Teorija brojeva i kriptografija

Andrej Dujella

PMF-Matematički odsjek, Sveučilište u Zagrebu
HAZU, Razred za matematičke, fizičke i kemijske znanosti
e-mail: duje@math.hr
URL: http://web.math.pmf.unizg.hr/~duje/

Teorija brojeva

Teorija brojeva je grana matematike koja se ponajprije bavi proučavanjem svojstava cijelih brojeva.

Ima vrlo dugu i bogatu povijest (Euklid, Euler, Gauss).

Dugo je smatrana "najčišćom" granom matematike, u smislu da je bila najdalja od bilo kakvih konkretnih primjena.

Danas je teorija brojeva jedna od najvažnijih grana matematike za primjene u kriptografiji i sigurnoj razmjeni informacija (od 1975. godine nadalje).

Navedimo neke teme i primjere problema iz teorije brojeva.

Djeljivost:

- Je li broj 123456789 djeljiv s 9?
- Naći broj koji pri dijeljenju sa 7 daje ostatak 2, pri dijeljenju s 11 daje ostatak 1, a pri dijeljenju s 13 daje ostatak 9.
- Naći ostatak pri dijeljenju broja 2^{100} sa 101.

Prosti brojevi i faktorizacija:

- Prirodan broj $p > 1$ je prost ako je djeljiv samo s 1 i sa samim sobom: $2, 3, 5, 7, 11, 13, 17, 19, \dots$.
- Koliko ima prostih brojeva?
- Je li broj 91 prost?
- Je li broj $2^{31} - 1$ prost?
- Rastaviti na proste faktore broj 1001.
- Rastaviti na proste faktore broj $2^{32} + 1$.
- Može li se svaki paran broj veći od 2 prikazati kao zbroj dva prosta broja?

($4 = 2+2$, $6 = 3+3$, $8 = 3+5$, $10 = 3+7$, $12 = 5+7$)

Najveći zajednički djelitelj:
- Odrediti $\text{nzd}(901, 1001)$ (bez faktorizacije).
- Naći cijele brojeve x i y takve da je $901x - 1001y = \text{nzd}(901, 1001)$.

Euklidov algoritam:

$$\begin{aligned} 1001 &= 901 \cdot 1 + 100 \\ 901 &= 100 \cdot 9 + 1 \end{aligned}$$

Dakle, $\text{nzd}(901, 1001) = 1$. Nadalje,

$$\begin{aligned} 1 &= 901 - 100 \cdot 9 = 901 - (1001 - 901 \cdot 1) \cdot 9 \\ &= 901 \cdot 10 - 1001 \cdot 9. \end{aligned}$$

Kineski teorem o ostatcima:

- Riješiti sustav kongruencija:

$x \equiv 91 \pmod{901}, \quad x \equiv 101 \pmod{1001}.$

Rješenje je oblika

$$x \equiv 1001x_1 + 901x_2 \pmod{901 \cdot 1001},$$

gdje je

$$1001x_1 \equiv 91 \pmod{901}$$

$$901x_2 \equiv 101 \pmod{1001}.$$

Iz Euklidovog algoritma znamo da je

$$901 \cdot 10 - 1001 \cdot 9 = 1,$$

pa je

$$x_1 \equiv -9 \cdot 91 \equiv 82 \pmod{901},$$

$$x_2 \equiv 10 \cdot 101 \equiv 9 \pmod{1001},$$

te je rješenje $x \equiv 90191 \pmod{901901}$.

Diofantske jednadžbe:

$$3x + 5y = 28 \quad \text{(linearna)}$$

$$x^2 + y^2 = z^2 \quad \text{(Pitagorina)}$$

$$x^2 - 2y^2 = 1 \quad \text{(Pellova)}$$

$$y^2 = x^3 + 17 \quad \text{(Mordellova)}$$

Nedavno je dokazano da ne postoji Diofantova petorka (ali ima beskonačno mnogo takvih šestorki u racionalnim brojevima).
Ako je $\{1, 3, 8, d\}$ Diofantova četvorka, onda je $d = 120$.

$$1 \cdot d + 1 = x^2, \quad 3 \cdot d + 1 = y^2, \quad 8 \cdot d + 1 = z^2$$

$(xyz)^2 = (d + 1)(3d + 1)(8d + 1)$ – eliptička krivulja

Diofantske aproksimacije:

Nejednadžba $\left|\sqrt{2}-\frac{a}{b}\right| < \frac{1}{2b^2}$ ima beskonačno mnogo rješenja: $\frac{a}{b} = 1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \frac{41}{29}, \frac{99}{70}, \ldots$, a nejednadžba $\left|\sqrt{2}-\frac{a}{b}\right| < \frac{1}{4b^2}$ niti jedno.

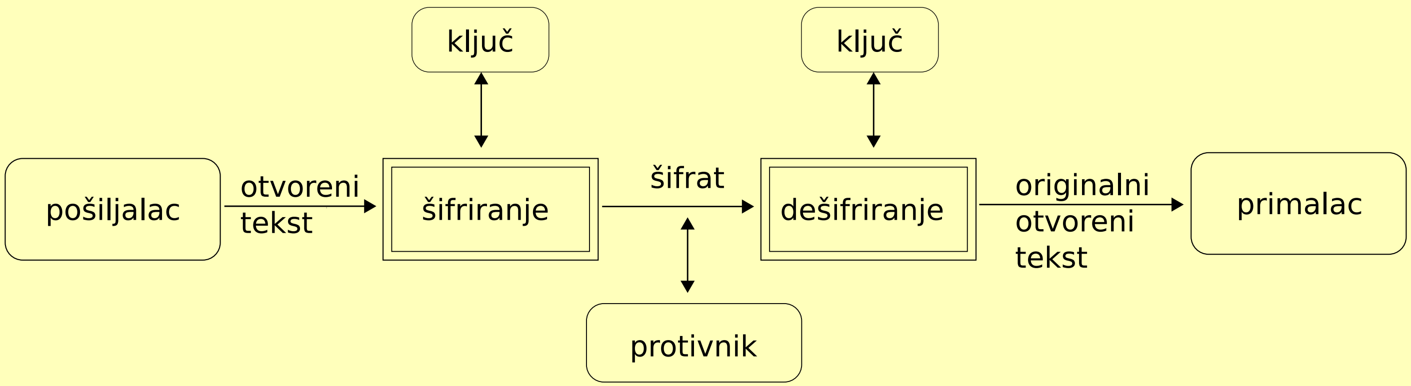
Uočimo da brojevi a, b iz prve nejednakosti zadovoljavaju Pellovu jednadžbu $a^2 - 2b^2 = \pm 1$.

Verižni razlomak: $\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \ddots}}$

$\sqrt{2} \approx 1.4142$,
$1 + 1/(2 + 1/2) = 7/5 = 1.4$,
$1 + 1/(2 + 1/(2 + 1/2)) = 17/12 \approx 1.4167$.

Kriptografija

Šifriranje ili kriptografija (tajnopis) je znanstvena disciplina koja se bavi proučavanjem metoda za slanje poruka u takvom obliku da ih samo onaj kome su namijenjene može pročitati.

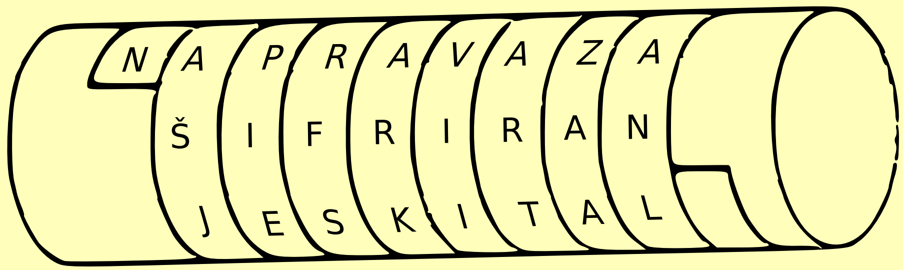


Glavne metode klasične kriptografije:

- transpozicija (premještanje) TAJNA $\mapsto$ JANAT
- supstitucija (zamjena) TAJNA $\mapsto$ UBKOB

Transpozicijske šifre

Skital (Sparta, 5. st. pr. Kr.)



Stupčana transpozicija

Poruka se piše po redcima, a čita po stupcima, ali s promijenjenim poretkom stupaca

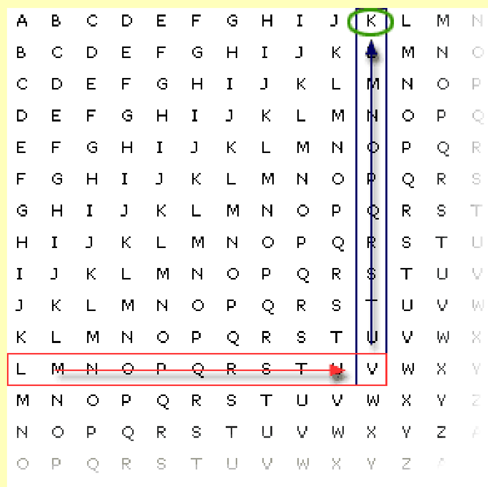
6	1	3	7	5	2	4
S	T	U	P	Č	A	N
A	T	R	A	N	S	P
O	Z	I	C	I	J	A

TTZASJURINPAČNISAOPAC

Supstitucijske šifre

Cezarova šifra (1. st. pr. Kr.)
- svako slovo se pomakne za k mjesta u alfabetu,
- Cezar je koristio šifru s $k = 3$

Vigenèreova šifra (16. st. – 19. st.)
- ključna riječ $(k_1, k_2, \ldots, k_m)$,
- slova se pomiču redom za $k_1, k_2, \ldots, k_m, k_1, k_2, \ldots$ mjesta



ENIGMA (1920. – 2. svjetski rat)

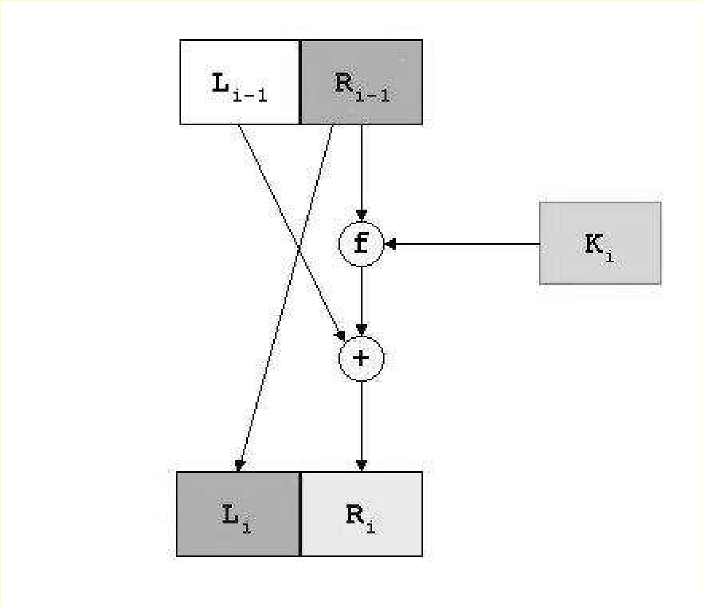
- najčuvenija naprava za šifriranje
- Kriptoanaliza: Marian Rejewski i Alan Turing

Hrvatska:

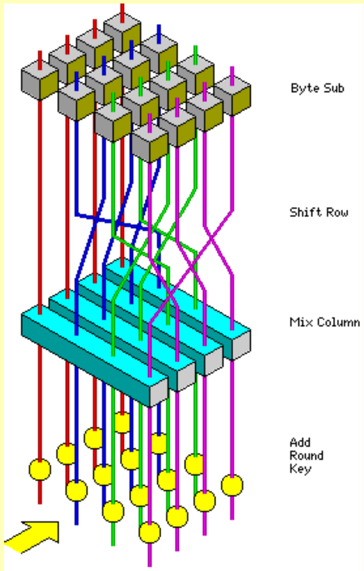
- Ivan Krstitelj Prus: Cryptographia nova seu Ars cryptographica noviter inventa (Nova kriptografija ili nedavno izmišljena kriptografska vještina, 1732.)

- Jules Verne: Mathias Sandorf - roman u kojem se opisuju urote u Austro-Ugarskoj, koje uključuju slanje šifriranih poruka; važan dio radnje se odvija u Hrvatskoj, posebno u Dubrovniku te pazinskom Kaštelu.

DES – Data Encryption Standard (1976. – 1998.)
- kombinira se supstitucija i transpozicija,
- ključna riječ ima 56 bitova (binarnih znamenaka $0, 1$),
- 16 rundi šifriranja



AES – Advanced Encryption Standard (2000. –)
- koristi operacije u polju $\mathbb{F}_{2^8}$,
- polje ima $2^8 = 256$ elemenata
- elementi polja su polinomi stupnja ≤ 7 s koeficijentima iz polja $\mathbb{F}_2 = \{0, 1\}$,
- operacije su zbrajanje polinoma u $\mathbb{F}_2[x]$ ($1+1=0$) i množenje polinoma modulo fiksni polinom osmog stupnja: $x^8 + x^4 + x^3 + x + 1$



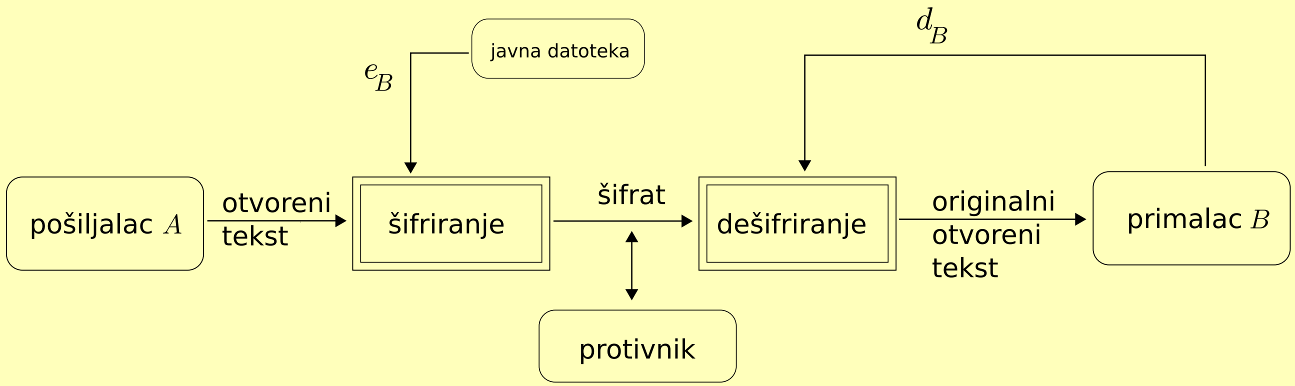
Kriptosustavi s javnim ključem

Sigurnost svih do sada navedenih kriptosustava leži u tajnosti ključa.

Problem: Kako sigurno razmijeniti ključ?

Ideja: javni ključ e_K za šifriranje, tajni (osobni) ključ d_K za dešifriranje.

Ovdje e_K mora biti tzv. jednosmjerna funkcija, tj. nju se računa lako, a njezin inverz jako teško.



Kriptosustavi s javnim ključem su puno sporiji od suvremenih simetričnih kriptosustava (npr. AES-a). Zato se u praksi ne koriste za šifriranje poruka, već za:

- razmjenu ključeva,
- digitalni potpis: $z = d_A(e_B(x))$, $e_A(z) = e_B(x)$.

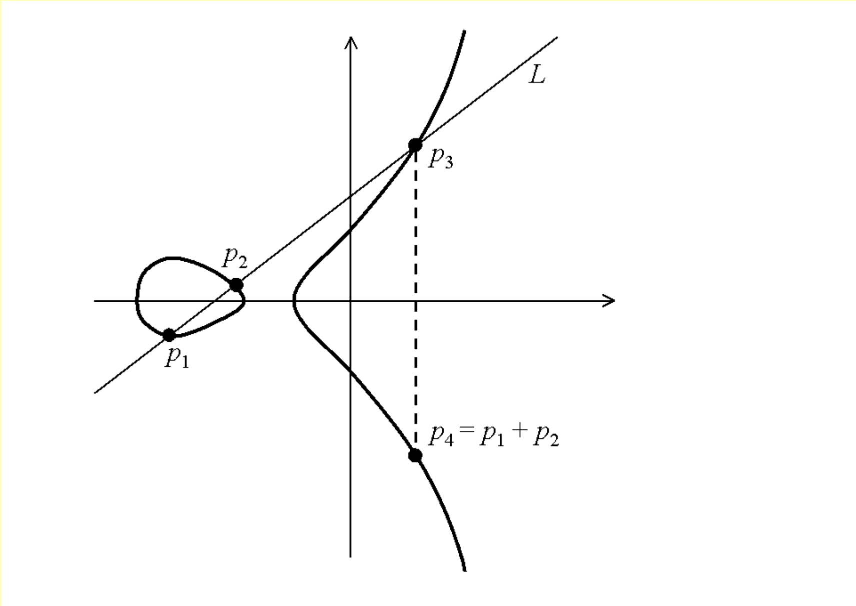
Osnova za kriptosustave s javnim ključem su “teški” matematički problemi:

- faktorizacija velikih složenih brojeva
- problem diskretnog logaritma (DLP)

$$a^x \equiv b \pmod{p}$$

- eliptički diskretni logaritam (ECDPL)

Eliptička krivulja: $y^2 = x^3 + ax + b$



ECDLP: $xP = \underbrace{P + \cdots + P}_{x \text{ pribrojnika}} = Q$ (nad $\mathbb{F}_p$ ili $\mathbb{F}_{2^k}$)

ECDLP je teži od DLP $\Rightarrow$ ista sigurnost uz kraći ključ (1024 $\longleftrightarrow$ 160)

Diffie–Hellmanov protokol za razmjenu ključeva

G je konačna ciklička grupa s generatorom g, tj.
$G = \{g, g^2, \ldots, g^{|G|}\}$

Ana i Branko žele se dogovoriti o jednom tajnom elementu grupe G, preko nesigurnog komunikacijskog kanala kojeg prisluškuje Eva.

Primjer: Grupa $\mathbb{F}_{11}^* = \{1, 2, \ldots, 10\}$ (operacija je množenje modulo 11) je ciklička grupa s generatorom 2.

x	1	2	3	4	5	6	7	8	9	10
2^x	2	4	8	5	10	9	7	3	6	1

Primjer: Grupa točaka na eliptičkoj krivulji

$$E : \quad y^2 = x^3 + x + 3$$

nad poljem $\mathbb{F}_7 = \{0, 1, 2, 3, 4, 5, 6\}$ je ciklička grupa s generatorom $P = (4, 1)$.

x	1	2	3	4	5	6
xP	$(4,1)$	$(6,6)$	$(5,0)$	$(6,1)$	$(4,6)$	$\mathcal{O}$

$$\begin{array}{llcll} & G,\ g & & G,\ g & \\ \text{Ana} & a & \longrightarrow & g^a & \text{Branko} \\ & g^b & \longleftarrow & b & \end{array}$$

Eva: G, g, g^a, g^b

Ana: $(g^b)^a = g^{ab}$ $\searrow$

razmijenili su ključ

Branko: $(g^a)^b = g^{ab}$ $\nearrow$

Eva g^a, g^b ? g^{ab}

Da bi protokol funkcionirao, grupa G treba biti takva da je u njoj potenciranje lako, a logaritmiranje teško.

Primjeri takvih grupa jesu multiplikativna grupa konačnog polja $\mathbb{F}_q^*$ i grupa $E(\mathbb{F}_q)$ točaka na eliptičkoj krivulji nad konačnim poljem (q mora imati barem 300 znamenaka kod $\mathbb{F}_q^*$, odnosno barem 50 znamenaka kod $E(\mathbb{F}_q)$).

RSA kriptosustav

(Rivest, Shamir, Adleman (1977))

- izaberemo tajno dva velika prosta broja p i q,
- izračunamo $n = p \cdot q$ i $\varphi(n) = (p-1)(q-1) = n+1-p-q$ (Eulerova funkcija),
- izaberemo e tako da je $e < \varphi(n)$ i $\mathrm{nzd}(e, \varphi(n)) = 1$,
- izračunamo tajno d takav da je $d \cdot e \equiv 1 \pmod{\varphi(n)}$ (linearna diofantska jednadžba $d \cdot e - t \cdot \varphi(n) = 1$, rješava se proširenim Euklidovim algoritmom.

(n, e) – javni ključ

(p, q, d) – tajni (osobni) ključ

šifriranje: $e_K(x) = x^e \bmod n$

dešifriranje: $d_K(y) = y^d \bmod n$

Provjera:
$d_K(e_K(x)) \equiv d_K(x^d) \equiv x^{de} \equiv x^{t\varphi(n)+1} \equiv (x^{\varphi(n)})^t \cdot x \equiv x \bmod n$ (Eulerov teorem)

- sigurnost leži u teškoći faktorizacije velikih brojeva: onaj tko zna ili može otkriti faktore p i q javno poznatog broja n, taj može izračunati $\varphi(n) = (p-1)(q-1)$, te saznati tajni eksponent d rješavajući linearnu diofantsku jednažbu $d \cdot e - t \cdot \varphi(n) = 1$.

- efikasnost: modularno potenciranje može se izvesti vrlo efikasno metodom uzastopnog kvadriranja (još se naziva i metoda “kvadriraj i množi” ili “binarne ljestve”) koja koristi binarni zapis broja n.

Recimo da želimo izračunati x^{13}. Binarni zapis od 13 je $(1\,1\,0\,1)_2$. Sada x^{13} možemo izračunati kao

$$x^{13} = x \cdot (x^2)^2 \cdot ((x^2)^2)^2.$$

Mogli bismo reći da smo binarni zapis čitali s desna na lijevo. Ako isti zapis pročitamo s lijeva na desno, onda imamo

$$x^{13} = x \cdot ((x \cdot x^2)^2)^2.$$

- Teško je faktorizirati veliki prirodan broj n.

- Možda i nije; npr. $n = 10^{200} = 2^{200} \cdot 5^{200}$,
$n = 9999 \cdots 9919 = x^2 - 9^2 = (x-9)(x+9)$

- Teško je faktorizirati n koji je produkt dva velika pažljivo odabrana prosta broja p i q (s barem stotinjak znamenaka)

- Kako naći (tajno) veliki prosti broj?

Prostih brojeva ima "puno", pa možemo krenuti od slučajno izabranog prirodnog broja zadane veličine te tražiti prvi veći prosti broj. Kako za dani veliki prirodni broj efikasno odrediti je li prost ili složen?
Čini se ("školskim" načinom - dijeleći redom s 2, 3, ...) da je to podjednako teško kao faktorizirati veliki prirodni broj slične veličine.

- Testiranje prostosti – može se puno brže nego "školski". Postoje polinomijalni ("efikasni") algoritmi koji ne koriste definiciju prostih brojeva, već neka njihova svojstva koja su jednostavna za provjeru.
Mali Fermatov teorem: $a^{p-1} \equiv 1 \pmod{p}$,
$x^2 \equiv 1 \pmod{p} \;\Rightarrow\; x \equiv \pm 1 \pmod{p}$.
(Miller-Rabinov vjerojatnosni test; Agrawal-Kayal-Saxena deterministički test; Lucas-Lehmerov test za Mersenne-ove brojeve)

Najveći danas poznati prosti broj je Mersenneov broj $2^{82589933} - 1$ koji ima $24\,862\,048$ znamenaka (pronađen 2018. godine).

- Faktorizacija: ne može puno brže nego "školski" (po onome što je danas poznato). Najbolji poznati algoritmi su subeksponencijalni. Osnovna ideja je izračunati nzd(n, y) za prikladno odabrani y (tako da rezultat bude $\neq 1, n$). (Metode kvadratnog sita, sita polja brojeva i eliptičkih krivulja.)

Broj RSA-768 sa 768 bitova (232 decimalnih znamenaka) je faktoriziran 2009. godine:

$$
\begin{aligned}
&1230186684530117755130494958384962720772853569595334792197\\
&3224521517264005072636575187452021997864693899564749427740\\
&6384592519255732630345373154826850791702612214291346167042\\
&9214311602221240479274737794080665351419597459856902143413\\
=\ &33478071698956898786044169848212690817704794983713768568912\\
&431388982883793878002287614711652531743087737814467999489\\
\times\ &36746043666799590428244633799627952632279158164343087642676\\
&032283815739666511279233373417143396810270092798736308917.
\end{aligned}
$$

Napadi na RSA koji ne koriste faktorizaciju modula n

1) mali javni eksponent e

Pretpostavimo da imamo tri korisnika s različitim vrijednostima javnog modula n_1, n_2, n_3, te da svi oni koriste isti javni eksponent $e = 3$. Nadalje, pretpostavimo da im netko želi poslati identičnu poruku m.

Eva sazna šifrate c_1, c_2, c_3 i riješi sustav kongruencija (koristeći Kineski teorem o ostatcima)

$$x \equiv c_1 \pmod{n_1}, \quad x \equiv c_2 \pmod{n_2}, \quad x \equiv c_3 \pmod{n_3}.$$

Na taj način, dobije broj x sa svojstvom da je $x = m^3$, pa Eva može izračunati originalnu poruku m računajući treći korijen iz x.

Za izbjegavanje ovog i sličnih napada, preporuča se korištenje javnog eksponenta $e = 65537 = 2^{16} + 1$.

2) relativno mali tajni eksponent d

Iz jednakosti $d \cdot e - t \cdot \varphi(n) = 1$, vidimo da je t/d jako dobra aproksimacija broja $e/\varphi(n)$. Broj

$$\varphi(n) = (p-1)(q-1) = n - p - q + 1$$

nije poznat napadaču, jer ne zna p i q. No, $\varphi(n) \approx n$ može poslužiti kao zadovoljavajuća aproksimacija. Tako dobivamo da je broj t/d, koji sadrži tajni podatak d, dobra aproksimacija broja e/n koji je u potpunosti javan.

Ako d nije velik ($d < \sqrt[4]{n}$), onda t/d mora biti neka konvergenta verižnog razlomka od e/n. Napad se može proširiti i na nešto veće vrijednosti od d primjenom preciznijih rezultata i naprednijih algoritama iz diofantskih aproksimacija.

Za razliku od korisnika koji šifrira otvoreni tekst, korisnik koji dešifrira poruku koja mu je namijenjena poznaje proste faktore p i q modula n. Može najprije "parcijalno" dešifrirati šifrat modulo p i modulo q tako da izračuna $x_p = y^d \bmod p$, $x_q = y^d \bmod q$, pa kombinirati ova dva rezultata pomoću Kineskog teorema o ostatcima da dobije otvoreni tekst modulo n.

Postupak dešifriranja se može ubrzati ako uočimo da se umjesto d može koristi eksponente d_p i d_q definirane s

$$d_p = d \bmod (p-1), \quad d_q = d \bmod (q-1),$$

jer iz Malog Fermatovog teorema slijedi da je $x_p = y^{d_p} \bmod p$, $x_q = y^{d_q} \bmod q$.

Iako eksponent za dešifriranje d ne bi smio biti manji od $n^{0.25}$, čini se da nema prepreka da eksponenti d_p i d_q budu znatno manji, a da to ne naruši sigurnost kriptosustava.