})$.

U pravilu je ovaj problem teško rješiti kada je dimenzija rešetke velika. Metode bazirane na rešetkama koje će biti opisane u nastavku se mogu gledati kao primjer SVP problema. Budući da se koriste rešetke malih dimenzija, SVP problem se onda može efikasno rješiti. Dovoljno je pronaći LLL-reduciranu bazu kako bi se pronašao najkraći vektor. Stoga je pronalazak LLL-reducirane baze ekvivalentno pronalasku najkraćeg vektora.

Kako bi rješili SVP problem potrebno je rješiti linearnu jednadžbu, što će biti opisano u nastavku. SVP problem je moguće rješiti i Coppersmithovom metodom u slučaju nelinearnih jednadžbi. Rezultati dobiveni ovim metodama su jači kada se koriste rešetke većih dimenzija, stoga njihova učinkovitost ovisi o mogućnostima računala. No i sa rešetkama male dimenzije dobivaju se dobri rezultati.

2.3.2. Rješavanje linearne jednadžbe

Ako je poznato da linearna multivarijantna jednadžba ima malo rješenje, tada ga možemo pronaći koristeći **heurističku metodu baziranu na rešetkama** koja pronalazi najkraći vektor u rešetki. Metoda je heuristička jer se zasniva na nedokazanoj pretpostavci najkraćeg vektora u rešetki.

Želimo doći do rješenja linearne jednadžbe oblika

$$Ax + By + Cz = w,$$

gdje su $x, y, z, w \in \mathbb{Z}$ nepoznate veličine, te x, z, w su male vrijednosti. Jednadžbu možemo zapisati u matričnom obliku $\vec{u}\mathcal{B} = \vec{v}$, pri čemu koristimo trivijalne jednadžbe $x = x$ i $z = z$

$$\begin{bmatrix} x & z & y \end{bmatrix} \begin{bmatrix} 1 & 0 & A \\ 0 & 1 & C \\ 0 & 0 & B \end{bmatrix} = \begin{bmatrix} x & z & w \end{bmatrix}. \tag{2}$$

Matrica $\mathcal{B}$ je matrica baze za neku rešetku $\mathcal{L}$ jer su njeni retci linearno nezavisni. Vektor $\vec{v} = [x, z, w]$ je linearna kombinacija redaka bazne matrice te je on vektor iz rešetke.

Ako je vektor $\vec{v}$ (i $-\vec{v}$) jedini najkraći vektori u rešetki, onda rješavamo linearnu jednadžbu tako da rješimo problem najkraćeg vektora u rešetki. To se može efikasno napraviti za male rešetke. Nakon pronalaska vektora $\vec{v}$, rješavamo matričnu jednadžbu $\vec{u}\mathcal{B} = \vec{v}$ i time dobivamo x, y, z, w.

Postoje dva kriterija na temelju kojih možemo zaključiti je li vektor $\vec{v}$ kandidat za najkraći vektor u rešetki $\mathcal{L}$:

1. vektor $\vec{v}$ mora biti najkraći bazni vektor u matrici baze. Neka je $X = \max\{|x|, |z|, |w|\}$. Kako bi vektor zadovoljavao $||\vec{v}|| \leq \sqrt{3}X$, prvi nužan uvjet da $\vec{v}$ bude najkraći vektor je
$$X \leq |A|, |B|, |C|.$$

2. vektor mora zadovoljavati Minkowskijevu granicu
$$||\vec{v}|| \leq \sqrt{3}\text{vol}(\mathcal{L})^{1/3}.$$
Koristeći $||\vec{v}|| \leq \sqrt{3}X$ i $\text{vol}(\mathcal{L}) = |\det \mathcal{B}| = |B|$, slijedi
$$X \leq |B|^{1/3}.$$

Ako vektor zadovoljava oba kriterija, onda je možda najkraći vektor u rešetki. Nadamo se da je on najmanji vektor i da ga možemo otkriti rješavanjem SVP problema u rešetki.

Pretpostavka 2.1. *Neka je $\mathcal{B}$ baza rešetke $\mathcal{L}$ i $v \in \mathcal{L}$. Ako je vektor v najkraći od svih vektora u bazi $\mathcal{B}$, te ako zadovoljava Minkowskijevu granicu za rešetku $\mathcal{L}$, onda je $\pm v$ jedini najkraći vektor u rešetki $\mathcal{L}$.*

Često se dogodi da vektor $\vec{v}$ ima komponente koje nisu uravnotežene i tada se povećaju granice ove metode. Primjerice, neka su X, Z, W redom gornje granice za x, z, w, te neka je $Z > X, W$, iz čega zaključujemo da je komponenta z najveća. Možemo matričnu jednadžbu (2) pomnožiti zdesna dijagonalnom matricom
$$\mathcal{D} = \begin{bmatrix} Z/X & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & Z/W \end{bmatrix}$$
i tako dobivamo novu jednadžbu $\vec{x}\mathcal{B}\mathcal{D} = \vec{x}\mathcal{B}' = \vec{v}'$ oblika
$$\begin{bmatrix} x & z & y \end{bmatrix} \begin{bmatrix} Z/X & 0 & AZ/W \\ 0 & 1 & CZ/W \\ 0 & 0 & BZ/W \end{bmatrix} = \begin{bmatrix} xZ/X & z & wZ/W \end{bmatrix}.$$

Novi vektor $\vec{v}'$, koji je vektor nove rešetke $\mathcal{L}'$, ima uravnotežene komponente, te vrijedi
$$||\vec{v}'|| \leq ((XZ/X)^2 + Z^2 + (WZ/W)^2)^{1/2} \leq \sqrt{3}Z.$$

Volumen nove rešetke $\mathcal{L}'$ se lako računa zbog konstrukcije matrice $\mathcal{D}$, a time i $\mathcal{B}'$
$$\text{vol}(\mathcal{L}') = |\det(\mathcal{B}')| = \frac{BZ^2}{XW},$$
te zaključujemo da se volumen povećao, kao i veličina vektora baze.
Dakle, kada vektor ima komponente koje nisu uravnotežene možemo koristiti ovu metodu kojom povećavamo granice nepoznatih vrijednosti, a da pri tome vektor zadovoljava svojstva

najkraćeg vektora.

Modularna linearna jednadžba

Prethodne ideje možemo primjeniti na rješavanje modularne linearne multivarijatne jednadžbe. Analogno pretpostavimo da želimo naći malo rješenje jednadžbe

$$Ax + By + Cz \equiv w \pmod{N},$$

odnosno

$$Ax + By + Cz = w + nN,$$

gdje je $n \in \mathbb{Z}$ nepoznata vrijednost. Na ovaj način dobivamo novu linearnu jednadžbu s još jednom nepoznanicom. Ideju iz prethodnog poglavlja primjenjujemo na novu jednadžbu, te konstruiramo matričnu jednadžbu

$$\begin{bmatrix} x & y & z & -n \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & A \\ 0 & 1 & 0 & B \\ 0 & 0 & 1 & C \\ 0 & 0 & 0 & N \end{bmatrix} = \begin{bmatrix} x & y & z & w \end{bmatrix},$$

i tražimo (x, y, z, w) najkraći vektor u rešetki koji je linearna kombinacija redaka bazne matrice.

Sljedeća tvrdnja se dugo vremena koristila za rješavanje modularne linearne jednadžbe, a da pri tome nije bila dokazana.

Tvrdnja 2.1. *Za modularnu linearnu jednadžbu oblika*

$$a_1x_1 + \cdots + a_nx_n \equiv 0 \pmod{N}, \tag{3}$$

gdje su a_i i N poznate veličine, te vrijedi $|x_i| \leq X_i$ pri čemu je cijeli broj $X_i > 0, \forall 1 \leq i \leq n$, je poznato da ima malo rješenje. Rješenja $x_1 \ldots, x_n$ se mogu izračunati kada vrijedi

$$\prod_{i=1}^{n} X_i \leq N.$$

Iako je ovaj rezultat odavno poznat, tek su 2008. godine njegov dokaz na temelju ove metode dali Herrmann i May.

Dokaz. Neka je $(a_n, N) = 1$. Jednadžbu (3) množimo s $-a_n^{-1} \mod N$ i dobivamo

$$b_1x_1 + \cdots + b_{n-1}x_{n-1} \equiv x_n \pmod{N},$$

pri čemu je $b_i = -a_n^{-1}a_i \mod N$ za $i = 1, \ldots, n$. Dobivenu kongruenciju zapišemo u obliku jednadžbe

$$b_1x_1 + \cdots + b_{n-1}x_{n-1} = -x_n - nN,$$

za neki cijeli broj n, te dobivamo sljedeću matričnu jednadžbu

$$\begin{bmatrix} x_1 & \cdots & x_{n-1} & n \end{bmatrix} \begin{bmatrix} N/X_1 & 0 & \cdots & 0 & b_1 N/X_1 \\ 0 & N/X_2 & \cdots & 0 & b_2 N/X_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & N/X_{n-1} & b_{n-1} N/X_{n-1} \\ 0 & 0 & \cdots & 0 & N^2/X_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_{n-1} & x_n \end{bmatrix}.$$

Vektor

$$\vec{v}' = \left(\frac{x_1 N}{X_1}, \cdots, \frac{x_{n-1} N}{X_{n-1}}, \frac{x_n N}{X_n} \right)$$

je vektor rešetke $\mathcal{L}'$. Da bi vektor $\vec{v}'$ bio najkraći vektor u rešetki, potrebno je da zadovoljava Minkowskijevu granicu iz teorema 2.5.
Granica svake komponente vektora $\vec{v}'$ je N, odnosno $||\vec{v}'|| \leq \sqrt{n} N$.
Volumen rešetke $\mathcal{L}'$ je produkt dijagonalnih elemenata, odnosno $\text{vol}(\mathcal{L}') = N^{n+1} \prod_{i=1}^{n} \frac{1}{X_i}$.
Uvrštavanjem prethodne dvije jednakosti u Minkowskijevu granicu dobivamo da je vektor $\vec{v}'$ najkraći vektor u rešetki $\mathcal{L}'$ ako vrijedi

$$\sqrt{n} N \leq \sqrt{n} \left(N^{n+1} \prod_{i=1}^{n} \frac{1}{X_i} \right)^{1/n},$$

odnosno

$$\prod_{i=1}^{n} X_i \leq N.$$

□

2.3.3. Coppersmithova metoda

Za nelinearne jednadžbe postoje dokazani rezultati koji pronalaze mala rješenja pojedinih tipova jednadžbi. Coppersmithova metoda se može koristiti za traženje malog korijena bivarijantnog polinoma nad $\mathbb{Z}$ kao i za traženje malog korijena univarijantnog polinoma nad $\mathbb{Z}_N$. Ideja koja se krije iza ove metode je transformirati modularnu diofantsku jednadžbu u jednadžbu na skupu $\mathbb{Z}$.

Ovu metodu je moguće proširiti za polinome s više varijabli, no ona se temelji na nedokazanim pretpostavkama, stoga su proširenja heuristička. Unatoč tome, metoda u praksi radi jako dobro.

Neka je N cijeli broj čiju faktorizaciju ne znamo, te f_N normirani polinom jedne varijable s cjelobrojnim koeficijentima stupnja d

$$f_N(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{d-1} x^{d-1} + x^d \in \mathbb{Z}[x] \tag{4}$$

za kojeg je poznato da postoji malo rješenje. Potrebno je efikasno naći rješenje $x_0 \leq X$ za koji je

$$f_N(x_0) \equiv 0 \pmod{N}$$

za što veću moguću granicu X.

Za $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + a_nx^n$ i $X > 0$ definiramo Euklidsku normu od $f(xX)$ na sljedeći način

$$||f(xX)|| = \Big(\sum_{i=0}^{n} |a_iX^i|^2\Big)^{1/2}.$$

Teorem 2.7 (**Howgrave - Graham teorem**). *Neka je $h(x) \in \mathbb{Z}[x]$ polinom stupnja ω, te $X > 0$. Pretpostavimo da vrijedi $||h(xX)|| < \frac{N}{\sqrt{\omega}}$. Ako $|x_0| < X$ zadovoljava $h(x_0) \equiv 0 \pmod N$, onda je x_0 cjelobrojni korijen od $h(x)$, odnosno $h(x_0) = 0$.*

Dokaz. Neka je $h(x) = h_0 + h_1x + \cdots + h_\omega x^\omega$. Tada vrijedi

$$\begin{aligned} |h(x_0)| &\leq |h_0| + |h_1X| + \cdots + |h_\omega X^\omega| \\ &\leq \sqrt{\omega} \cdot \sqrt{h_0^2 + h_1^2X^2 + \cdots + h_\omega^2X^{2\omega}} \\ &= \sqrt{\omega} \cdot ||h(xX)|| \\ &< \sqrt{\omega} \cdot \frac{N}{\sqrt{\omega}} = N \end{aligned}$$

pri čemu smo koristili nejednakost trokuta, Cauchy-Schwarz nejednakost, te pretpostavku teorema $||h(xX)|| < \frac{1}{\sqrt{\omega}}N$.
Iz pretpostavke $h(x_0) \equiv 0 \pmod N$ i dobivene nejednakosti $|h(x_0)| < N$ slijedi da je x_0 cjelobrojni korijen od $h(x)$. □

Dakle, osnovna ideja je konstrirati novi polinom $h(x)$ koji će imati isti korijen kao i $f(x)$, ali male koeficijente. Tada se kongruencija modulo N zamjeni običnom jednakošću, a problem se rješava nalaženjem cjelobrojnih nultočaka novog polinoma.

Coppersmith je pokazao da se mogu naći korijeni manji od $N^{1/d-\epsilon}$ na način da se konstruira novi polinom kao kombinacija polinoma $f_N(x)$ modulo N^m za neki cijeli broj m.

Teorem 2.8 (**Coppersmithov teorem**). *Neka je N cijeli broj čiju faktorizaciju ne znamo, te $b \geq N^\beta$. Neka je $f_b(x)$ normiran polinom jedne varijable sa stupnjem d i neka je $c > 1$ konstanta. Korijeni x_0 za koje vrijedi $f_b(x_0) \equiv 0 \pmod b$ i $|x_0| \leq cN^{\beta^2/d}$ se mogu naći u polinomijalnom vremenu $\log(N)$, c i broju korijena.*

Neka je $f_N(x)$ polinom čiji je korijen x_0 modulo N. Za neki fiksni m i t definiramo polinome

$$f_{(i,j)}(x) = x^i f_N(x)^j N^{m-j}, \qquad j = 0, \ldots, m,\ i = 0, \ldots, t-1.$$

Budući da x_0 zadovoljava $f_N(x_0) \equiv 0 \pmod N$, onda slijedi i da je $f_N(x_0)^k \equiv 0 \pmod{N^k}$ za $k \geq 1$. Stoga vrijedi i $f_{(i,j)}(x) \equiv 0 \pmod{N^m}$ gdje je $0 \leq i < d,\ 0 \leq j < m$.

Svaka cjelobrojna linearna kombinacija polinoma $f_{(i,j)}(x)$ ima korijen x_0 modulo N^m. Cilj nam je pronaći polinom

$$h(x) = \sum_{i=0}^{d-1} \sum_{j=0}^{m} a_{i,j} f_{(i,j)}(x), \qquad a_{i,j} \in \mathbb{Z},$$

koji će zadovoljavati Howgrave-Graham teorem 2.7, a zatim pronaći rješenje polinoma $h(x)$ u skupu $\mathbb{Z}$. Uočimo da vrijedi da je stupanj polinoma $f_{(i,j)}(x)$ manji ili jednak $dm + t - 1$.

Neka je $\omega \geq (m+1)d - 1$. Potrebno je pažljivo odabrati ω polinoma $f_{(i,j)}(xX)$, kako bi bazna matrica $\mathcal{B}$ bila trokutasta $\omega \times \omega$ matrica, a time rešetka $\mathcal{L}$ punog ranga. Konstruiramo matricu baze tako da retci matrice odgovaraju koeficijentima polinoma $f_{(i,j)}(xX)$ za $j = 0, \ldots, m$ i $i = 0, \ldots, d-1$.

$$\begin{array}{c} \\ f_{(0,0)} \\ \vdots \\ f_{(d-1,0)} \\ \vdots \\ f_{(0,j)} \\ \vdots \\ f_{(d-1,j)} \\ \vdots \\ f_{(0,m)} \\ \vdots \\ f_{(d-1,m)} \end{array} \begin{array}{c} \begin{array}{cccccccccccc} 1 & \ldots & x^{d-1} & \ldots & x^{d_j} & \ldots & x^{(j+1)d-1} & \ldots & x^m & \ldots & x^{(m+1)d-1} \end{array} \\ \left[\begin{array}{cccccccccccc} N^m & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & N^m X^{d-1} & & & & & & & & \\ & & & \ddots & & & & & & & \\ * & \ldots & * & * & N^{m-j}X^d & & & & & & \\ * & \ldots & * & * & * & \ddots & & & & & \\ * & \ldots & * & * & * & * & N^{m-j}X^{(d+1)j-1} & & & & \\ * & \ldots & * & * & * & * & * & \ddots & & & \\ * & \ldots & * & * & * & * & * & * & X^{dm} & & \\ * & \ldots & * & * & * & * & * & * & * & \ddots & \\ * & \ldots & * & * & * & * & * & * & * & * & X^{(m+1)d-1} \end{array}\right] \end{array}$$

Primjenom teorema 2.6, gdje najkraći vektor u reduciranoj bazi odgovara koeficijentima polinoma $h(xX)$ s korijenom x_0 modulo N, dobivamo

$$||h(xX)|| \leq 2^{(\omega-1)/4} \mathrm{vol}(\mathcal{L})^{1/\omega},$$

gdje je $\mathrm{vol}(\mathcal{L}) = N^{m(m+1)d/2} X^{\omega(\omega-1)/2}$. Prema teoremu 2.7

$$||h(xX)|| < \frac{N^m}{\sqrt{\omega}},$$

gdje je $\omega = d(m+1)$. Kombiniranjem prethodne dvije nejednakosti zaključujemo da ako je x_0 cjelobrojni korijen polinoma $h(x)$ vrijedi sljedeće

$$\mathrm{vol}(\mathcal{L}) \leq \gamma N^{m\omega}, \tag{5}$$

za $\gamma = 2^{-\omega(\omega-1)/4} \omega^{-\omega/2}$. Dobivena nejednakost se smatra nužnim uvjetom za primjenu Coppersmithove metode. Ako je nužan uvjet zadovoljen, onda su svi korijeni $|x_0| < X$ polinoma $f_N(x)$ modulo N ujedno i korijeni $h(x)$.

Sljedećim teoremom dan je poseban slučaj Coppersmithove metode kada je $b = N$ i $c = 1$.

Teorem 2.9. *Neka je N cijeli broj čiju faktorizaciju ne znamo, te neka je $f_N(x)$ normirani polinom jedne varijable stupnja d. Postoji algoritam koji pronalazi rješenje kongruencije $f_N(x_0) \equiv 0 \pmod{N}$ takvo da je $|x_0| \leq N^{1/d-\epsilon}$. Složenost tog algoritma je polinomijalna $\log N$ i $1/\epsilon$.*

Coppersmithova metoda koju smo opisali za modularne polinome s jednom varijablom mogu se proširiti na modularne polinome s više varijabli. U nastavku se nalaze generalizacije prethodnih rezultata.

Teorem 2.10 (**Generalizacija Howgrave-Graham teorema**). *Neka je* $h(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ *polinom stupnja* ω, *te* $X_i > 0$, *za* $1 \leq i \leq n$. *Ako za* $(y_1, \ldots, y_n) \in \mathbb{Z}^n$, *gdje su* $|y_i| < X_i, 1 \leq i \leq n$, *vrijedi* $h(y_1, \ldots, y_n) \equiv 0 \pmod{N}$ *i*

$$||h(x_1X_1, \ldots, x_nX_n)|| < \frac{N}{\sqrt{\omega}},$$

onda je $(y_1, \ldots, y_n)$ *cjelobrojni korijen od* $h(x_1, \ldots, x_n)$, *odnosno* $h(y_1, \ldots, y_n) = 0$.

Ukoliko polinom zadovoljava uvjete teorema 2.10, onda kažemo da polinom zadovoljava **Howgrave-Graham granicu**.

Analogno prethodno obrađenom problemu za polinome s jednom varijablom, za polinome s n varijabli moramo pronaći n malih polinoma koji zadovoljavaju Howgrave-Graham granicu. Mali korijeni se pronalaze rješavanjem sustava jednadžbi, pri čemu uočavamo da polinomi moraju biti algebarski nezavisni kako bi izračunali korijen. Za dva polinoma kažemo da su algebarski nezavisni ako i samo ako rezultanta dva polinoma nije nula.

Rezultanta dva polinoma $f = f_0 + f_1x + \ldots + f_nx^n$ i $g = g_0 + g_1x + \ldots + g_mx^m$, $n, m > 0$ je determinanta $(m+n) \times (m+n)$ Sylvesterove matrice, odnosno

$$Res(f, g) = \det \begin{bmatrix} f_m & f_{m-1} & \cdots & f_0 & & & & \\ & f_m & f_{m-1} & \cdots & f_0 & & & \\ & & \ddots & \ddots & \ddots & \ddots & & \\ & & & & f_m & \cdots & \cdots & f_0 \\ g_n & g_{n-1} & \cdots & \cdots & g_0 & & & \\ & g_n & g_{n-1} & \cdots & \cdots & g_0 & & \\ & & \ddots & \ddots & \ddots & \ddots & & \\ & & & g_n & g_{n-1} & \cdots & \cdots & g_0 \end{bmatrix}.$$

Ukoliko koristimo Coppersmithovu tehniku onda moraju biti zadovoljene sljedeće dvije pretpostavke da bi metoda bila uspješna.

Pretpostavka 2.2. *Polinomi čiju malu nultočku želimo odrediti imaju jedinstvenu malu nultočku u skupu* $\mathbb{Z}$ *ili* $\mathbb{Z}_N$.

Pod pojomom "male nultočke" smatramo brojeve s kojima u praksi možemo računati, odnosno koje aktualna računala mogu koristiti u kriptosustavima.

Pretpostavka 2.3. *Polinomi dobiveni iz LLL-reducirane baze su algebarski nezavisni.*

2.4. Kriptologija

Kriptografija i kriptoanaliza zajedno čine granu zananosti koja se naziva **kriptologija**. Kriptoanaliza se razvijala u skladu s kriptografijom i zapravo su dvije suprotne strane unutar jedne znanosti. Kako se povećavala kompleksnost kriptografije, tako ju je kriptoanaliza pratila, te je danas nezamisliva bez računala. Da bi kriptografija bila sigurna, potrebno je da se prilagodi mogućim kriptoanalizama.

U nastavku su navedeni osnovni pojmovi unutar ove znanosti.

2.4.1. Kriptografija

Znanstvena disciplina koja se bavi proučavanjem metoda sigurne komunikacije koja se odvija u nesigurnom komunikacijskom kanalu naziva se **kriptografija** (grč. kriptós - skriven i gráfo - pisati).

Dva subjekta žele međusobno razmjenjivati poruke u nesigurnom komunikacijskom kanalu. Svjesni su da u tom kanalu svatko može pristupiti njihovim porukama te ih pročitati, stoga žele stvoriti sigurnije okruženje kako bi mogli nesmetano slati poruke. Subjekti između kojih se odvija komunikacija nazivaju se pošiljatelj i primatelj. Pošiljatelj želi poslati poruku koja se naziva **otvoreni tekst** (eng. plaintext) tako da ga transformira pomoću unaprijed dogovorenog ključa. Dobiveni tekst se naziva **šifrat** (eng. ciphertext), a postupak transformacije po unaprijed dogovorenom pravilu ili ključu šifriranje. Pošiljatelj šalje šifrat primatelju nesigurnim komunikacijskim kanalom gdje mu svatko može pristupiti, pa tako i protivnik. Ukoliko šifrat dođe u protivnikove ruke, s poteškoćama će pokušati otkriti otvoreni tekst. S druge strane, primatelj je upućen u postupak šifriranja, te će na jednostavniji način šifrat dešifrirati korištenjem unaprijed dogovorenog ključa i tako dobiti otvoreni tekst koji mu je poslao pošiljatelj.

Dakle, možemo reći da je kriptosustav skup svih poruka, šifrata, ključeva, te kriptografskog algoritma, dok je šifra uređeni par funkcije šifriranja i funkcije dešifriranja koje ovise o unaprijed dogovorenom parametru, odnosno ključu. Matematički definiran pojam kriptosustava se nalazi u nastavku.

Definicija 2.5. *Kriptosustav je uređena petorka* $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ *gdje je* $\mathcal{P}$ *konačan skup svih otvorenih tekstova,* $\mathcal{C}$ *konačan skup svih šifrata,* $\mathcal{K}$ *konačan skup svih mogućih ključeva,* $\mathcal{E}$ *skup svih funkcija šifriranja, te* $\mathcal{D}$ *skup svih funkcija dešifriranja.*
Za svaki ključ $K \in \mathcal{K}$ *postoji* $e_K \in \mathcal{E}$ *i odgovarajući* $d_k \in \mathcal{D}$*, gdje za funkcije* $e_K \colon \mathcal{P} \to \mathcal{C}$*,* $d_K \colon \mathcal{C} \to \mathcal{P}$ *vrijedi* $d_K(e_K(x)) = x$*, za svaki* $x \in \mathcal{P}$*.*

Klasifikacija kriptosustava obzirom na tajnost ključeva je sljedeća:

- **Kriptosustave s tajnim ključem** gdje pošiljatelj i primatelj tajno odabiru ključ K te na temelju njega generiraju funkcije za šifriranja i dešifiranje. Zbog korištenja istog ključa za šifriranje i dešifriranje, moguće je iz poznavanja funkcije šifriranja lako dobiti funkciju dešifriranja i obrnuto. Stoga ovaj tip kriptosustava nazivamo još i **simetrični**

kriptosustavi.
Kod simetričnih kriptosustava potrebno je prije uspostave komunikacije sigurnim kanalom razmijeniti tajni ključ. Što dovodi do mane simetričnih kriptosustava, a to je situacija kada pošiljatelj i primatelj nisu u mogućnosti sigurnosnim kanalom razmijeniti tajni ključ. Vrlo važno je sačuvati tajnost ključeva, jer sigurnost simetričnih kriptosustava ovisi upravo o tome.

- **Kriptosustave s javnim ključem** gdje svaki korisnik ima dva ključa, jedan javni koji se koristi za šifriranje i drugi tajni za dešifriranje. Budući da je ključ za šifriranje javno objavljen svatko može pomoću njega šifrirati tekst, no samo osoba koja poznaje tajni ključ za dešifriranje može dobiveni šifrat trasnformirati u otvoreni tekst. Dakle, ključ koji se koristi za dešifrianje je tajan, dok je ključ za šifriranje javan bez narušavanja tajnosti ključa za dešifriranje. Budući da ključ za dešifriranjeje nije moguće otkriti u nekom razumnom vremenu iz poznavanja ključa za šifriranja, drugi naziv za opisane sustave su **asimetrični kriptosustavi**.

Ideja kriptosustava s javnim ključem predstavljena je 1976. godine radom *"New Directions in Cryptography"* autora **Whitfield Diffiea i Martin Hellmana**. Predložili su da se umjesto dotadašnjeg jednog ključa, postupak ekripcije i dekripcije obavlja parom različitih ključeva od kojih je jedan javan, a drugi tajan. Na temelju javnog ključa generira se funkcija enkripcije koja je javna i jednostavna za izračunati, dok se odgovarajuća funkcija dekripcije kreira na temelju tajnog ključa. Funkcija dekripcije mora biti lako rješiva ukoliko znamo tajni ključ, no ako to nije slučaj, odnosno ukoliko imamo samo javno poznate podatke tada ju je komplicirano i neisplativo rješavati.

Neka je K skup svih mogućih korisnika. Funkcije šifriranja e_K i dešifriranja d_K kriptosustava s javnim ključem trebaju zadovoljavati sljedeća svojstva:

1. $\forall K$ vrijedi $d_K = e_K^{-1}$,
2. $\forall K$ e_K je javna, a d_K je poznata samo korisniku K,
3. $\forall K$ e_K je osobna jednosmjerna funkcija.

e_K nazivamo **javnim**, a d_K **tajnim ključem**.

Promotrimo što znači treće svojstvo. Prethodno smo predstavili funkciju šifriranja na način da ju je lako izračunati, dok bi nam računanje inverza bio problem. Stoga tražimo da je funkcija šifriranja e_K **jednosmjerna funkcija**, odnosno jednostavno možemo izračunati vrijednost u pojedinoj točki, dok iz zadane vrijednosti teško nalazimo točku. Matematički definirano, funkcija f je jednosmjerna funkcija (eng. one-way function) ako za svaki argument x iz domene funkcije f je jednostavno izračunati odgovarajuću vrijednost $f(x)$, te gotovo za svaki y iz slike f je potpuno neučinkovito računati jednadžbu $y = f(x)$ za neki argument x.

Ukoliko se inverz jednosmjerne funkcije lako računa zbog poznavanja dodatne tajne informacije (eng. trapdoor), tada funkciju f nazivamo **osobna jednosmjerna funkcija** (eng. one-way trapdoor function). Poznavanje tajnog ključa je trapdoor informacija pomoću kojeg se inverz funkcije šifriranja e_K računa u razumnom vremenu što je prema svojstvu 1. upravo funkcija dešifriranja d_K.

Diffie-Hellman protokol za razmjenu ključeva

Diffie i Hellman unijeli su novosti u svijet kriptografije nakon što su uočili da je u nekim grupama potenciranje puno jednostavnije od logaritmiranja i time započeli kriptografiju javnog ključa. Za svoj rad [1] iz 1976. godine Diffie i Hellman su dobili *Turingovu nagradu*[1] za 2015. godinu.

Pogledajmo razmjenu ključeva između Anje i Borisa u nesigurnom javnom komunikacijskom kanalu. Pretpostavimo da su javnim kanalom odabrali dva velika broja p i g takvi da su $(p, g) = 1$.

- Anja generira slučajan prirodan broj $a \in \{1, 2, \ldots, p-1\}$, te šalje Borisu $A = g^a \pmod p$
- Boris generira slučajan prirodan broj $b \in \{1, 2, \ldots, p-1\}$, te šalje Anji $B = g^b \pmod p$
- Anja računa tajni ključ $B^a \pmod p = (g^b)^a \pmod p = g^{ab} \pmod p$
- Boris računa tajni ključ $A^b \pmod p = (g^a)^b \pmod p = g^{ab} \pmod p$
- Dobili su isti tajni ključ $K = g^{ab}$ jer vrijedi
  $B^a \pmod p = (g^b)^a \pmod p = g^{ab} \pmod p = (g^a)^b \pmod p = A^b \pmod p$.

Uočavamo da su javno objavljeni podaci: p, g, A i B. Pretpostavimo da Cvjetko želi saznati njihov tajni ključ K. Kako bi rješio taj problem, treba iz poznatih podataka g i g^a saznati a, odnosno problem svodi na problem diskretnog logaritma $\log_g g^a \equiv a \pmod p$ koji nema jedinstveno rješenje. Kada Cvjetko sazna tajni element a, može bez poteškoća izračunati tajni ključ.

Pozadina problema je u tome da je G ciklička grupa reda p, a g je generator grupe. Stoga se unutar grupe može lako obaviti diskretno potenciranje, dok se diskretni logaritam u cikličkoj grupi teže računa (moguće je napraviti u eksponencijalnom vremenu).

Ukratko objasnimo kako bi izgledala komunikacija između Anje i Borisa. Oboje imaju svoj javni i tajni ključ. Pretpostavimo da Boris želi poslati Anji poruku x. U tom slučaju će iskoristiti Anjin javno objavljeni ključ e_A i pomoću njega kriptirati svoju poruku. Anja

[1] je godišnja nagrada koju dodjeljuje Association for Computing Machinery (ACM) za osobu koja je dala doprinos tehničko-računalnoj zajednici. Nagrada je dobila naziv po britanskom matematičaru Alanu Turingu koji je jedan od utemeljitelja modernog računalstva. Često se Turingovova nagrada smatra Nobelovom nagradom za računalstvo. Prva nagrada je dodjeljena 1966. godine.

od Borisa dobiva šifrat $y = e_A$ kojeg može dekriptirati svojim tajnim ključem d_A. Anja na ovaj način može pročitati otvoreni tekst koji joj je Boris poslao: $d_A(y) = d_A(e_A(x)) = x$.

Budući da je Anjin javni ključ javno objavljeni i tako dostupan svima, svatko može Anji poslati poruku koristeći njen javni ključ. Kako bi spriječili lažno predstavljanje, Diffie i Hellman su u radu [1] uveli pojam **digitalnog potpisa**. Boris dakle treba digitalno potpisati svoju poruku kako bi Anja bila sigurna da je dobivena poruka uistinu njegova. Boris, kao i u prethodnom primjeru, javno objavljenim Anjinim ključem šifrira poruku $e_A(x) = y$. Prije samog slanja poruke Anji, Boris digitalno potpisuje šifrat $d_B(y) = d_B(e_A(x)) = z$. Budući da nitko osim Borisa ne zna njegov tajni ključ, taj korak govori Anji da je Boris pošiljatelj poruke. Na dobivenu poruku z Anja najprije primjeni javno dostupan Borisov ključ e_B, a zatim svoj tajni ključ d_A. Vrijedi: $d_A(e_B(z)) = d_A(e_B(d_B(e_A(x)))) = x$.

Napomenimo da je digitalni potpis snažniji od običnog potpisa, budući da digitalni potpis ovisi i o samoj poruci, dok se običan potpis može kopirati i zalijepiti na razne dokumente, te tako biti prikazan kao važeći.

Najveća je prednost asimetričnih kriptosustava je u tome što je jednostavno stvoriti sigurni komunikacijski kanal između osoba koje žele uspostaviti komunikaciju, gdje je distribucija ključeva rješena na elegantniji način nego kod simetričnih. No unatoč ovoj velikoj prednosti, asimetrični algoritmi nisu predviđeni za šifriranje podataka većeg opsega budući da su oni sporiji od simetričnih algoritama. Danas se zapravo za šifriranje podataka koriste simetrični sustavi, dok se asimetrični sustavi koriste za šifriranje ključa simetričnog algoritma.

2.4.2. Kriptoanaliza

Suprotnost kriptografiji je **kriptoanaliza** (grč. kryptós – skriven i analýein – odriješiti), znanstvena disciplina koja proučava metode otkrivanja šifrata bez znanja o tajnim informacijama potrebnim za dekriptiranje. Ideja je pronalazak ranjivih točaka pojedinih kriptosustava s ciljem otkrivanja tajnih informacija potrebnih za dekripciju.

Osnovna pretpostavka je da kriptoanalitičar zna koji je kriptosustav korišten, što je poznato kao *Kerckhoffsovo načelo*[2]. Ova pretpostavka zapravo kaže da je tajnost šifre u korištenom ključu, a ne odabranom kriptosustavu. Sigurnost ne treba biti u klimavoj pretpostavci o nepoznavanju korištenog kriptosustava, budući da kriptoanalitičar može provjeriti nekoliko mogućih kriptosustava, a da se pri tome složenost procedure znatno ne mijenja.

Kriptoanalitičke napade možemo klasificirati obzirom na informacije koje kriptoanalitičar ima na raspolaganju:

- **Poznat samo šifrat** (eng. ciphertext only attack (COA))
  Model napada u kojem kriptoanalitičar na temelju jednog ili više šifrata koji su šifrirani

[2]Auguste Kerckhoffs (1835-1903), nizozemski lingvist i kriptograf, autor *La Cryptographie Militaire*

istim algoritmom pokušava otkrit nepoznati otvoreni tekst ili ključ kojim je napravljeno šifriranje. Ovaj tip napada se često pojavljuje u praksi.
No, kriptoanalitičar zapravo ima neka saznanja o otvorenom tekstu, primjerice, poznato je koji jezik je korišten pri pisanju otvorenog teksta, statistička svojstva korištenog jezika, te vjerojatnost pojavljivanja pojedinih riječi.

- **Poznat otvoreni tekst** (eng. known plaintext attack (KPA))
  Kriptoanalitičar ima pristup šifratu i njemu odgovarajućem otvorenom tekstu, te na temelju toga otkriva ključ ili algoritam kojim je šifrirana poruka.

- **Odabran otvoreni tekst** (eng. chosen plaintext attack (CPA))
  Kriptoanalitičar može dobiti šifrat za neograničen broj odabranih otvorenih tekstova. Napad je jak, ali i manje realan.

- **Odabran šifrat** (eng. chosen ciphertext attack (CCA))
  Kriptoanalitičar poznaje funkciju dešifriranja, te za jedan ili više odabranih šifrata može saznati odgovarajući otvoreni tekst. Ovaj napad je tipičan za kriptosustave s javnim ključem, stoga se na temelju dobivenih informacija pokušava otkriti ključ koji je korišten za dešifriranje.

Prethodni napadi su **pasivnog karaktera** i ignoriraju postojanje stvarnog svijeta. U stvarnom svijetu ne možemo ignorirati **aktivne napade**, odnosno mogućnost ucjene, potkupljivanja, krađe i slično kako bi se otkrio otvoreni tekst. Aktivni napadi se vrlo često kombiniraju s "pravim" pasivnim napadima na kriptosustave.

3. RSA kriptosustav

Matematičari **Ron Rivest, Adi Shamir i Len Adleman** 1977. godine su kreirali prvi kriptosustav s javnim ključem koji je dobio naziv po inicijalima tvoraca – RSA kriptosustav.

Prva objava njihovog rada bila je u *Scientific American*[3] u kolumni *"Mathematical Games"* **Martina Gardnera** [5]. Naime, Gardner je pomno pripremao svoje kolumne mjesecima unaprijed, no kada je dobio rad navedenog trojca, kolumnu je u kratkom roku napisao i objavio, što nas upućuje na veličinu njihovog rada. „A new kind of cipher that would take millions of years to break" je postala jedna od najpoznatijih Gardnerovih kolumni.
Osnovna ideja je korištenje dva velika prosta broja p i q s najmanje 40 znamenka, te se izračuna njihov produkt N. Moguće je da N bude javno dostupan upravo zbog pretpostavke da će faktoriziranje N biti nesavladiv zadatak. Javan je i neparan broj e. Ukoliko netko želi poslati tajnu poruku m osobi koja je odabrala p i q treba pronaći ostatak koji m^e daje pri djeljenju s N i poslati ga primaocu. Jednostavan matematički trik dopušta osobi koja je primila c da sazna poruku m iz poznavanja faktorizacije N.
Kako bi dokazali uspješnost svoga rada, trojac je Gardneru dao 128-znamenkastu kodiranu poruku koja je koristila 129-znamenkast N sastavljen od tajnih 64-znamenkastih i 65-znamenkastih prostih brojeva p i q. Odabran $e = 9007$ je bio javan. Za otkrivanje izvorne poruke ponuđena je nagrada od 100$, no naslov članka je upućivao na to da će osvajanje nagrade biti vrlo težak zadatak. U Gardnerovoj kolumni se našao i citat od Edgar Allan Poe: *"Yet it may be roundly asserted that human ingenuity cannot concoct a cipher which human ingenuity cannot resolve."*. Izazov je riješen u travnju 1994.

9686	9613	7546	2206
1477	1409	2225	4355
8829	0575	9991	1245
7431	9874	6951	2093
0816	2982	2514	5708
3569	3147	6622	8839
8962	8013	3919	9055
1829	9451	5781	5154

Slika 1: *Poruka šifrirana RSA kriptosustavom objavljena u Gardnerovoj kolumni vrijedna 100$*

Trojac Rivest, Shamir i Adleman su dobitnici *Turingove nagrade* za 2002. godinu.

U nastavku ćemo matematički definirati RSA kriptosustav.

[3]časopis o znanosti, informacijskim tehnologijama i politici osnovan 1845. godine. Među autorima članaka nalaze se znanstvenici dobitnici Nobelovih nagrada.

3.1. Definicija RSA kriptosustava

Definicija 3.1 (**RSA kriptosustav**). *Neka je N produkt dva velika prosta broja p i q, $\mathcal{P} = \mathcal{C} = \mathbb{Z}_n$, te $\mathcal{K} = \{(N, p, q, e, d) \colon ed \equiv 1 \pmod{\varphi(N)}\}$ gdje je $\varphi(N) = (p-1)(q-1)$ Eulerova funkcija. Za svaki $K \in \mathcal{K}$ definiramo pravilo kriptiranja $e_K \colon \mathbb{Z}_N \to \mathbb{Z}_N$*

$$e_K(x) = x^e \mod N$$

i pravilo dekriptiranja $d_K \colon \mathbb{Z}_N \to \mathbb{Z}_N$

$$d_K(y) = y^d \mod N$$

za $x, y \in \mathbb{Z}_N$.

Produkt $N = pq$ se naziva **RSA modul**, p i q **RSA prosti brojevi**. Broj e iz prethodne definicije nazivamo **enkripcijski ili javni**, a d **dekripcijski ili tajni eksponent**. Par (N, e) se naziva **javni ključ**, a trojka (p, q, d) **tajni ključ**.

Iz kongruencije koju zadovoljavaju enkripcijski i dekripcijski eksponent

$$ed \equiv 1 \pmod{\varphi(N)}$$

dobivano RSA **ključnu jednadžbu**

$$ed = 1 + k\varphi(N),$$

za neki cijeli broj k.

Budući da je tajni eksponent definiran kao multipikativni inverz od javnog eksponenta modulo $\varphi(N)$ dobivamo dovoljan uvjet pravila dešifriranja koji omogućuje otkrivanje otvorenog teksta iz bilo kojeg šifrata.

Nužan uvjet je da su javni i tajni eksponent međusobno inverzni modulo Carmichaelova lambda funkcije $\lambda(N)$.

Definicija 3.2. ***Carmichaelova lambda funkcija*** *od N je najmanji broj takav da je*

$$a^{\lambda(N)} \equiv 1 \pmod{N},$$

pri čemu je cijeli broj a relativno prost s N.

Neka je N produkt potencija prostih brojeva $p_i^{\alpha_i}$, tada vrijedi

$$\lambda(N) = [\lambda(p_1^{\alpha_1}), \ldots, \lambda(p_k^{\alpha_k})] = [(p_1 - 1)p_1^{\alpha_1 - 1}, \ldots, (p_k - 1)p_k^{\alpha_k - 1}].$$

Vrijedi sljedeće:

$$\begin{aligned}\varphi(N) &= (p-1)(q-1)\\ &= (p-1,q-1)[p-1,q-1]\\ &= (p-1,q-1)\lambda(N),\end{aligned} \tag{6}$$

odnosno, $\varphi(N)$ je višekratnik od $\lambda(N)$, stoga u ključnoj jednadžbi možemo koristiti $\varphi(N)$. Uočimo da ako je N prost broj, onda su vrijednosti Carmichaelove lambda funkcije od N i Eulerove funkcija od N jednake.

U radu se u pojedinim primjerima promatra RSA kriptosustav u kojemu je privatni eksponent nekad definiran kao multiplikativni inverz tajnog eksponenta modulo $\lambda(N)$, a ponekad modulo $\varphi(N)$.

Također ćemo promatrati **uravnotežene proste brojeve** što podrazumijeva da su prosti brojevi RSA kriptosustava približno jednake veličine. Pretpostaviti ćemo da vrijedi sljedeće

$$4 < \frac{1}{2}N^{\frac{1}{2}} < p < N^{\frac{1}{2}} < q < 2N^{\frac{1}{2}} \tag{7}$$

što je ekvivalentno

$$p < q < 2p.$$

Za proste brojeve koji su uravnoteženi, Eulerova funkcija zadovoljava

$$\begin{aligned}|N-\varphi(N)| &= |N-(p-1)(q-1)|\\ &= |N-(N-p-q-1)|\\ &= |p+q-1|\\ &< 3N^{\frac{1}{2}}.\end{aligned}$$

Ukoliko uvedemo oznaku

$$s = N - \varphi(N) \tag{8}$$

onda vrijedi

$$s < 3N^{1/2}. \tag{9}$$

Modul N i Eulerova funkcija od N imaju približno jednu polovinu zajedničkih značajnih bitova. Posljedica toga je

$$\varphi(N) < N < 2\varphi(N). \tag{10}$$

Kada je jedan eksponent definiran kao multiplikativni inverz drugog eksponenta, za očekivati je da će veličina eksponenta biti približno najveće moguće veličine, te tada kažemo da je eksponent pune veličine. U RSA kriptosustavu eksponenti su definirani modulo $\varphi(N)$. Dakle, eksponent pune veličine imao bi približnu veličinu kao i Eulerova funkcija $\varphi(N)$, odnosno modul N, budući da su $\varphi(N)$ i N jednake veličine. Ukoliko su eksponenti definirani modulo $\lambda(N)$, budući da su $\lambda(N)$ i $\varphi(N)$ s velikom vjerojatnošću približno jednake veličine, očekuje se da je eksponent pune veličine približno jednake veličine kao i N.

3.2. Implementacija RSA kriptosustava

Ključeve RSA kriptosustava generiramo na sljedeći način:

1. **Generiranje velikih prostih brojeva**: velike proste brojeve p i q koji imaju najmanje 512 znamenaka (sigurnim se smatra ukoliko prosti brojevi imaju 1024 znamenaka)

2. **Modul RSA kriptosustava**: produkt je dva velika prosta broja $N = pq$

3. **Eulerova funkcija**: izračuna se Eulerova funkcija $\varphi(N) = (p-1)(q-1)$

4. **Javni ključ**: e odabiremo iz $[3, \varphi(N)\rangle$ pri čemu je e relativno prost s Eulerovom funkcijom od N, odnosno $(e, \varphi(N)) = 1$

5. **Tajni ključ**: zbog prethodne definicije javnog ključa, tajni je ključ multiplikativni inverz od e definiran modulo $\varphi(N)$. Tajni ključ d možemo efikasno odrediti koristeći prošireni Euklidov algoritam, odnosno, postoji cijeli broj a takav da je $e \cdot d + a \cdot \varphi(N) = 1$. Euklidov algoritam, kao i prošireni Euklidov algoritam trebaju $\mathcal{O}(\ln^2 N)$ operacija.

Postupak šifriranja

Za postupak šifriranja koristi se javni ključ (N, e), te za dani otvoreni tekst $m \in \mathbb{Z}_N$ modularnim potenciranjem dobivamo šifrat $c = m^e \mod N$. To efikasno možemo izračunati metodom "kvadriraj i množi" čiji je algoritam u nastavku.

Algorithm 3 Kvadriraj i množi

Input: m, $0 \leq e < N$ gdje je $e = \sum_{i=0}^{t} e_i 2^i$ binarni zapis broja e
$c = 1$
for $i = l-1, \ldots, 0$ **do**
    $c = c^2 \pmod N$
    if $e_i = 1$ **then**
        $c = c \cdot m \pmod N$
    end if
end for
Output: $c = m^e \pmod N$

Budući da je ukupan broj množenja manji ili jednak $2l$, ukupan broj operacija je $\mathcal{O}(\log e \cdot \log^2 N)$, što nam kaže da imamo polinomijalan algoritam za šifriranje.

Postupak dešifriranja

Algoritam dešifriranja koristi tajni ključ (p, q, d) te za dani šifrat $c \in \mathbb{Z}_N$ dobivamo otvoreni tekst $m \equiv c^d \mod N$.

Ukoliko nam je poznat tajni ključ, onda koristimo polinomijalan algoritam "kvadriraj i množi" kako bi došli do otvorenog teksta.

U nastavku je primjer u kojem je opisano kako se odvija komunikacija u kojoj je korišten RSA kriptosustav. U primjeru nisu odabrani veliki prosti brojevi, stoga se ovaj primjer RSA kriptosustava ne smatra sigurnim.

Primjer 3.1. *Kao i u prethodnim primjerima, komunikacija se odvija između Anje i Borisa. Boris želi poslati Anju poruku "Napad na RSA s malim tajnim eksponentom".*

Svaki od sudionika komunikacije ima svoj vlastiti javni i tajni ključ. Pogledajmo na koji način je Anja generirala svoje ključeve. Anja za odabrane proste brojeve $p = 59$ i $q = 97$ računa njihov produkt $n = pq = 59 \cdot 97 = 5723$ te $\varphi(5723) = (59-1)(97-1) = 5568$. Potom bira $e = 2011$, te dobiva javni ključ $(5723, 2011)$ kojeg javno objavljuje.
Boris svako slovo svoje poruke pretvara u njegov numerički ekvivalent za engleskom alfabetu (A=00 , B=01, ..., Z=25), te dobiva sljedeće

$$m = 13001500031300171800181200110812190009130812041018151413041319141 2$$

Dijeli na četveroznamenkaste blokove krećući slijeva na desno.

$$(1300, 1500, 0313, 0017, 1800, 1812, 0011, 0812, 1900, 0913, 0812, 0410, 1815, 1413, 0413, 1914, 12)$$

Boris svaki blok šifrira prema pravilu $c = m^e \mod n$ i dobiva šifrat

$$\begin{aligned} c &= (0215, 4028, 5216, 2726, 0073, 3351, 1527, 2170, \\ &\quad 3909, 0553, 2170, 3373, 2406, 4022, 2242, 2408, 1667) \\ &= (02154028521627260073335115272170390905532170337324064022224224081667) \end{aligned}$$

kojeg šalje Anji.

Generiranjem ključeva Anja je za tajni eksponent dobila vrijednost $d = 2323$. Anjin tajni ključ je $(59, 97, 2323)$. Anja sada dijeli dobiveni šifrat od Borisa na blokove, te pravilom dekriptiranja $m = c^d \mod n$ dobiva otvoreni tekst "Napad na RSA s malim tajnim eksponentom".

$\diamond$

3.3. Sigurnost RSA kriptosustava

U prethodnom potpoglavlju smo vidjeli da je implementacija RSA kriptosustava jednostavna, a ako nam je poznat tajni eksponent d možemo efikasno provesti postupak dešifriranja.

Sigurnost RSA kriptosustava je u tome da je funkcija šifriranja jednosmjerna funkcija. Osoba koja je generirala svoj javni ključ, zna i svoj tajni ključ što je trapdoor informacija na temelju koje može pronaći inverz funkcije šifriranja, te time iz šifrata dobila otvoreni tekst. Napadač nema trapdoor informaciju te joj je neisplativo tražiti inverz funkcije šifriranja.

Trapdoor u RSA kriptosustavu je faktorizacija RSA modula $N = pq$ što prizlazi iz toga da je javni eksponent multiplikativni inverz tajnog eksponenta modulo $\varphi(N)$ koja je funkcija s parametrima p i q.

Kako je algoritam dekriptiranja vezan uz poznavanje faktorizacije modula $N = pq$ važno je sačuvati tajnost odabranih prostih brojeva p i q. Ovo nas upućuje da je upravo jedan od napada na RSA kriptosustav faktorizacija od N. Ukoliko napadač otkrije proste brojeve koji u produktu daju N tada nema prepreka u računanju $\varphi(N) = (p-1)(q-1)$, te konačno otkrivanju tajnog eksponenta d pomoću Euklidovog algoritma.

Jasno je da tajnost prostih brojeva nema veliku ulogu u sigurnosti RSA kriptosustava ukoliko su oni male veličine. U primjeru 3.1 ako želimo faktorizirati $N = 5723$ na način da dijelimo N sa svim prostim brojevima manjim od $\sqrt{N}$, vrlo brzo i lako se pronalaze odgovarajući prosti faktori. Postoje razni algoritmi za faktorizaciju stoga je važno da takvi algoritmi budu potpuno neučinkoviti, odnosno da računalima treba mnogo vremena za otkrivanje rješenja algoritma.

U slučajevima kada se komunikacija odvija između dva uređaja različitih snaga jedna od ideja kako bi se minimiziralo računanje slabijeg uređaja je odabir malog javnog i/ili tajnog eksponenta. Stoga razlikujemo napade na RSA kriptosustav s malim tajnim eksponentom i napade na RSA s malim javnim eksponentom. U svakom kriptosustavu primjenom grube sile možemo probiti sustav ako je privatni ili tajni ključ izabran iz malog skupa.

U nastavku ćemo pokazati kako je odabir malog tajnog eksponenta d u odnosu na N zapravo loša ideja, a započeti ćemo s Wienerovim napadom. Pri proučavanju napada na RSA kriptosustav zanemariti ćemo postojanje stvarnog fizičkog svijeta i promatrati ćemo RSA kriptosustav samo u matematičkom svijetu. To znači da fokus sigurnosti nije na ljudskoj komponenti gdje se manipulacijom otkrivaju tajni podaci ili na fizičkim svojstvima uređaja na kojima je RSA implementiran.

4. Napadi na RSA s malim tajnim eksponentom

Ako je javni ili tajni ključ odabran iz malog skupa moguće je primjenom grube sile razbiti svaki kriptosustav. Konkretno za RSA kriptosustav za koji privatni eksponent zadovoljava $d < 2^l$, pri čemu l ovisi o trenutnom stanju u računalu, moguće je pogoditi parametar d. Primjerice, moguće je otkriti sve tajne eksponente $d \leq 2^{60}$, dok je za $d \leq 2^{80}$ neisplativo. Napadi koji su obrađeni mogu otkriti mnogo veći tajni ekponent nego što je to moguće grubom silom. Točnije, obrađeni napadi mogu razbiti RSA kriptosustave kojima je tajni eksponent $d \leq N^{\delta}$, pri čemu je $\delta \geq \frac{1}{4}$.

4.1. Wienerov napad

Kroz ovo potpoglavlje pretpostavit ćemo da su javni i tajni eksponenti definirani modulo $\lambda(N)$. Iz ključne jednadžbe

$$ed = 1 + k\lambda(N)$$

slijedi

$$0 < k = \frac{ed-1}{\lambda(N)} < \frac{ed}{\lambda(N)} < \min\{e, d\}.$$

Budući da proučavamo RSA s malim tajnim eksponentom vrijedi da je $k < d$. Analogno vrijedi ako su eksponenti definirani modulo $\varphi(N)$.

Kriptolog **Michael J. Wiener** predstavio je 1990. godine algoritam koji u polinomijalnom vremenu koristeći verižne razlomke probija RSA kriptosustav s malim tajnim eksponentom. Algoritam je dobio naziv **Wienerov napad**, a predstavljen je sljedećim teoremom.

Teorem 4.1. *Neka je $N = pq$ RSA modul, e valjan javni eksponent, te d njegov odgovarajući tajni eksponent definiran modulo $\lambda(N)$. Neka je k cijeli broj koji zadovoljava $ed = 1 + k\lambda(N)$, $g = (p-1, q-1)$, $g_0 = \frac{g}{(g,k)}$ i $k_0 = \frac{k}{(k,g)}$. Ako tajni eksponent zadovoljava*

$$d < \frac{pq}{2(p+q-1)g_0k_0} = \frac{N}{2sg_0k_0}, \tag{11}$$

tada se N može faktorizirati u polinomijalnom vremenu u $\log(N)$ i $\frac{g}{k}$.

Dokaz. Za RSA modul $N = pq$ vrijedi

$$\lambda(N) = [p-1, q-1] = \frac{N-s}{g},$$

gdje smo koristili (6), (8) i $g = (p-1, q-1)$. Slijedi da se ključna jednadžba može zapisati u obliku

$$ed = 1 + k\lambda(N) = 1 + \frac{k}{g}(N-s) = 1 + \frac{\frac{k}{(g,k)}}{\frac{g}{(g,k)}}(N-s),$$

gdje je k neki prirodan broj. Ako uvedemo oznake $g_0 = \frac{g}{(g,k)}$ i $k_0 = \frac{k}{(k,g)}$, ključna jednadžba je oblika

$$ed = 1 + \frac{k_0}{g_0}(N-s).$$

Dijeljenjem obje strane jednadžbe s dN dobivamo

$$\left|\frac{e}{N}-\frac{k_0}{dg_0}\right| = \left|\frac{1}{dN}-\frac{k_0 s}{dg_0 N}\right| < \frac{k_0 s}{dg_0 N} < \frac{1}{2(dg_0)^2},$$

gdje prva nejednakost vrijedi jer je $g_0 < s$, a sve veličine su pozitivne, dok druga nejednakost vrijedi zbog (11).

Iz dobivene nejednakosti $\left|\frac{e}{N}-\frac{k_0}{dg_0}\right| < \frac{1}{2(dg_0)^2}$ zbog teorema 2.3 zaključujemo da je $\frac{k_0}{dg_0}$ neka konvergenta razvoja u verižni razlomak od $\frac{e}{N}$. Ukoliko je $c_i = \frac{a_i}{b_i}$ i-ta konvergenta od $\frac{e}{N}$, onda neka za neki j vrijedi $\frac{k_0}{dg_0} = \frac{a_j}{b_j}$.

Korištenjem (8) ključna jednadžba poprima oblik $ed = 1 + \frac{k_0}{g_0}\varphi(N)$, te dobivamo sljedeće

$$\varphi(N) = e\left(\frac{dg_0}{k_0}\right) - \frac{g_0}{k_0} = \left\lfloor e\left(\frac{b_j}{a_j}\right)\right\rfloor - \left\lfloor\frac{g_0}{k_0}\right\rfloor.$$

Možemo izračunati $\varphi(N)$ ako znamo točnu konvergentu c_j, a vrijednost $\lfloor\frac{g_0}{k_0}\rfloor$ možemo pogoditi. Postupak koji koristimo da izračunamo $\varphi(N)$ je sljedeći. Počevši od $m = 0$ tražimo kandidate za $\varphi(N)$ kroz iteracije konvergenti u razvoju verižnog razlomka od $\frac{e}{N}$ i računamo $\varphi' = \lfloor\frac{e}{c_i}\rfloor + m$. Ukoliko niti jedan kandidat ne daje $\varphi(N)$, odnosno ne faktorizira modul N, povećamo m za 1 i ponavljamo postupak. Na ovaj način na kraju prolazeći kroz kandidate $\varphi' = \lfloor\frac{e}{c_j}\rfloor + \lfloor\frac{g_0}{k_0}\rfloor$ dolazimo do faktorizacije modula.

Test za traženje faktorizacije modula N i kandidata za $\varphi(N)$ se može napraviti u polinomijalnom vremenu $\log(N)$. Budući da je ukupan broja konvergenti od $\frac{e}{N}$ reda $\mathcal{O}(\log N)$, a testiramo najviše $\lfloor\frac{g_0}{k_0}\rfloor = \lfloor\frac{g}{k}\rfloor$ kandidata za svaku konvergentu.

□

Dovoljan uvjet prethodnog teorema nije uobičajen uvjet koji uglavnom povezujemo s Wienerovim napadom. Za RSA kriptosustav s nasumično generiranim prostim brojevima i s malim privatnim eksponentom, odnosno uz pretpostavke da je javni eksponent otprilike iste veličine kao modul, da su prosti brojevi uravnoteženi, te da je g_0 mali, najčešći uvjet teorema 4.1 je

$$d < \frac{1}{c}N^{\frac{1}{4}},$$

za neku malu konstantu $c > 1$. Ova granica (grubo govoreći $d < N^{\frac{1}{4}}$) je referentna točka za Wienerov napad, te se ponekad naziva **Wienerova granica**.

RSA modul N kojeg smatramo sigurnim ima 1024 bita, stoga slijedi da ukoliko želimo da Wienerov napad ne bude učinkovit potrebno je da tajni eksponent d ima najmanje 256 bita.

Na sljedećem primjeru ilustrirati ćemo napad na RSA kriptosustav s malim privatnim eksponentom.

Primjer 4.1. *Neka je Borisov javni ključ* $(1778824363, 26030317)$. *Koristeći Borisov javni ključ Anja šifrira poruku i dobiveni šifrat šalje Borisu. Cvjetko je došao u posjed Anjine*

poslane poruke, te na temelju Borisovog javnog ključa pokušava dešifrirati poruku. Odlučio se iskoristiti Wienerov napad.

Cvjetko ima pristup javnom ključu, te mu se nameće ideja da promatra razvoj u verižni razlomak od $\frac{e}{N}$ u kojem se pojavljuje d. Budući da je $N \approx \varphi(N)$, moguće je da verižni razlomci $\frac{e}{N}$ i $\frac{e}{\varphi(N)}$ imaju neke zajedničke konvergente. Razvija $\frac{e}{N}$ u verižni razlomak i dobiva sljedeće

$$\frac{e}{N} = [0; 68, 2, 1, 32, 1, 19, 3, 1, 1, 1, 1, 2, 1, 2, 2, 3],$$

a pripadne konvergente su

$$0, \frac{1}{68}, \frac{2}{137}, \frac{3}{205}, \frac{37}{6697}, \frac{40}{6902}, \frac{797}{137835}, \ldots$$

Cvjetko potom za svakog kandidata računa $\varphi' = \left\lfloor \frac{e}{c_i} \right\rfloor + m$ i provjerava može li se za izračunati φ' faktorizirati modul N. Prvo postavlja vrijednost m na 0 i za prve tri konvergente ne uspjeva faktoritirati modul N.
Za konvergentu $c_4 = \frac{3}{205}$ računa $\varphi' = \left\lfloor \frac{e}{c_4} \right\rfloor + 0 = 1778738328$ i rješavanjem sustava jednadžbi

$$\begin{aligned} \varphi' &= (p-1)(q-1) \\ N &= pq \end{aligned}$$

dobiva proste brojeve $p = 34549$ i $q = 51487$.
Bez poteškoća saznaje tajni ključ $(34549, 51487, 205)$ pri čemu je $\lambda(N) = 296456388$, $g = 6$, $k = 18$. Uočavamo da je uvjet Wienerovog napada ispunjen jer je $d = 205 < 3445.93$.

Cvjetko dobivenim ključem dešifrira poruke koje je Anja poslala Borisu i ukoliko dobije suvislu poruku može sa sigurnošću reći da je razbio RSA kriptosustav.

$\diamond$

Postoji tri načina kojima se granica privatnog eksponenta može umanjiti i time napad oslabiti.

1. Korištenje prostih brojeva koji nisu uravnoteženi tako da $s = p + q - 1$ postane veliko.

2. Korištenje prostih brojeva s velikim $g = (p-1, q-1)$ tako da g_0 također postane veliko.

3. Korištenje javnog eksponenta $e > N$ tako da $k \approx ed/N$ i k_0 postane veliko. Konstrukcija javnog eksponent veći od modula jednostavno se dobiva dodavanjem višekratnika od $\lambda(N)$ u postojeći javni eksponent. Bez obzira koliko je d mali, Wienerov napad postaje potpuno neučinkovit kada je $e > N^{3/2}$. Jasno je da velika vrijednost e povećava vrijeme potrebno za šifriranje poruke.

Posljednji način utječe na jačinu Wienerov napada u oba smjera. Veći javni eksponent oslabljuje napad, dok ga manji javni eksponent ojačava.
Razmotrimo RSA sa uravnoteženim prostim brojevima, malim g_0, malim eksponentom $d = N^\delta < N^{1/2}$ i javnim eksponentom $e = N^\alpha$ za $1/2 < \alpha < 1$. Budući da je $ed = 1 + k\lambda(N)$

imamo $k \approx n^{\alpha+\delta-1}$. Uvrštvajući u dovoljan uvjet Wienerovog napada (11) i ignoriranjem malih konstanti dobivamo

$$N^{\delta} < N^{1-1/2-(\alpha+\delta-1)-\epsilon},$$

odnosno

$$\delta < \frac{3}{4} - \frac{\alpha}{2} - \epsilon,$$

gdje je $\epsilon > 0$ mala konstanta koja predstavlja sve male faktore koje smo ignorirali.
U tipičnom slučaju, kada je $e \approx N$, vrijedi $\alpha \approx 1$ i $\delta < 1/4$.
Za veće javne eksponente granica od δ se smanjuje sve dok ne nestane za $\alpha = 3/2$, te u tom slučaju napad ne može ništa jamčiti. Za manje javne eksponente granica se povećava sve do $\delta < 1/2$ za $\alpha = 1/2$.

4.1.1. Proširenje Wienerovog napada

Za RSA s malim privatnim eksponentom, uravnoteženim prostim brojevima, javnim eksponentom pune veličine $e \approx N$ i malim g_0 možemo upotrijebiti Wienerov napad ako javni eksponent d zadovoljava

$$d < \frac{1}{c} N^{\frac{1}{4}} = N^{\frac{1}{4}-\epsilon},$$

pri čemu je $\epsilon > 0$. Tada jedna od konvergenti u razvoju od $\frac{e}{N}$ u verižni razlomak sadrži dovoljno informacija za faktoriziranje modula.

Verheul i van Tilborg metoda

Verheul i van Tilborg smatraju da u slučaju da tajni eksponent prelazi Wienerovu granicu, odnosno kada je tajni eksponent nekoliko bitova veći od $N^{\frac{1}{4}}$, ipak postoje neke informacije koje se mogu dobiti iz konvergenata u razvoju od $\frac{e}{N}$ u verižni razlomak. 1997. godine su pokazali da postoje uzastopne konvergente, $c_j = \frac{a_j}{b_j}$ i $c_{j+1} = \frac{a_{j+1}}{b_{j+1}}$, takve da je

$$\frac{k_0}{dg_0} = \frac{xa_{j+1} + ya_j}{xb_{j+1} + yb_j}, \tag{12}$$

gdje su x i y nenegativni cijeli brojevi. Osnovna ideja Verheula i van Tilborga o proširenju Wienerovog napada je iscrpno pretraživanje x i y dok se ne pronađe $\frac{k_0}{dg_0}$.

No, **Dujella** je pokazao da nije potrebno među svim konvergentama provjeravati je li uvjet (12) zadovoljen, već se odabir može suziti na tri. Ako je l najveći neparan cijeli broj tako da konvergenta $c_l = \frac{a_l}{b_l}$ zadovoljava

$$\frac{a_l}{b_l} > \frac{e}{N} + \frac{2.121e}{N\sqrt{N}},$$

tada su odgovarajuće konvergente c_j, c_{j+1} gdje je $j \in \{l, l+1, l+2\}$.

Za odgovarajuće uzastopne konvergente preostaje odrediti nepoznate veličine x i y u (12). Primjenom grube sile tražimo odgovarajuće x i y. Za svakog para kandidata za x, y,

koristeći (12), računaju se kandidati za $\frac{k_0}{dg_0}$. Primjenom prethodno opisanog Wienerovog napada pokuša se faktorizirati modul izračunavanjem kandidata za $\varphi(N)$.

Ako za privatni eksponent vrijedi $d = DN^{1/4}$, onda postoji velika vjerojatnost da su x i y omeđeni s $4D$, odnosno $x, y < 4D$ (dokaz tvrdnje se nalazi u [2]). Budući da postoji $16D^2$ mogućnosti za kandidate x, y, imamo eksponencijalnu $\mathcal{O}(D^2)$ složenost. Za svakog kandidata, testiranje je moguće napraviti u polinomijalnom vremenu $\log(N)$.

Primjer 4.2. *(preuzeto iz [2]) Neka je dan javni ključ* $(7978886869909, 4603830998027)$. *Cvjetko je odlučio primjeniti prošireni Wienerov napad kako bi otkrio tajni ključ.*
Iz razvoja u verižni razlomak

$$[0, 1, 1, 2, 1, 2, 1, 18, 10, 1, 3, 3, 1, 6, 57, 2, 1, 2, 14, 7]$$

Cvjetko dobiva pripadne konvergente

$$0, 1, \frac{1}{2}, \frac{3}{5}, \frac{4}{7}, \frac{11}{19}, \frac{15}{26}, \frac{281}{487}, \frac{2825}{4896}, \ldots.$$

Uočava da je

$$\frac{11}{19} > \frac{e}{N} + \frac{2.121e}{N\sqrt{N}},$$

stoga zaključuje da tajni eksponent traži u obliku $26u + 19v$ ili $487x - 26y$ ili $4896w + 487z$. Otkriva tajni ključ $d = 5936963$ za $x = 12195$ i $y = 77$.

$\diamond$

Dujellina metoda

O aktualnosti napada na RSA kriptosustav govori da je hrvatski matematičar Andrej Dujella 2004. godine unaprijedio Verheul i van Tilborgovu metodu. Dujellina metoda faktorizira modul koristeći meet-in-the-middle tehniku kako bi se utvrdili x i y.

Radi jednostavnosti pretpostaviti ćemo da je poznat g_0. Ukoliko g_0 nije poznat, metodu možemo ponavljati s pogađanjem g_0 dok točan pogodak ne dovede do faktorizacije modula. Iz (12) uočavamo da točan par konvergenata daje $dg_0 = xb_{j+1} + yb_j$, te za svaki otvoreni tekst poruke m vrijedi

$$m^{e(xb_{j+1}+yb_j)} \equiv m^{edg_0} \equiv m^{g_0} \pmod{N}.$$

Prebacivanjem x i y na različite strane ekvivalencije dobivamo

$$(m^{eb_{j+1}})^x \equiv m^{g_0}(m^{-eb_j})^y \pmod{N}.$$

Za neki fiksirani otvoreni tekst poruke m, primjerice $m = 2$, definiraju se konstante X i Y kao $X = m^{eb_{j+1}} \mod N$ i $Y = (m^{eb_j})^{-1} \mod N$ tako da

$$X^x \equiv m^{g_0}Y^y \pmod{N}.$$

Ideja napada je prvo za svaki $0 \leq x' < 4D$ izračunati $X^{x'} \mod N$ i pohraniti ga u sortiranu listu. Potom, za svaki $0 \leq y' < 4D$ računati $m^{g_0} Y^{y'} \mod N$ i provjeravati nalazi li se dobivena vrijednost u spremljenoj listi. Ako se dogodi podudaranje, dobivene vrijednosti x' i y' su korištene u (12), te možemo računati kandidate za $\varphi(N)$ kako je opisano u Wienerovom napadu. Ako je $x' = x$ i $y' = y$ onda je izračunata točna vrijednost $\varphi(N)$ te je modul lako faktorizirati.

Ako su x i y manji od $4D$, složenost Dujelline metode je $\mathcal{O}(D \log D)$, gdje dominira stvaranje i sortiranje liste. Jasno je da je ovo napredak u odnosu na prethodnu Verheul i van Tilborg metodu koja radi na principu grube sile. Osnovna operacija koja se koristi za svakog kandidata, modularno potenciranje, je mnogo skuplja kod ove metode, ali na kraju je složenost mnogo bolja za dovoljno velik D.

Kao i kod svih meet-in-the-middle napada imamo prostorno-vremenski kompromis, te je za ovu metodu potrebno $\mathcal{O}(D)$ prostora za razliku metode grube sile koja zahtjeva $\mathcal{O}(1)$ prostora.

Napad efikasno radi za vrijednosti D do 2^{30}, odnosno

$$d < 2^{30} N^{\frac{1}{4}}.$$

4.2. Wienerov napad s rešetkama

Wienerov napad se, osim korištenja verižnih razlomaka, može prikazati na nekoliko različitih načina koristeći rešetke. U poglavlju 2.3. nalaze se osnovni pojmovi i rezultati potrebni za razumjevanje napada baziranih na rešetkama.

Metoda bazirana na heurističkom pristupu, te pristupu dokazivanjem koriste rešetke. Napadi Boneh i Durfee, iz sljedećeg poglavlja, su također metode bazirane na rešetkama.

Pretpostavka ovog poglavlja je da su prosti brojevi p i q uravnoteženi, odnosno da su prosti brojevi približno jednake veličine, te da su javni i tajni eksponent definirani modulo $\varphi(N)$.

4.2.1. Heuristički pristup

Prvo ćemo prikazati da se Wienerov napad može prikazati kao heuristički napad rešetkama koristeći metodu iz potpoglavlja 2.3.2.. Ovaj pristup se sastoji u tome da konstruiramo rešetku s malim vektorom koji sadrži informacije potrebne za faktoriziranje modula. Točnije, ukoliko je odabrani vektor najkraći vektor rešetke i ukoliko ga možemo pronaći, onda možemo faktorizirati modul. Koristeći Minkowskijevu granicu iz teorema 2.5 možemo naći dovoljan uvjet na veličinu privatnog eksponenta potrebnog za izvršenje napada.

Ključnu jednadžbu $ed = 1 + k(N - s)$ i trivijalnu jednadžbu $dN^{1/2} = dN^{1/2}$ zapisujemo

u matričnom obliku $x\mathcal{B}_1 = v$, pri čemu je $\mathcal{B}_1$ matrica baze rešetke $\mathcal{L}_1$, na sljedeći način

$$\begin{bmatrix} d & -k \end{bmatrix} \begin{bmatrix} N^{1/2} & e \\ 0 & N \end{bmatrix} = \begin{bmatrix} dN^{1/2} & 1-ks \end{bmatrix}.$$

Vidimo da je $v = (dN^{1/2}, 1-ks)$ cjelobrojna linearna kombinacija redaka bazne matrice, odnosno v je vektor u rešetki $\mathcal{L}_1$.

Potrebno je još pokazati da je v kratki vektor u $\mathcal{L}_1$. Rešetka $\mathcal{L}_1$ ima dimenziju $\dim(\mathcal{L}_1) = 2$ i volumen $\mathrm{vol}(\mathcal{L}_1) = N^{3/2}$. Iz teorema 2.5 znamo da norma najkraćeg vektora u $\mathcal{L}_1$ ne prelazi Minkowskijevu granicu $\sqrt{2}N^{3/4}$.
Uz pretpostavke $k < d$ i $s < 3\sqrt{N}$, norma vektora v zadovoljava

$$\|v\|^2 = (dN^{1/2})^2 + (1-ks)^2 < d^2N + k^2s^2 < d^2N + 9k^2N < 10d^2N.$$

Dovoljan uvjet da bi v zadovoljio Minkowskijevu granicu je dan s $\sqrt{10}dn^{1/2} < \sqrt{2}n^{3/4}$, odnosno jednostavnije

$$d < \frac{1}{\sqrt{5}}N^{1/4}.$$

Sada se nadamo da pretpostavka 2.1 vrijedi za $\mathcal{L}_1$. Odnosno, nadamo se da kada privatni eksponent d zadovoljava ovu granicu da je vektor v najkraći vektor u $\mathcal{L}_1$ te da su svi ostali vektori, osim $-v$, mnogo veći.

Najkraći vektor u dvodimenzionalnoj rešetki se pronalazi Gaussovim algoritmom (algoritam 1). Kada pronađemo vektor v, računa se vektor $x = (d, -k)$ iz $x\mathcal{B}_1 = v$. Konačno možemo izračunati Eulerovu funkciju $\varphi(N) = \frac{1}{k}(ed-1)$, a potom i faktorizirati modul N.

U potpoglavlju 2.3.2. vidjeli smo da je korisno u ovom tipu heurističkog napada da komponente traženog vektora budu uravnotežene. To smo napravili na način da umjesto trivijalne jednadžbe $d = d$ koristimo $dN^{1/2} = dN^{1/2}$ što dovodi do maksimiziranja volumena rešetke, a da pri tome ne mijenja normu traženog vektora v. Dolazi i do povećanja Minkowskijeve granice koju vektor v treba zadovoljavati.

Za odabir 1024-bitni RSA modul, napad nam sugerira da ako tajni eksponent zadovoljava

$$d < \frac{1}{\sqrt{5}}N^{\frac{1}{4}} \approx N^{0.2489}$$

tada bi RSA trebao biti nesiguran.

4.2.2. Pristup dokazivanjem

Uspjeh heurističkog napada zapravo nije bilo iznenađenje. May [9] je pristupio problemu iz druge perspektive, te dao dokaz napada baziranom na rešetkama. To je uspio nametanjem blagih ograničenja na proste brojeve RSA kriptosustava. Sljedećim teoremom je prikazan njegov napad.

Teorem 4.2. *Neka je $N = pq$ RSA modul s uravnoteženim prostim brojevima koji zadovoljavaju $p + q < \frac{3}{\sqrt{2}} N^{\frac{1}{2}}$. Neka je (N, e) valjan javni ključ, te d njegov odgovarajući javni eksponent definiran modulo $\varphi(N)$. Ako privatni eksponent zadovoljava*

$$d < \frac{1}{3} N^{\frac{1}{4}},$$

onda se modul može faktorizirati u polinomijalnom vremenu $\log(N)$.

Za izvođenje ovog napada, May pristupa problemu koristeći Coppersmithovu tehniku. Redukcijom ključne jednadžbe $ed = 1 + k(N - s)$ modulo N dobivamo

$$ed + ks - 1 \equiv 0 \pmod{N}.$$

Znamo da polinom $f_N(x, y) = ex + y$ modulo N ima mali korijen $(x_0, y_0) = (d, ks - 1)$, odnosno vrijedi $f_N(d, ks - 1) \equiv 0 \mod N$. Da bi razbili RSA, potrebno je naći korijen ovog polinoma modulo N.

Za $m = 1$ definiramo polinome $f_{1,0}(x, y) = Nx$ i $f_{0,1}(x, y) = f_N(x, y) = ex + y$ koji imaju korijen (x_0, y_0) modulo N. Polinom koji je linearna kombinacija polinoma $f_{1,0}(x, y)$ i $f_N(x, y)$, također ima korijen (x_0, y_0) modulo N. Neka je $k < d$, tada vrijedi

$$\begin{aligned} |x_0| &= d < \frac{1}{3} N^{1/4} \\ |y_0| &= |ks - 1| < ds < d(p + q) < \frac{1}{\sqrt{2}} N^{3/4}, \end{aligned}$$

te dobivamo granice $X = \frac{1}{3} N^{1/4}$ i $Y = \frac{1}{\sqrt{2}} N^{3/4}$.

Modularna jednadžba ima cjelobrojno rješenje koje možemo naći rješavanjem jednadžbe u skupu cijelih brojeva koristeći Howgrave-Graham teorem.

Konstruiramo rešetku $\mathcal{L}_2$ koristeći koeficijente polinoma $f_{1,0}(xX, yY)$ i $f_N(xX, yY)$, te dobivamo matricu baze

$$\mathcal{B}_2 = \begin{bmatrix} NX & 0 \\ eX & Y \end{bmatrix}.$$

Svi vektori iz rešetke $\mathcal{L}_2$ su cjelobrojna linearna kombinacija polinoma s korijenom (x_0, y_0) modulo N. May je pokazao da ako je $(c_0, c_1)\mathcal{B}_2 = v$ najkraći vektor u rešetki $\mathcal{L}_2$ onda vrijedi

$$(d, ks - 1) = (|c_1|, |c_0 N + c_1 e|).$$

Kratki vektor možemo efikasno naći Gaussovim algoritmom i tako doći do tajnog eksponenta d, a time i faktorizirati modul.

Složenost algoritma koji je kreirao May je $\mathcal{O}(\log^2(N))$ zbog Gaussovog algoritma i traženja najvećeg zajedničkog djelitelja kako bi se faktorizirao modul.

4.3. Boneh i Durfee napad rešetkama

U prethodnom poglavlju razmotrili smo verzije Wienerovog napada koju smo rješavali reduciranjem ključne jednadžbe modulo N. No Boneh i Durfee pokazali su da se mnogo jači napad može dobiti tako da se reducira ključna jednadžba modulo javni eksponent e. Odnosno, rješavanje

$$-k(N-s) \equiv 1 \pmod{e},$$

za nepoznat k i s, vodi do heurističkog napada na RSA s privatnim eksponentom $d < N^{0.292}$. Ovo je najjači poznati napad na RSA s malim tajnim eksponentom koji ne zahtjeva da prosti brojevi zadovoljavaju posebna svojstva.

U ovom poglavlju ćemo pretpostaviti da su u napadima baziranim na rešetkama prosti brojevi uravnoteženi, te da su javni i privatni eksponent definirani modulo $\varphi(N)$.

4.3.1. Napad rešetkama

Prikazati ćemo Boneh i Durfeeov napad baziran na rešetkama s malim tajnim eksponentom RSA na proizvoljnom javnom eksponentu u sljedećem napadu. Rezultat se oslanja na nedokazivim pretpostavkama, stoga je naveden samo kao napad.

Napad 4.1. *Za svaki $\epsilon > 0$ postoji n_0 tako da za svaki $n > n_0$ vrijedi: neka je $N = pq$ N-bitni RSA modul s uravnoteženim prostim brojevima, neka je (N, e) valjan javni ključ, a d odgovarajući tajni eksponent definiran modulo $\varphi(N)$. Neka je $e = N^\alpha$ i $d = N^\delta$. Ako tajni eksponent zadovoljava*

$$\delta < \frac{7}{6} - \frac{1}{3}\sqrt{1+6\alpha} - \epsilon,$$

i ako vrijede pretpostavke 2.2 i 2.3, onda se modul N može faktorizirati u u polinomijalnom vremenu $\log(N)$.

Objašnjenje Neka su k i s poznate vrijednosti. Reducirana ključna jednadžba $ed = 1 + k(N-s)$ modulo e

$$1 + kN - ks \equiv 0 \pmod{e}$$

je motivacija da pogledamo male korijene polinoma $f_e(x,y) \in \mathbb{Z}[x,y]$ danog s

$$f_e(x,y) = Nx + xy + 1.$$

Korijen od $f_e(x,y)$ modulo e je $(x_0, y_0) = (k, -s)$.

Neka je $e = N^\alpha$ i $d = N^\delta$. Koristeći nejednakosti (9) i (10) (jer su prosti brojevi uravnoteženi), dobivamo granice

$$|x_0| = k = \frac{ed-1}{\varphi(N)} < \frac{ed}{\frac{1}{2}N} = 2N^{\alpha+\delta-1} \tag{13}$$

$$|y_0| = |-s| = p + q - 1 < 3N^{1/2}. \tag{14}$$

Ideja je da, koristeći polinom $f_e(x, y)$ i granice X i Y, konstruiramo rešetku čiji elementi odgovaraju polinomu s korijenom (x_0, y_0) modulo neka potencija od e. Za neki fiksni cijeli broj $m > 0$, definiramo polinome

$$\begin{aligned} g_{i,k}(x, y) &:= x^i f_e^k(x, y) e^{m-k} \\ h_{j,k}(x, y) &:= y^j f_e^k(x, y) e^{m-k}. \end{aligned}$$

Funkcija $g_{i,k}(x, y)$ je umnožak baznog polinoma $f_e(x, y)$ i neke potencije od x, stoga ju nazivamo x-pomak. Iz istog razloga funkciju $h_{j,k}(x, y)$ nazivamo y-pomak. Iz ove konstrukcije primjećujemo da za svaki $i, j \geq 0$ i $0 \leq k \leq m$, $h_{j,k}(x, y)$ vrijedi

$$f_e(x_0, y_0) \equiv 0 \pmod{e} \longrightarrow \begin{cases} g_{i,k}(x_0, y_0) \equiv 0 \pmod{e^m} \\ h_{j,k}(x_0, y_0) \equiv 0 \pmod{e^m}. \end{cases} \tag{15}$$

Konstruiramo matricu baze $\mathcal{B}$ za rešetku $\mathcal{L}$ koristeći koeficijente polinoma $g_{i,k}(xX, yY)$ i $h_{j,k}(xX, yY)$, gdje su X i Y granice za x_0 i y_0. Svaki vektor u ovoj rešetki je cjelobrojna linearna kombinacija koeficijenata vektora $g_{i,k}(xX, yY)$ i $h_{j,k}(xX, yY)$. Iz (15) slijedi da svaki od tih polinoma $f(x, y)$ zadovoljava

$$f_e(x_0, y_0) \equiv 0 \pmod{e} \longrightarrow f(x_0, y_0) \equiv 0 \pmod{e^m}.$$

Dakle, svaki vektor v iz rešetke $\mathcal{L}$ odgovara polinomu $f(x, y)$ koji ima korijen (x_0, y_0) modulo e^m.

Boneh i Durfee koriste posebnu konstrukciju za matricu baze $\mathcal{B}_{BD}$ koja se sastoji od koeficijenta vektora

$$\begin{aligned} \{g_{i,k}(xX, yY) &\mid 0 \leq k \leq m,\ 0 \leq i \leq m - k\} \\ \{h_{j,k}(xX, yY) &\mid 0 \leq k \leq m,\ 1 \leq j \leq t\}, \end{aligned}$$

za neki cijeli broj $t > 0$. To su bazni vektori za rešetku $\mathcal{L}_{BD}$ kojih ima $\omega = (m + 1)(m + 2)/2 + t(m + 1)$ i svi su u parovima linearno nezavisni, stoga vrijedi $\dim(\mathcal{L}_{BD}) = w$.

Retci i stupci matrice baze su poredani tako da je matrica donje trokutasta. Retke čine redom:

- polinomi x-pomaka $g_{i,k}(xX, yY)$ koji su poredani u rastućim vrijednostima indeksa $l = i + k$ za $l = 0, \ldots, m$. Za svaku vrijednost l, polinomi su nadalje poredani s rastućim vrijednostima od $k = 0, \ldots, l$.
- polinomi y-pomaka $h_{j,k}(xX, yY)$ pojavljuju s rastućim vrijednostima od $j = 1, \ldots, t$. Za svaku vrijednost j polinomi su dalje poredani s rastućim vrijednostima od $k = 0, \ldots, m$.

S ovakvim poretkom baznih vektora, svaki novi bazni vektor ima točno jedan novi monom koji nije bio prisutan u prethodnim baznim vektorima.

Stupci odgovaraju određenim monomima, pri čemu stupac n odgovara novom monomu koji je uveden u retku n. Kako se povećava indeks stupca, povećava se i eksponent u monomu, tako da:

- prvih $w_x = (m+1)(m+2)/2$ stupaca odgovara monomima $x^{i+k}y^k$ koji se pojavljuju u rastućim vrijednostima od $i = 0, \ldots, m$, a za svaku vrijednost od i, pojavljuju se u rastućim vrijednostima od $k = 0, \ldots, m$. Ovi stupci odgovaraju monomima koji se nalaze u polinomu x-pomaka.
- ostalih $\omega_y = t(m+1)$ stupaca odgovara monomima $x^k y^{j+k}$ koji se pojavljuju u rastućim vrijednostima od $j = 1, \ldots, t$, te se svaka vrijednost j pojavljuje se u rastućim vrijednostima od $k = 0, \ldots, m$. Ti stupci odgovaraju monomima koji se pojavljuju u polinomima y-pomaka.

U nastavku je prikazana matrica baze $\mathcal{B}_{BD}$ na prethodno opisan način kada je $m = 2$ i $t = 1$.

$$\mathcal{B}_{BD} = \begin{array}{c} \\ g_{0,0} \\ g_{1,0} \\ g_{0,1} \\ g_{2,0} \\ g_{1,1} \\ g_{0,2} \\ h_{1,0} \\ h_{1,1} \\ h_{1,2} \end{array} \begin{array}{c} \begin{array}{ccccccccccc} 1 & x & xy & x^2 & x^2y & x^2y^2 & y & xy^2 & x^2y^2 \end{array} \\ \left[\begin{array}{ccccccccc} e^2 & & & & & & & & \\ 0 & e^2X & & & & & & & \\ e & eNX & eXY & & & & & & \\ 0 & 0 & 0 & e^2X^2 & & & & & \\ 0 & eX & 0 & eNX^2 & eX^2Y & & & & \\ 1 & 2NX & 2XY & N^2X^2 & 2NX^2Y & X^2Y^2 & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & e^2Y & & \\ 0 & 0 & eNXY & 0 & 0 & 0 & eY & eXY^2 & \\ 0 & 0 & 2NXY & 0 & N^2X^2Y & 2NX^2Y^2 & Y & 2XY^2 & X^2Y^3 \end{array}\right] \end{array}$$

Budući da je bazna matrica donjetrokutasta kvadratna matrica, za računanje volumena rešetke $\mathcal{L}_{BD}$, potrebno je izračunati produkt dijagonalnih elemenata.

$$\begin{aligned} \mathrm{vol}(\mathcal{L}_{BD}) &= \left(\prod_{k=0}^{m}\prod_{i=0}^{m-k} X^{i+k}Y^k e^{m-k}\right)\left(\prod_{k=0}^{m}\prod_{j=1}^{t} X^k Y^{j+k} e^{m-k}\right) \\ &= (eX)^{m(m+1)(m+2)/3+tm(m+1)/2} Y^{m(m+1)(m+2)/6+t(m+1)(m+t+1)/2}, \end{aligned}$$

gdje prvi dvostruki produkt odgovara doprinosu iz polinoma x-pomaka, a drugi dvostruki produkt odgovara doprinosu iz polinoma y-pomaka.

Primjenom teorema 2.6 dobivamo LLL-reduciranu bazu s kratkim vektorima $p_1(x, y)$ i $p_2(x, y)$ koji zadovoljavaju

$$||p_1(xX, yY)|| \leq ||p_2(xX, yY)|| \leq 2^{\omega/4}\mathrm{vol}(\mathcal{L})^{1/(\omega-1)}.$$

Ako veći od tih polinoma zadovoljava Howgrawe-Graham granicu za bivarijantne polinome (teorem 2.10)

$$||p_2(xX, yY)|| < \frac{e^m}{\sqrt{\omega}},$$

onda dobivamo sljedeću nejednakost

$$2^{\omega/4}\mathrm{vol}(\mathcal{L})^{1/(\omega-1)} < \frac{e^m}{\sqrt{\omega}}.$$

Dakle, ukoliko vrijedi

$$\mathrm{vol}(\mathcal{L}) < \gamma e^{m(\omega-1)}, \tag{16}$$

pri čemu je $\gamma = 2^{-\omega(\omega-1)/4}\omega^{-(\omega-1)/2}$ konstanta kada su m i t fiksirani, onda je (x_0, y_0) cjelobrojni korijen od $p_1(x, y)$ i $p_2(x, y)$.

Ovo je nužan uvjet napada. Kada volumen od $\mathcal{L}_{BD}$ zadovoljava nejednakost (16), oba polinoma $p_1(x, y)$ i $p_2(x, y)$ imaju korijen (x_0, y_0) koji je cjelobrojan. Ako su ti polinomi i algebarski nezavisni možemo izračunati rezultantu od $p_1(x, y)$ i $p_2(x, y)$ obzirom na varijablu x za dobivanje univarijantnog polinoma $p_{1,2}(y)$. Cjelobrojni korijen $y_0 = -s$ polinoma $p_{1,2}(y)$ se pronalazi uobičajenim tehnikama za polinome jedne varijable. Kada je s poznat možemo izračunati $\varphi(N) = N - s$ i faktorizirati modul N.

Iako je nužan uvjet savršeno valjan, nemamo uvid o uspjehu napada na temelju ulaznih parametara kao α i δ. Kako bi izvukli više korisnih nužnih uvjeta napraviti ćemo neke pretpostavke i pojednostavljenja.

1. Ignoriramo faktor $\gamma = 2^{-\omega(\omega-1)/4}\omega^{-(\omega-1)/2}$ u (16), jer za odabrane m i t, γ je fiksna i može biti zanemariva ako je N dovoljno velik, te stoga i e dovoljno velik.

2. Uvodimo supstituciju $X = 2N^{\alpha+\delta-1}$, $Y = 3N^{1/2}$ i $e = N^{\alpha}$ u nužan uvjet (5), te neka je $t = \tau m$ za neki realni $\tau > 0$.

Tada nužan uvjet postaje:

$$N^{(\tau^2/4+(\alpha+\delta/2-1/4)\tau+2\alpha/3+\delta/3-1/4)m^3+\mathcal{O}(m^3))} < N^{(\alpha\tau+\alpha/2)m^3+\mathcal{O}(m^3)},$$

gdje smo uključili doprinos iz m^3. Za dovoljno velik m, što možemo zahtjevati obzirom na veliki N tako da je γ još uvijek zanemarivo mala, možemo napraviti doprinose $\mathcal{O}(m^3)$ dovoljno male. Ignoriranjem malih vrijednosti i promatranjem samo koeficijenta uz m^3, uvjeti eksponenata se mogu pojednostavniti

$$\frac{1}{4}\tau^2 + \Big(\frac{\delta}{2} - \frac{1}{4}\Big)\tau + \frac{\alpha}{6} + \frac{\delta}{3} - \frac{1}{4} < 0.$$

Lijeva strana nejednakosti je minimizirana kada je τ dan s $\tau_{min} = \frac{1}{2} - \delta$, koju uvrštavanjem u nejednadžbu dobivamo

$$\delta < \frac{7}{6} - \frac{1}{3}\sqrt{1 + 6\alpha} - \epsilon, \tag{17}$$

gdje $\epsilon > 0$ predstavlja sve ignorirane male vrijednosti koja može biti blizu nule za dovoljno velike N i m.

Dakle, za dovoljno velik N, ako privatni eksponent $d = N^{\delta}$ zadovoljava nužan uvjet (17) mogu se pronaći dva polinoma sa korijenom $(x_0, y_0) = (k, -s)$. Ako su ti polinomi algebarski nezavisni, onda se može izračunati $y_0 = -s$ i iskoristiti za faktoriziranje modula po prethodno opisanom načinu, gdje svi izračuni mogu biti napravljeni u polinomijalnom vremenu od $\log(N)$.

□

U slučaju malih tajnih eksponenata RSA javni eksponent je približno iste veličine kao i modul, te korištenjem približne vrijednosti $\alpha \approx 1$, dovoljan uvjet (17) napada 4.1 postaje

$$\delta < \frac{7}{6} - \frac{1}{3}\sqrt{7} - \epsilon \approx 0.2847 - \epsilon.$$

Dakle, za dovoljno velik modul, RSA sustav se smatra nesigurnim kada tajni eksponent zadovoljava prethodnu granicu.

4.3.2. Napad podrešetkama

Boneh i Durfee su pokazali da granica $\delta < 0.2847$ u napadu 4.1 može biti poboljšana pronalaskom kratkog vektora u posebno odabranoj podrešetki od $\mathcal{L}_{BD}$. Motivacija za poboljšanje je u tome da neki bazni vektori $\mathcal{B}_{BD}$ više doprinose volumenu rešetke od drugih.

Iz nužnog uvjeta dobivenog u prethodnom napadu

$$\mathrm{vol}(\mathcal{L}_{BD}) < \delta e^{m(\omega-1)},$$

zaključujemo da prosječni doprinos svakog baznog vektora na volumen mora biti manji od $e^{m(\omega-1)/\omega} < e^m$. Intuitivno, ako bazne vektore s dijagonalnim elementima većim od e^m uklonimo iz matrice baze, volumen nastale rešetke bi se smanjio, te se granica od δ povećava. Na toj ideji će se temeljiti poboljšanje napada. No, pri uklanjanju baznih vektora iz matrice baze dobivamo rešetku koja nije punog ranga, stoga se volumen računa na netrivijalan način. Boneh i Durfee su prevladali tu prepreku tako da računaju granicu volumena koristeći rezultate posebne klase matrica takozvane **geometrijski progresivne matrice**.

Neka su a i b pozitivni cijeli brojevi, te M $(a+1)b \times (a+1)b$ matrica. Stupci matrice M su indeksirani s (i,j), a retci (k,l). Stupac koji odgovara paru indeksa (i,j) je $(bi+j)$-ti stupac matrice M, a redak koji odgovara paru indeksa (k,l) je $(bk+l)$-ti redak matrice. Stoga ćemo s $M(i,j,k,l)$ označiti elemente matrice koji odgovaraju stupcu s parom indeksa (i,j) i retku s parom indeksa (k,l). Dijagonalni elementi matrice su oblika $M(k,l,k,l)$.

Definicija 4.1. *Neka su $C, D, c_0, c_1, c_2, c_3, c_4, \beta \in \mathbb{R}$, te $\beta \geq 1$. Za matricu M kažemo da je geometrijski progresivna matrica s parametrima $(C, D, c_0, c_1, c_2, c_3, c_4, \beta)$ ako za svaki $i, k = 0, \ldots, a$ i $j, l = 1, \ldots b$ vrijedi sljedeće:*

1. $|M(i,j,k,l)| \leq CD^{c_0+c_1 i+c_2 j+c_3 k+c_4 l}$
2. $M(i,j,k,l) = D^{c_0+c_1 k+c_2 l+c_3 k+c_4 l}$
3. $M(i,j,k,l) = 0$ *za* $i > k$ *ili* $j > l$
4. $\beta c_1 + c_3 \geq 0$ *i* $\beta c_2 + c_4 \geq 0$.

Boneh i Durfee su dali rezultat za granicu volumena podrešetke koja pripada rešetki čija je matrica baze geometrijski progresivna matrica. Geometrijska interpretacija determinante matrice je volumen paralelepipeda koji je razapet retcima matrice. Stoga je njihov rezultat prilagođen na granicu determinante podmatrice umjesto na granicu volumena podrešetke.

Sljedećim teoremom dana je granica determinante matrice koja nije kvadratna.

Teorem 4.3. *Neka je M $(a+1)b \times (a+1)b$ geometrijska progresivna matrica s parametrima $(C, D, c_0, c_1, c_2, c_3, c_4, \beta)$ i neka je B realan broj. Definiramo skup redaka od M čiji dijagonalni elementi nisu veći od B*

$$S_B = \{(k,l) \in \{0,\dots,a\} \times \{1,\dots,b\} : M(k,l,k,l) \leq B\},$$

te neka je $s = |S_B|$. Ako se matrica M_B sastoji od redaka $(k,l) \in S_B$ iz M, onda vrijedi

$$\det(M_S) \leq ((a+1)b)^{s/2}(1+C)^{s^2} \prod_{(k,l)\in S_b} M(k,l,k,l).$$

U sljedećem napadu se nalazi njihov rezultat generaliziran za proizvoljne javne eksponente.

Napad 4.2. *Za svaki $\epsilon > 0$ postoji N_0 takav da za svaki $N > N_0$ vrijedi: Neka je $N = pq$ RSA modul s uravnoteženim prostim brojevima, neka je $e = N^\alpha$ valjan javni eksponent, $d = N^\delta$ odgovarajući privatni eksponent modulo $\varphi(N)$. Ako tajni eksponent zadovoljava*

$$\delta < \frac{2 - \sqrt{2\alpha}}{2} - \epsilon,$$

i ako pretpostavke 2.2 i 2.3 vrijede, onda se modul N može faktorizirati u polinomijalnom vremenu $\log(N)$.

Objašnjenje. Analogno prethodnom napadu, koristeći granice (13) i (14) za x_0 i y_0, konstruiramo matricu baze $\mathcal{B}_{BD}$ za neki fiksni $m > 0$ i $t = (1-2\delta)m$.

Potom konstruiramo novu rešetku $\mathcal{L}'$ s matricom baze $\mathcal{B}'$ tako da iz $\mathcal{B}_{BD}$ uklonimo bazne vektore koji odgovaraju polinomu y-pomaka čiji je dijagonalni element veći od e^m. Budući da je $\mathcal{L}'$ podrešetka od $\mathcal{L}$, svaki vektor iz $\mathcal{L}'$ je linearna kombinacija polinoma s korijenom $(x_0, y_0) = (k, -s)$ modulo e^m.

Prema teoremu 2.6 za izračunatu LLL-reduciranu bazu za $\mathcal{L}'$ možemo naći dva linearno nezavisna vektora koji zadovoljavaju

$$||f_1'(xX, yY)|| \leq ||f_2'(xX, yY)|| \leq 2^{\omega'/4}\mathrm{vol}(\mathcal{L}')^{1/(\omega'-1)},$$

gdje je $\dim(\mathcal{L}') = \omega'$. Ako je zadovoljena Howgrave-Graham granica iz teorema 2.10

$$||f_2'(xX, yY)|| < \frac{e^m}{\sqrt{\omega}},$$

onda je (x_0, y_0) cjelobrojni korijen od oba polinoma $f_1'(x, y)$ i $f_2'(x, y)$.
Iz prethodne dvije nejednakosti dobivamo nužan uvjet napada

$$\text{vol}(\mathcal{L}') < \gamma' e^{m(\omega'-1)}, \tag{18}$$

gdje je $\gamma' = 2^{-\omega'(\omega'-1)/4}\omega^{-(\omega'-1)/2}$ konstanta za fiksni m. Dakle, ukoliko je ova granica zadovoljena, te ako su dva polinoma algebarski nezavisna, možemo efikasno naći y_0 i faktorizirati modul.

□

Pogledajmo kako izgleda rešetka bazne matrice $\mathcal{B}_{BD}$ s dimenzijom $\omega = \omega_x + \omega_y$, gdje su ω_x i ω_y, broj polinoma x-pomaka, odnosno y-pomaka.

$$\mathcal{B}_{BD} = \begin{bmatrix} \mathcal{M}_x & 0 \\ \mathcal{M}_{yx} & \mathcal{M}_y \end{bmatrix},$$

gdje je $\mathcal{M}_x$ donjetrokutasta $\omega_x \times \omega_x$ matrica koja odgovara polinomima x-pomaka, $\mathcal{M}_{yx}$ $\omega_y \times \omega_x$ matrica čijih ω_x stupaca odgovara polinomima y-pomaka, $\mathcal{M}_y$ donjetrokutasta $\omega_y \times \omega_y$ matrica koja odgovara ostatku polinoma y-pomaka. Uklanjanjem polinoma y-pomaka iz odgovarajućih redaka $\mathcal{B}_{BD}$ dobivamo baznu matricu

$$\mathcal{B}' = \begin{bmatrix} \mathcal{M}_x & 0 \\ \mathcal{M}'_{yx} & \mathcal{M}'_y \end{bmatrix}.$$

Boneh i Durfee su pokazali da je matrica $\mathcal{M}_y$ geometrijski progresivna matrica s parametrima

$$(m^{2m}, e, m, \frac{1}{2} + \delta, -\frac{1}{2}, -1, 1, 2),$$

pri čemu su ignorirane konstante X, Y, te $e \approx N$. Korištenjem teorema 4.3 su izračunali determinantu matrice $\mathcal{M}'_y$ koja nije kvadratna i tako došli do najjačeg poznatog napada na RSA s malim tajnim eksponentom.

U RSA kriptosustavu s malim tajnim ekponentom, javnim ekspoenent približno jednake veličine kao i modul, te korištenjem aproksimacije $\alpha \approx 1$, dovoljan uvijet napada 4.2 postaje

$$\delta < \frac{2 - \sqrt{2}}{2} - \epsilon \approx 0.2929 - \epsilon.$$

Ova granica za δ je originalan Boneh i Durfeeov rezultat. Za dovoljno velik modul, RSA se smatra nesigurnim ukoliko tajni eksponent zadovoljava tu granicu.

5. Zaključak

U radu su predstavljeni neki napadi na RSA kriptosustav koji ima mali tajni eksponent. Napade možemo podijeliti na napade koji u izvođenju koriste verižne razlomke, za koje postoje dokazi, te napade koji u izvođenju koriste rešetke, koji su u pravilu heuristički.

Najpoznatiji napad na RSA kriptosustav je Wienerov napad koji koristi verižne razlomke. Wienerova granica tajnog eksponenta je $N^{0.25}$. Proširenjem Wienerovog napada Verheul i van Tilborg otkrivaju tajni eksponent koji je nekoliko bitova veći od Wienerove granice. Svoj doprinos u napadima koji koriste verižne razlomke je dao i Dujella. Metodom meet-in-the-middle otkriva tajni eksponent koji je manji do $2^{30}N^{0.25}$

Wienerov napad koji se može prikazati kao heuristička metoda bazirana na rešetkama otkriva tajni eksponent do granice $N^{0.2489}$.

Razmatrani napadi bazirani na rešetkama su heuristički, osim Mayovog napada koji je za svoj napad baziran na rešetkama dokazao koristeći ograničenja na proste brojeve [9].

Unatoč tome što su napadi bazirani na rešetkama u pravilu heuristički, u praksi se smatraju dosta uspješnim. Tako je najjači napad na RSA kriptosustav Boneh i Durfeeov napad. Heurističkom napadom baziranom na rešetkama uspjeli su razbiti RSA kriptosustav kada je veličina tajnog eksponenta manja od $N^{0.2847}$. Granicu su uspjeli povećati na $N^{0.2929}$ napadom baziranim na podrešetkama. Pretpostavke napada su da je N proizvoljno velik i da je snaga računala beskonačno velika, što zapravo nije moguće u stvarnosti. Za velike vrijednosti fiksiranih parametara m i t, rešetke postaju veće, pa je dominantna operacija u napadima računanje LLL-reducirane baze što dovodi do opterećenja računala.

U praksi se može dogoditi da navedeni napadi mogu razbiti RSA kriptosustav, odnosno otkriti tajni eksponent d koji je za nekoliko bitova veći od teorijski izračunatih granica.

Većina napada na RSA s malim tajnim eksponentom se bazira na uravnoteženim prostim brojevima, odnosno prostim brojevima koji su približno jednakih veličina. Postoje dokazi koji govore da ukoliko je razlika između prostih brojeva mala, onda su napadi jači. Stoga pri odabiru prostih brojeva RSA kriptosustava treba izbjegavati uravnotežene brojeve. Dakle, pri odabiru prostih brojeva treba biti oprezan i birati ih na način da modul N bude otporan na metode faktorizacije.

Što je veći modul, veća je sigurnost RSA kriptosustava. No, u tom slučaju se šifriranje i dešifriranje u RSA kriptosustavu odvija sporije. Stoga pri odabiru veličine modula se moramo pitati koliko zaštićeni podaci moraju biti i koliko jake su moguće prijetnje. Najveći RSA modul je faktoriziran 2009. godine. Dugačak je 768 bita (232-znamenkst broj), a za njegovo razbijanje je korišteno stotinjak računala i dvije godine. Stoga je veličina RSA modula koja se trenutno smatra sigurnom 1024-bita.

Napadi na RSA kritposustav tema su proučavanja mnogih matematičara koji su predstavili svoju verziju napada. No unatoč brojnim istraživanjima, nije pronađena metoda koja u potpunosti razbija RSA kriptosustav, nego se svi napadi temelje na tome da se pronađe neka njegova slabost. Stoga su nam svi poznati napadi na RSA pokazatelji na što treba paziti i

što treba izbjegavati pri izboru parametara i implementacije RSA. Za sada RSA kriptosustav možemo smatrati sigurnim kriptosustavom.

Literatura

[1] W. Diffie, M. E. Hellman, *New Directions in Cryptography*, IEEE Transactions on Information Theory archive Volume 22, Issue 6 (1976), 644-654

[2] A. Dujella, *Continued fractions and RSA with small secret exponent*, Tatra Mt. Math. Publ. 29 (2004), 101-112

[3] A. Dujella, A variant of Wiener's attack on RSA Computing, Computing 85 (2009), 77-83

[4] A. Dujella, *Diofantske aproksimacije i primjene*, Skripta, PMF-Matematički odjel, Sveučilište u Zagrebu, poslijediplomski kolegij (2011/2012)

[5] M.Gardner, *Mathematical Games: A New Kind of Cipher that Would Take Millions of Years to Break*, Scientific American. 237(1977), 120-124

[6] B. Ibrahimpašić, *RSA kriptosustav*, Osječki matematički list, 5(2005), 101-112

[7] M.Hinek, *Cryptanalysis of RSA and Its Variants, Cryptography and Network Security Series*, Chapman & Hall/CRC, Boca Raton, 2009.

[8] I. Matić, *Uvod u teoriju brojeva*, Odjel za matematiku Sveučilišta J. J. Strossmayera u Osijeku, 2015.

[9] A. May, *New RSA Vulnerabilities Using Lattice Reduction Methods*, PhD thesis, University of Paderborn, 2003.

[10] A. Nitaj, *Diophantine and Lattice Cryptanalysis of the RSA Cryptosystem*, Artificial Intelligence, Evolutionary Computing and Metaheuristics Volume 427 of the series Studies in Computational Intelligence, 2013, 139-168

Sažetak

U ovom radu predstavljeni su rizici RSA kriptosustava sa malim tajnim eksponentom. Rad se sastoji od tri cjeline. Prvi dio sadrži matematičku osnovu potrebnu za razumijevanje teme diplomskog rada. U sklopu drugog dijela dana je teorijska osnova za shvaćanje RSA kriptosustava, prvog kriptosustava s javnim ključem, kao i sama implementacija RSA kriptosustava. Također, u ovom dijelu opisana je sigurnost RSA kriptosustava te potencijalni rizici. Posljedni dio sadrži neke od napada na RSA kriptosustav s malim tajnim eksponentom. Primjenom grube sile moguće je probiti svaki kriptosustav, no taj posao nije efikasan. Javljaju se algoritmi koji koriste neke slabosti kriptosustava, te u razumnom vremenu razbijaju kriptosustav. Obrađeni napadi probijaju RSA kriptosustav s malim tajnim eksponentom do N^{δ} za $\delta \geq 1/4$. Najpoznatiji od njih je Wienerov napad koji koristi verižne razlomke, a najjači je Boneh i Durfee napad koji koristi rešetke.

Ključne riječi: kriptologija, RSA kriptosustav, tajni eksponent, Wienerov napad, verižni razlomak, rešetke, Boneh i Durfee napad

Attacks on RSA cryptosystem with small private exponent

Summary

This paper presents the risks of RSA with small private exponent. The paper consists of three parts. The first part provides the mathematical basis needed for the understanding of the thesis. Within the second part, the theoretical basis for the understanding of RSA, the first public key cryptosystem, is given, as well as its implementation. Further more, this part explains the security of RSA and the potential risks. The last part contains some of the attacks on the RSA cryptosystem with small private exponent. Every cryptosystem can be broken using brute-force search attack. However, this is not practical. Advanced algorithms use some of the weaknesses of cryptosystem, and are able to break cryptosystem in a reasonable time. All of the attacks can efficiently break RSA with private exponents up to N^{δ} for some $\delta \geq 1/4$. The best known of these algorithms is the Wiener's attack that uses a continued fractions, but the strongest is Boneh and Durfee's attack that uses the lattice.

Key words: cryptology, RSA cryptosystem, private exponent, Wiener's attack, continued fraction, lattice, Boneh and Durfee's attack

Životopis

Rođena sam 28. prosinca 1990. godine u Slavonskom Brodu. U Županji sam završila Osnovnu školu Ivana Kozarca i opću gimnaziju u Gimnaziji Županja. 2009. godine sam obrazovanje nastavila na Sveučilišnom preddiplomskom studiju matematike na Odjelu za matematiku u Osijeku. Titulu sveučilišne prvostupnice matematike stekla sam 2012. godine, a potom upisala Sveučilišni diplomski studij matematike, smjer Financijska matematika i statistika.