Teorija brojeva i kriptografija

Andrej Dujella

PMF-Matematički odsjek, Sveučilište u Zagrebu
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. 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:

- Prirodan broj $p > 1$ je prost ako je djeljiv samo s 1 i sa samim sobom: $2, 3, 5, 7, 11, 13, 17, 19, \ldots$.
- Koliko ima prostih brojeva?
- Je li broj $2^{31} - 1$ prost?
- Je li broj $2^{32} + 1$ prost?
- 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,

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

Diofantske jednadžbe:

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

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

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

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

$$x^4 - 12x^3y + 20x^2y^2 + 12xy^3 + y^4 = 1 \quad \text{(Thueova)}$$

Ne postoji Diofantova šestorka.
Diofantovih petorki ima najviše konačno mnogo.
Ako je $\{1, 3, 8, d\}$ Diofantova četvorka, onda je $d = 120$.

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

$$(xyz)^2 = (d+1)(3d+1)(8d+1) - \text{eliptička krivulja}$$

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.

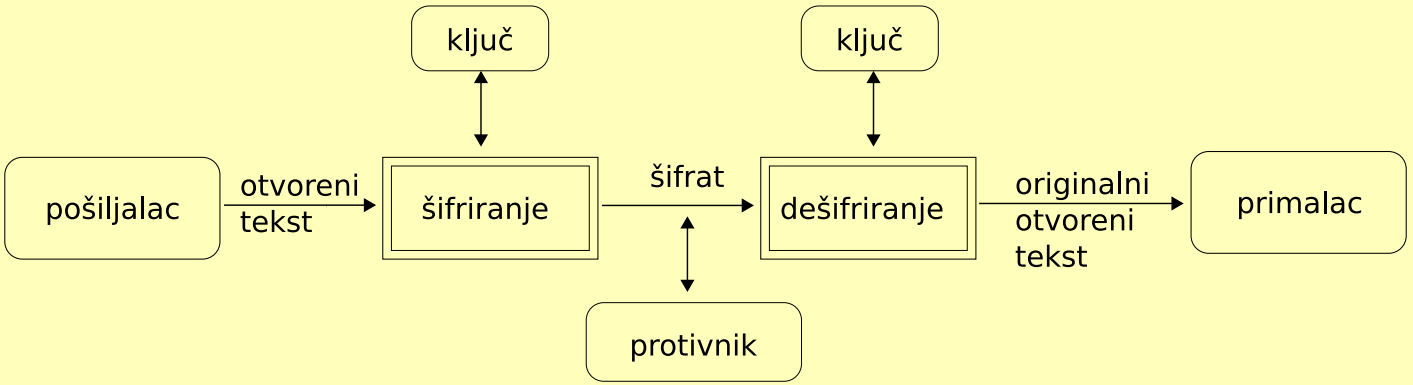
Uočimo da brojevi a, b iz prve nejednakosti zadovoljavaju
Pellovu jednadžbu $a^2 - 2b^2 = \pm 1$.

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

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

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.

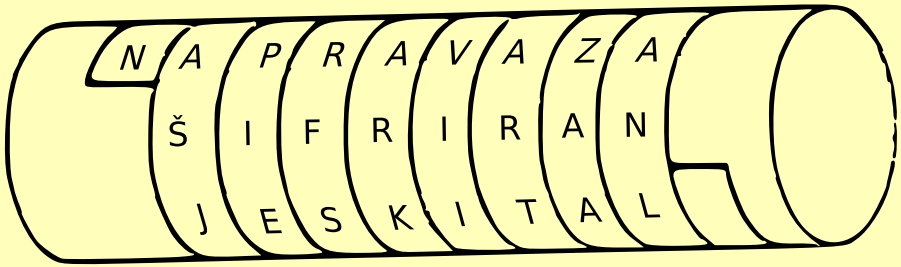


Glavne metode klasične kriptografije:

- 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

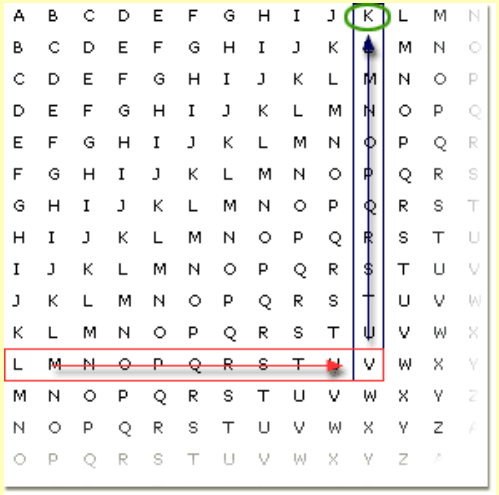
6	1	3	7	5	2	4
S	T	U	P	Č	A	N
A	T	R	A	N	S	P
O	Z	I	C	I	J	A

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

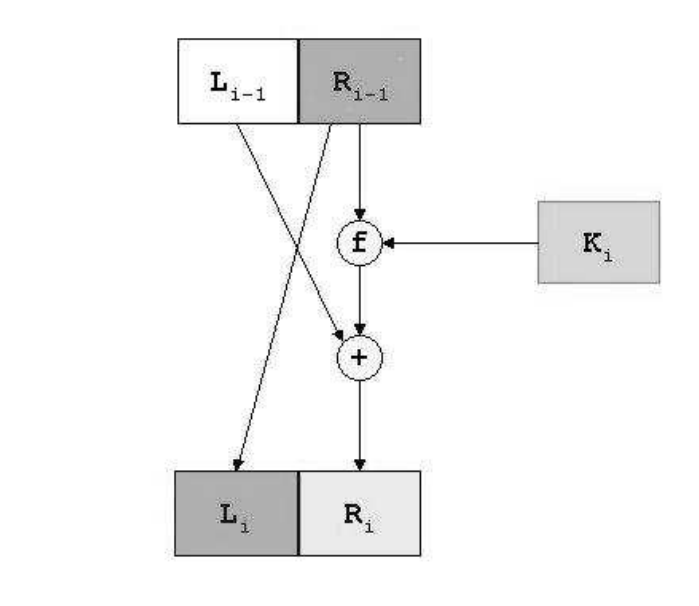


ENIGMA (1920. – 2. svjetski rat)

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

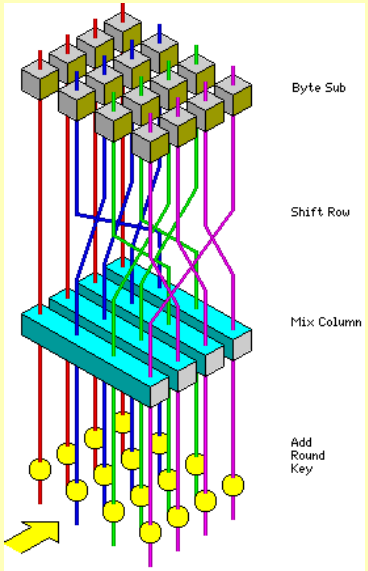
DES – Data Encryption Standard (1976. – 1998.)

- 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: $x^8 + x^4 + x^3 + x + 1$



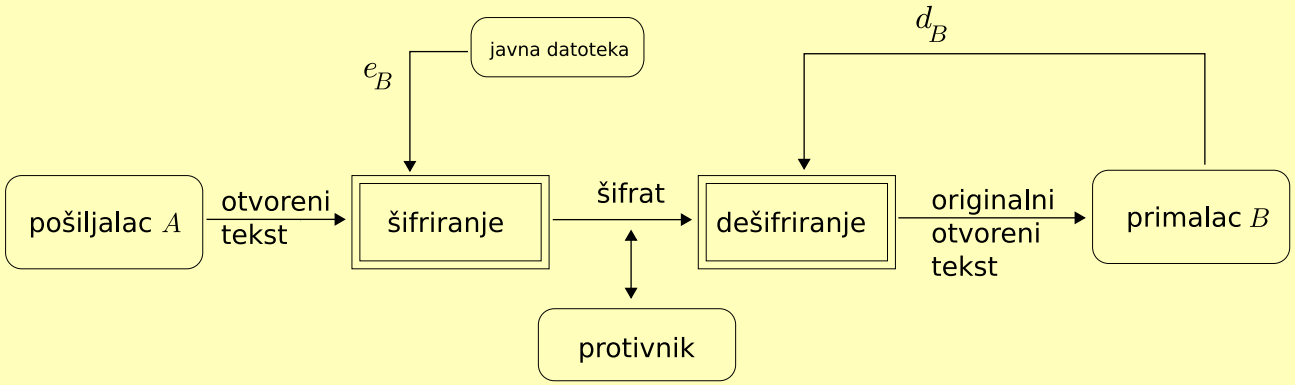
Kriptosustavi s javnim ključem

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: $z = d_A(e_B(x))$, $e_A(z) = e_B(x)$.

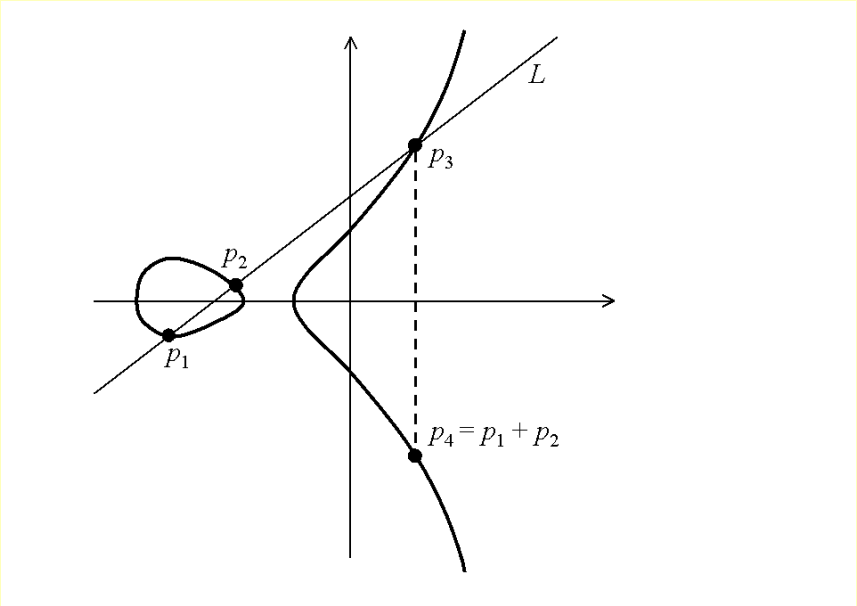
Osnova za kriptosustave s javnim ključem su "teški" matematički problemi:

- 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.

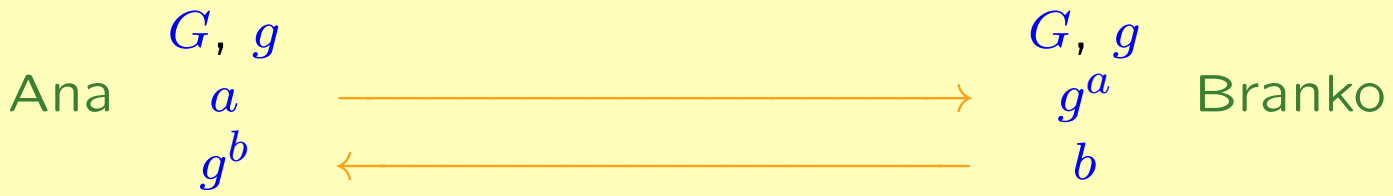
x	1	2	3	4	5	6	7	8	9	10
2^x	2	4	8	5	10	9	7	3	6	1

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	$(4, 1)$	$(6, 6)$	$(5, 0)$	$(6, 1)$	$(4, 6)$	$\mathcal{O}$



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

Ana: $(g^b)^a = g^{ab}$ ↘

Branko: $(g^a)^b = g^{ab}$ ↗

razmijenili su ključ

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.

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 $\text{nzd}(e, \varphi(n)) = 1$,
- 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 Mersenneove brojeve)

Najveći danas poznati prosti broj je Mersenneov broj $2^{57885161} - 1$ koji ima $17\,425\,170$ znamenaka (pronađen 2013. 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:

$$
\begin{aligned}
&123018668453011775513049495838496272077285356959533479219732245215172640050726365751874520219978646938995647494277406384592519255732630345373154826850791702612214291346167042921431160222124047927473779408066535141959745985690214341 \\
= &33478071698956898786044169848212690817704794983713768568912431388982883793878002287614711652531743087737814467999489 \\
\times &36746043666799590428244633799627952632279158164343087642676032283815739666511279233373417143396810270092798736308917.
\end{aligned}
$$

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

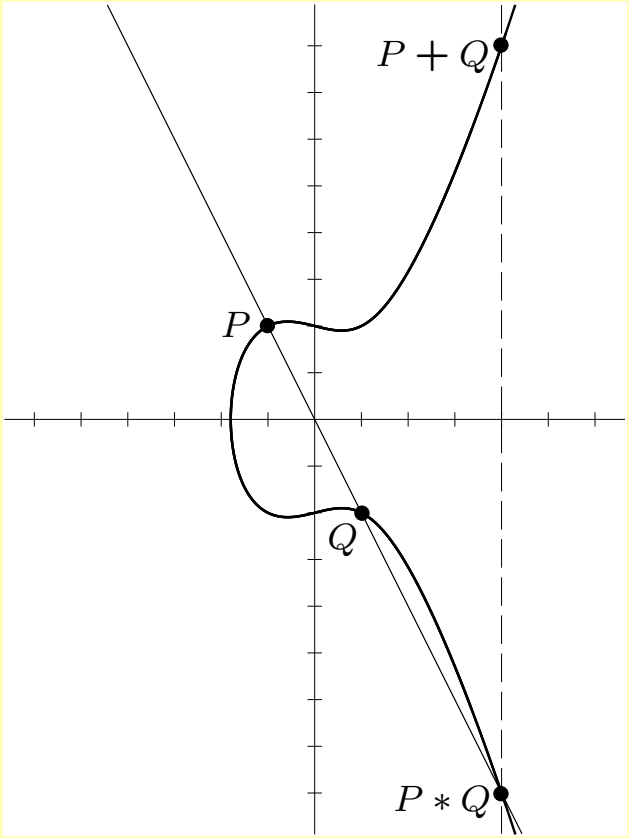
Iz jednakosti $d \cdot e - t \cdot \varphi(n) = 1$, vidimo da je t/d jako dobra aproksimacija broja $e/\varphi(n)$. Broj

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

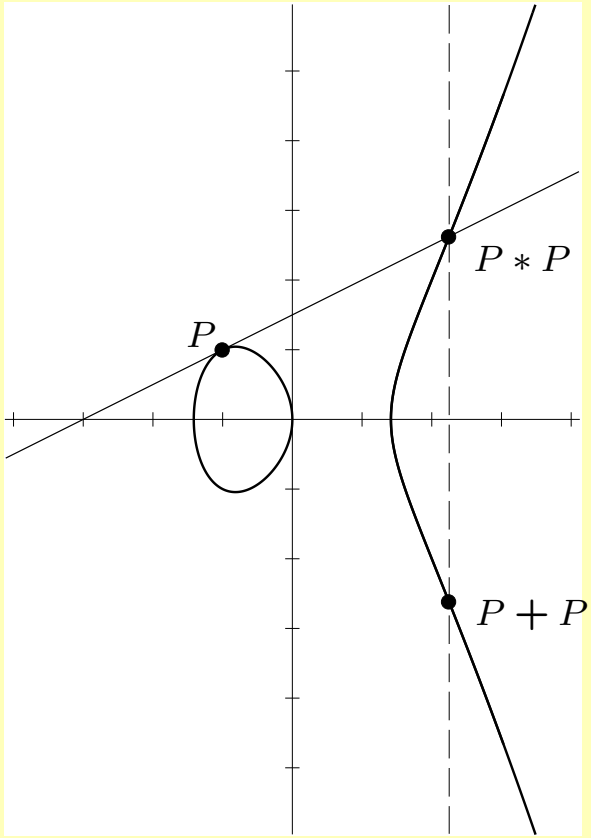
nije poznat napadaču, jer ne zna p i q. No, $\varphi(n) \approx n$ može poslužiti kao zadovoljavajuća aproksimacija. Tako dobivamo da je broj t/d, koji sadrži tajni podatak d, dobra aproksimacija broja e/n koji je u potpunosti javan.

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.

Eliptičke krivulje u kriptografiji



sekanta



tangenta

Koblitz, Miller (1985): ideja o korištenju eliptičkih krivulja u kriptografiji javnog ključa

U grupi točaka na eliptičkoj krivulji nad konačnim poljem razlika u težini između potenciranja i logaritmiranja još je veća nego u standarno korištenoj grupi $\mathbb{F}_p^*$.

Ista sigurnost uz manju duljinu ključa (umjesto ključa duljine 1024 bita, što je standardna duljina kod RSA kriptosustava te onih koji koriste $\mathbb{F}_p^*$, dovoljan je ključ duljine 160 bitova). To je osobito važno kod onih primjena (kao što su npr. pametne kartice) kod kojih je prostor za pohranu ključeva vrlo ograničen.

Za problem eliptičkog diskretnog logaritma (osim za neke vrlo specijalne eliptičke krivulje) nisu poznati bolji algoritmi od algoritama za općenite grupe, kojima je složenost $O(\sqrt{n})$, gdje je n red grupe. To su algoritam "malih i velikih koraka - BSGS ($x = \lceil n \rceil a + b$, $0 \leq a, b < \sqrt{n}$) i Pollardov ρ-algoritam (povezan s paradoksom rođendana).

Za problem diskretnog logaritma u grupi $\mathbb{F}_p^*$ postoji dosta efikasniji (subeksponencijalni) algoritmi zasnovani su na index calculus metodi.

Ideja (uspješna): shvatiti elemente od $\mathbb{F}_p^*$ kao elemente od $\mathbb{Z}$, te ih prikazati kao produkt malih prostih brojeva $\{p_1, p_2, \ldots, p_m\}$ (faktorska baza).

Ideja (neuspješna): shvatiti eliptičku krivulju na $\mathbb{F}_p$ kao krivulju nad $\mathbb{Q}$, koja može imati veći broj generatora (broj generatora zove se rang). Za realizaciju ideje trebale bi nam krivulje ranga barem 180, a danas nije poznata niti jedna krivulja ranga većeg od 28.

Druga razlika, i potencijalna prednost, grupe $E(\mathbb{F}_p)$ u odnosu na grupu $\mathbb{F}_p^*$ jest da je red grupe $\mathbb{F}_p^*$ potpuno određen s p (jednak je $p-1$), dok red grupe $E(\mathbb{F}_p)$, za različite krivulje E i fiksirani p, može poprimiti bilo koju vrijednost unutar Hasseovog intervala

$$\langle p+1-2\sqrt{p},\, p+1+2\sqrt{p}\rangle.$$

Jedna primjena ove ideje je u faktorizaciji velikih prirodnih brojeva. Polazište je Pollardova $p-1$ metoda (1974) koja koristi Mali Fermatov teorem. Neka je n prirodni broj čiji prosti faktor p želimo pronaći. Kad bismo znali neki višekratnik od $p-1$, onda bi p mogli naći (Euklidovim algoritmom) kao zajednički djelitelj od a^m-1 i n.

Pitanje je, međutim, kako naći višekratnik od $p-1$ kad ne znamo p. To možemo efikasno napraviti u slučaju kada broj $p-1$ ima samo male proste faktore, što naravno ne mora biti općenito zadovoljeno.

No, pokazuje se da je unutar Hasseovog intervala uvijek moguće pronaći broj koji je dovoljno "gladak", a time i eliptičku krivulju nad $E(\mathbb{F}_p)$ dovoljno glatkog reda. Prevođenjem Pollardove metode u grupu eliptičkih krivulja (što je predložio Lenstra (1987)), dobiva je jedna od najefikasnijih (subeksponencijalnih) danas poznatih metoda za faktorizaciju.

Daljnja poboljšanja ove metode dobivaju se promatranjem eliptičkih krivulja koje imaju veliku torzijsku grupu (grupu točaka konačnog reda) nad $\mathbb{Q}$ ili nad poljima algebarskih brojeva malog stupnja, čime se može unaprijed osigurati da red od $E(\mathbb{F}_p)$ ima neki netrivijalni faktor (red torzijske grupe).

Rekordni rangovi za eliptičke krivulje nad $\mathbb{Q}$:

rang $\geq$	godina	autori
3	1938	Billing
4	1945	Wiman
6	1974	Penney & Pomerance
7	1975	Penney & Pomerance
8	1977	Grunewald & Zimmert
9	1977	Brumer - Kramer
12	1982	Mestre
14	1986	Mestre
15	1992	Mestre
17	1992	Nagao
19	1992	Fermigier
20	1993	Nagao
21	1994	Nagao & Kouya
22	1997	Fermigier
23	1998	Martin & McMillen
24	2000	Martin & McMillen
28	2006	Elkies

`http://web.math.pmf.unizg.hr/~duje/tors/rankhist.html`

$$B(T) = \sup\{\mathrm{rank}\,(E(\mathbb{Q})) \,:\, E(\mathbb{Q})_{\mathrm{tors}} \cong T\}$$

T	$B(T) \geq$	autori
0	28	Elkies (2006)
$\mathbb{Z}/2\mathbb{Z}$	19	Elkies (2009)
$\mathbb{Z}/3\mathbb{Z}$	13	Eroshkin (2007,2008,2009)
$\mathbb{Z}/4\mathbb{Z}$	12	Elkies (2006)
$\mathbb{Z}/5\mathbb{Z}$	8	D. & Lecacheux (2009), Eroshkin (2009)
$\mathbb{Z}/6\mathbb{Z}$	8	Eroshkin (2008), D. & Eroshkin (2008), Elkies (2008), D. (2008), D. & Peral (2012)
$\mathbb{Z}/7\mathbb{Z}$	5	D. & Kulesz (2001), Elkies (2006), Eroshkin (2009), D. & Lecacheux (2009), D. & Eroshkin (2009)
$\mathbb{Z}/8\mathbb{Z}$	6	Elkies (2006), D., MacLeod & Peral (2013)
$\mathbb{Z}/9\mathbb{Z}$	4	Fisher (2009)
$\mathbb{Z}/10\mathbb{Z}$	4	D. (2005,2008), Elkies (2006)
$\mathbb{Z}/12\mathbb{Z}$	4	Fisher (2008)
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$	15	Elkies (2009)
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$	9	D. & Peral (2012)
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$	6	Elkies (2006)
$\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z}$	3	Connell (2000), D. (2000,2001,2006,2008), Campbell & Goins (2003), Rathbun (2003,2006), Flores, Jones, Rollick & Weigandt (2007), Fisher (2009)

`http://web.math.pmf.unizg.hr/~duje/tors/tors.html`