Rozhledy matematicko-fyzikální

Zuzana Masáková

Prvočísla v akci

Rozhledy matematicko-fyzikální, Vol. 90 (2015), No. 1-2, 66–77

Persistent URL: http://dml.cz/dmlcz/146618

Prvočísla v akci

Zuzana Masáková, FJFI ČVUT, Praha

Abstract. Mathematicians have been fascinated by the concepts of elementary number theory, such as integers, integer division, prime and relatively prime numbers, since the times of Ancient Greece. Problems based on these concepts are easy to pose, but their solutions are often very complicated. Some of these problems, e.g. the twin prime hypothesis, though very old, still represent a great challenge for mathematicians. Elementary number theory might seem full of mathematical recreations which have been nothing more than a source of entertainment for centuries. Although we do not consider this purpose pointless, we shall try to convince the reader that many applications of every day use are actually based on number theory principles. In fact, elementary number theory is a basic material for the construction of cryptographic systems used in e-mail communication or for credit card password authentication.

Šifrování s veřejně přístupným klíčem

Snad od chvíle, kdy lidé zjistili, jak důležité je komunikovat s ostatními, si zároveň uvědomovali, že možná ještě důležitější může být informace před nepovolanýma ušima umět skrývat. Šifrovacích metod bylo v historii vymyšleno spoustu. Asi nejjednodušší je substituční kódování, při kterém se nahrazuje každé písmeno jiným, a klíčem je tabulka nahrazovacích pravidel. Už například při substituci

$$A \mapsto B,\ B \mapsto C,\ C \mapsto D,\ D \mapsto E \text{ atd.}$$

zašifrovaný text

BI, KBL NJMVKJ NBUFNBUJLV!

vypadá na první pohled úplně nečitelně, ve skutečnosti je ale tento typ šifry i bez klíče velice snadno rozluštitelný, například s využitím znalosti frekvencí písmen a skupin písmen v přirozeném jazyce.

Mnohem bezpečnější je šifra pomocí tzv. jednorázové náhodné pásky. Ta má ovšem nevýhodu, že vyžaduje bezpečný způsob, jak předat druhé straně velmi dlouhý klíč, který nelze použít opakovaně.

My se ovšem zaměříme na šifrovací metodu, která využívá vlastností přirozených čísel a prvočísel. Naznačíme, jak taková metoda funguje. Představme si následující situaci: Andula a Bohouš hrají po telefonu šachy.[1)] Hodina pokročila a hra stále není u konce. Proto se rozhodnou ji přerušit. Jak ale zařídit, aby jeden z nich neměl na rozmyšlení svého tahu celou noc? Na šachových turnajích se poslední tah zaznamená do obálky a uschová u rozhodčího. Andula s Bohoušem ale nemají nezávislou třetí osobu. Musí zajistit, aby Andula mohla svůj tah zapsat do zprávy, kterou Bohouš nebude moci sám bez nějakého „klíče“ rozluštit. Zároveň ale Andula nesmí mít možnost si svůj tah přerozmyslet.

Řešení jejich problému je snadné: Andula s Bohoušem se dohodnou na nějakém způsobu, jak zakódovat šachový tah do čísel. Například chce-li Andula sdělit tah jezdce na c2, může napsat třeba 1132, protože „11“ znamená „J“ (jedenácté písmeno v abecedě bez diakritiky), „3“ znamená „c“ a „2“ znamená „2“. Takto zakódují každý tah do čtyřciferného čísla.

V dalším kroku Andula doplní čtyřčíslí 1132 na stomístné prvočíslo $p = 1132\cdots$. S pomocí počítače je to velmi snadné. Program Maple na osobním počítači našel nejmenší takové prvočíslo okamžitě,

$$p = 11320000000000000000000000000000000000000000000000000\\ 00000000000000000000000000000000000000000000000000321.$$

Samozřejmě je lepší, aby Andula mezi všemi stomístnými prvočísly vybrala nějaké náhodně. To lze provést tak, že Maple nechá testovat prvočíselnost n v náhodně zvoleném intervalu stomístných čísel s daným čtyřčíslím na začátku. Vyjde například

$$p = 11322242417046519313981641539571886053400080187484459427552056869842622684442996417214180694607189 89.$$

Pak Andula obdobným způsobem najde ještě jedno stomístné prvočíslo q, řekněme větší než p. Tato dvě prvočísla mezi sebou vynásobí a Bohoušovi sdělí výsledek $n = pq$. Ráno Andula Bohoušovi prozradí q, ten snadno vydělí $p = n/q$ a přečte si z prvního čtyřčíslí Andulin zamýšlený tah.

Teď je jasné, že Andula si svůj tah do rána nemůže přerozmyslet. Kdyby chtěla zaměnit původní tah za jiný, musela by dodat jiné stomístné prvočíslo $\tilde{p}$. Bohouš by snadno ověřil, že ho chce Andula ošálit.

[1)] V anglické kryptologické literatuře jsou osoby A a B komunikující v šifrovacích úlohách pojmenovávány zpravidla Alice a Bob.

Jednoduše by vyzkoušel, zda je n dělitelné $\tilde{p}$. To by mu na papíře nějakou dobu trvalo, ale Maple to zvládne ještě rychleji než cobydup.

Kdyby ale chtěl podvádět Bohouš, musel by bez klíče q sám zjistit, jaký je rozklad čísla n na prvočinitele. A v tom je zakopaný pes: Ačkoliv samotné dělení je rychlé, i v nejrychlejším známém algoritmu pro hledání rozkladu na prvočinitele je počet kroků exponenciálně závislý na počtu cifer rozkládaného čísla. Rozložit n na součin pq by tak Bohoušovi mohlo trvat i s použitím nejrychlejšího počítače několik miliard let. Oba hráči tak mohou být ujištěni, že hra je spravedlivá.

Trochu jiná situace než v předchozím příkladě by byla, kdyby si Andula s Bohoušem chtěli posílat zprávy tak, aby je slídilka Evelína nemohla číst.[2)] Průběh jejich dopisování by se dal ilustrovat asi takto: Protože obyčejná obálka poslaná poštou není před Evelínou dostatečně zabezpečená, má Andula na veřejném místě schránku, ke které jen ona vlastní klíč. Když jí chce Bohouš napsat zprávu, donese dopis do schránky a zabouchne dvířka. Takto velmi zjednodušeně by se dal popsat princip, na kterém je založen algoritmus RSA, který se používá například při zabezpečené internetové komunikaci, například v protokolu SSL.

RSA

Šifrovací algoritmus RSA dostal název podle tří matematiků, kteří ho v roce 1977 navrhli. Byli to Ronald Linn Rivest, Adi Shamir a Leonard Max Adleman. Zkratka RSA odpovídá počátečním písmenům jejich příjmení v pořadí, v jakém byli uvedeni na článku, který algoritmus popisoval.

V algoritmu budeme využívat tzv. Eulerovu funkci φ, která danému přirozenému číslu n přiřazuje počet celých kladných čísel menších nebo rovných n nesoudělných s n. Například když p je prvočíslo, pak všechna přirozená k ostře menší než p jsou s p nesoudělná, takže $\varphi(p) = p - 1$.

Pro zajímavost spočítejme ještě hodnotu Eulerovy funkce pro přirozené číslo n, které je součinem dvou různých prvočísel p a q. Potřebujeme zjistit počet čísel soudělných s $n = pq$. Číslo je soudělné s n, pokud je násobkem nějakého jeho prvočíselného dělitele. Je-li n součin dvou prvočísel $n = pq$, pak soudělné s ním jsou jen násobky p a násobky q. Násobky p menší nebo rovny n jsou $p, 2p, 3p, \ldots, qp$, je jich tedy q. Po-

[2)] V anglické literatuře je odposlouchávajícím padouchem zásadně Eva, tedy Eve, s odkazem na slovo ‚eavesdropper'.

dobně násobky q menší nebo rovny n jsou $q, 2q, 3q, \ldots, pq$, jejich počet je p. Čísel nesoudělných s pq je tedy $\varphi(pq) = pq - q - p + 1$, přičemž bereme v úvahu, že poslední ze soudělných čísel, tj. číslo pq, je zároveň násobek p i q, a my ho chceme odečíst pouze jednou. Odtud už snadno vytknutím plyne, že

$$\varphi(pq) = (p-1)(q-1). \tag{1}$$

Vraťme se nyní k algoritmu RSA. Nejprve vysvětlíme postup, poté na příkladě ilustrujeme, že pracuje tak, jak má. Andula zvolí dvě velká prvočísla p, q. Jejich součinem získá $n = pq$. Pak zvolí náhodné i nesoudělné s číslem $\varphi(n) = (p-1)(q-1)$, v rozmezí $1 < i < \varphi(n)$. Při zvolení náhodného čísla i z $\{2, 3, \ldots, \varphi(n) - 1\}$ je třeba ověřit nesoudělnost i a $\varphi(n)$. To se provádí Eukleidovým algoritmem hledání $\mathrm{nsd}(i, \varphi(n))$. Z něj zároveň vyjdou koeficienty j a k takové, že $ij + k\varphi(n) = \mathrm{nsd}(i, \varphi(n)) = 1$. Takové j lze navíc zvolit opět v množině $\{2, 3, \ldots, \varphi(n) - 1\}$.

Andula tedy nyní má $n = pq$ a $ij \equiv 1 \mod \varphi(n)$.[3)] Čísla n, i tvoří veřejný klíč. Naopak čísla $p, q, \varphi(n), j$ si nechává pro sebe. Bohouš nyní může s pomocí veřejného klíče posílat Andule zašifrované zprávy tak, aby je nikdo, kromě Anduly, nemohl rozluštit. Nejprve převede svou zprávu na číslo v množině $\{1, \ldots, n-1\}$. To by šlo například tak, že text, digitalizovaný na posloupnost nul a jedniček, rozseká na bloky takové délky m, aby $2^m < n$. Každý z bloků pak odpovídá číslu $x < n$. Aby text vždy šel rozdělit na tyto bloky, je třeba ho doplnit na délku dělitelnou m například posloupností cifer $10\cdots0$, případně 1. Pokud text sám má délku dělitelnou m, přidá se jeden celý blok $10\cdots0$ tak, aby při dekódování nebyl deformován význam. Bude totiž jasné, že ze zaslané zprávy je nutné vždy odebrat poslední jedničku a všechny nuly za ní.

Zašifrování probíhá takto: Bohouš umocní zprávu x na exponent i z veřejného klíče a spočítá zbytek po dělení číslem n. Získané číslo $y \equiv x^i \mod n$ pak pošle Andule.

Andula přečte y a zprávu rozšifruje pomocí privátního klíče j, který si schovala. Provede $y^j \mod n$. Výsledkem je původní zpráva, tedy číslo x.

Předveďme na příkladě, že uvedený postup opravdu funguje.

Příklad. Zvolme „velká“ prvočísla $p = 47$, $q = 67$. Potom $n = 47 \cdot 67 = 3149$ a $\varphi(n) = 46 \cdot 66 = 3036$. Jako šifrovací exponent do veřejného

[3)] Tento zápis znamená, že zbytek po dělení čísla ij číslem $\varphi(n)$ je roven 1. Obecně budeme psát $a \equiv b \mod m$, když čísla a a b mají stejný zbytek po dělení číslem m.

klíče zvolme například $i = 13$. K ověření nesoudělnosti i s číslem $\varphi(n)$ použijeme Eukleidův algoritmus:

$$\begin{aligned} 3036 &= 233 \cdot 13 + 7 \\ 13 &= 1 \cdot 7 + 6 \\ 7 &= 1 \cdot 6 + 1 \end{aligned}$$

Proto $\mathrm{nsd}(3036, 13) = 1$ a zároveň lze z těchto rovnic postupně odvodit, že

$$1 = 7 - 6 = 7 - (13 - 7) = 2(3036 - 233 \cdot 13) - 13 = 2 \cdot 3036 - 467 \cdot 13.$$

Máme tedy řešení $k = 2$, $j = -467$ rovnice $3036k + 13j = 1$, jenže bychom rádi řešení jiné, takové, kde $j \in \mathbb{N}$, $j < 3036$. To najdeme snadno:

$$1 = (2 - 13) \cdot 3036 + (-467 + 3036) \cdot 13 = -11 \cdot 3036 + 2569 \cdot 13.$$

Zvolíme-li nyní $j = 2569$, dostaneme

$$ij = 13 \cdot 2569 = 33397 \equiv 1 \mod 3036.$$

Veřejným klíčem je tedy $n = 3149$, $i = 13$. Číslo $j = 2569$ si naopak Andula pečlivě ukryje.

Bohouš chce šifrovat dejme tomu zprávu zdigitalizovanou na

$$01000001010|01000010011|11010010101, \tag{2}$$

kde už jsme pro přehlednost vyznačili rozdělení na bloky tak, aby každý blok odpovídal binárnímu zápisu čísla menšího než $n = 3149$. Proto bereme bloky délky 11. (Největší číslo s jedenácti binárními ciframi je totiž $2^{11} - 1 = 2047 < n$.) Bloky v Bohoušově zprávě tedy odpovídají číslům $x = 522$, $x = 531$ a $x = 1684$.

Zašifrujeme $y \equiv x^j \mod n$, tj. konkrétně

$$\begin{aligned} 522^{13} &\equiv 1077 \mod 3149, \\ 531^{13} &\equiv 1901 \mod 3149, \\ 1685^{13} &\equiv 1761 \mod 3149. \end{aligned}$$

Zašifrovanou zprávu tedy vyšleme ve tvaru

$$10000110101|11101101101|11011100001.$$

Andula zprávu přijme a rozšifruje $y^j \mod n$, tj.

$$1077^{2569} \equiv 522 \mod 3149,$$
$$1901^{2569} \equiv 531 \mod 3149,$$
$$1761^{2569} \equiv 1685 \mod 3149.$$

Poněkud se divíme, že Andule s Bohoušem stálo za to vynaložit tolik úsilí, aby si zašifrovali zprávu „AHOJ“, která je digitalizovaná v podobě posloupnosti (2) s pomocí ASCII kódu (po odebrání poslední 1, která dorovnávala čtyři osmibitové bloky na délku dělitelnou 11). Binárně totiž máme

$$01000001 = (65)_2 = \mathrm{A},$$
$$01001000 = (72)_2 = \mathrm{H},$$
$$01001111 = (79)_2 = \mathrm{O},$$
$$01001010 = (74)_2 = \mathrm{J}.$$

Poznamenejme, že ve skutečnosti je samozřejmě nutné zvolit mnohem větší prvočísla, než jsme předvedli v předchozím příkladě. Zde by totiž získání utajeného dešifrovacího klíče nebylo vůbec těžké. Stačilo by faktorizovat n na prvočinitele $n = pq$, odtud zjistit $\varphi(n) = (p-1)(q-1)$ a pak najít j podobně, jako jsme to udělali v příkladu my.

Důkaz platnosti šifrovacího algoritmu, totiž faktu, že zašifrované $y \equiv x^i \mod n$ lze rozluštit jako $x \equiv y^j \mod n$, je elementární, zvídavý čtenář se může pokusit ho sám nalézt. Jako nápovědu uvedeme, že k ověření $x^{ij} \equiv x \mod n$ se využívá tzv. Eulerova věta.

Věta 1 (Eulerova). *Mějme navzájem nesoudělná přirozená čísla n, a. Pak $a^{\varphi(n)} \equiv 1 \mod n$.*

Sluší se dodat, že při výpočtech v rámci algoritmu RSA je třeba pracovat s obrovskými čísly. Už v našem „malém“ příkladě není výhodné při šifrování mocnit

$$522^{13} = 213656711189633749233078088598986752$$

a pak dělit číslem 3149. Pro zrychlení výpočtu se využívá modulární aritmetika. Všimneme si, že

$$13 = 1 + 12 = 1 + 2 \cdot 6 = 1 + 2 \cdot 2 \cdot 3 = 1 + 2 \cdot 2 \cdot (1 + 2), \qquad (3)$$

takže $522^{13} \mod 3149$ získáme postupem

$$\begin{aligned} 522^2 &\equiv 272484 \equiv 1670 \mod 3149 \\ 522^3 &\equiv 522 \cdot 1670 \equiv 2616 \mod 3149 \\ 522^6 &\equiv 2616^2 \equiv 679 \mod 3149 \\ 522^{12} &\equiv 679^2 \equiv 1287 \mod 3149 \\ 522^{13} &\equiv 522 \cdot 1287 \equiv 1077 \mod 3149, \end{aligned}$$

kde následující řádek je vždy druhá mocnina nebo 522 násobek předchozího v souladu s (3). V každém kroku rovnou provedeme zbytek po dělení číslem 3149, takže tímto postupem nemusíme nikdy pracovat s čísly většími než 3148^2.[4)]

Testování prvočíselnosti

Je zřejmé, že podstatnou složkou šifrovacích problémů je možnost snadného hledání náhodných velkých prvočísel. Jak tedy rozhodovat o tom, jestli dané číslo je prvočíslo, nebo jestli je složené?

Je 123456 prvočíslo? Ne! Vidíme, že je sudé.

Je 1234567 prvočíslo? Ne. Ovšem nejmenší netriviální dělitel je až 127.

Je 1234577 prvočíslo? Tentokrát ano, ale abychom to ověřili, museli bychom např. zkoušet dělitelnost všemi prvočísly menšími nebo rovnými číslu $\lfloor\sqrt{1234567}\rfloor = 1111$[5)]. Těch je 186.

Je $n = 1111222233334444555566667777888899967$ prvočíslo? Tady bychom s běžným testováním neuspěli. Prvočísel menších nebo rovných $\sqrt{n}$ je totiž více než $2{,}4 \cdot 10^{16}$. Kdyby nám každé dělení trvalo 1 milisekundu, pak by celé prověřování zabralo téměř půl milionu let. Výsledkem by bylo, že dané číslo je opravdu prvočíslem. Otázkou tedy je, jak rozhodování o prvočíselnosti daného n urychlit.

Nejjednodušší je tzv. Fermatův test, založený na malé Fermatově větě, která je vlastně speciálním případem věty Eulerovy, když volíme za n prvočíslo.

Věta 2 (malá Fermatova). *Mějme prvočíslo n a s ním nesoudělné přirozené číslo a. Pak $a^{n-1} \equiv 1 \mod n$.*

[4)] Ve skutečnosti lze mocnění velkých přirozených čísel v modulární aritmetice počítat ještě rychleji, s využitím takzvané Montgomeryho redukce.

[5)] Symbol $\lfloor x \rfloor$ značí tzv. dolní celou část reálného čísla x, například $\lfloor 2{,}7 \rfloor = 2$, $\lfloor -2{,}7 \rfloor = -3$.

Malá Fermatova věta říká, že je-li n prvočíslo a a není jeho násobek, pak $a^{n-1} \equiv 1 \bmod n$, jinými slovy $a^{n-1} - 1$ je dělitelné číslem n. Přeformulujeme-li tvrzení pomocí obrácené implikace, vidíme, že najdeme-li $a \in \{1, 2, \ldots, n-1\}$ takové, že n nedělí $a^{n-1} - 1$, pak je nutně n složené. A opravdu: pro namátkou vybraná složená lichá čísla n stačí dokonce zvolit $a = 2$ a Fermatův test prokáže složenost n. Např. $n = 9$ nedělí $2^8 - 1 = 255$ nebo $n = 21$ nedělí $2^{20} - 1 = 1048575$. V těchto příkladech Fermatův test potvrdil složenost čísla n už pro $a = 2$. Co ale můžeme říct o n, pokud nastane $2^{n-1} \equiv 1 \mod n$? Implikaci v malé Fermatově větě nelze obrátit, takže takové n může být prvočíslo, ale také nemusí.

Příklad. Mějme $n = 341 = 11 \cdot 31$. Ověříme, že číslo 341 dělí $2^{340} - 1$. Je totiž

$$2^{340} - 1 = (2^5)^{68} - 1 = \underbrace{(2^5 - 1)}_{31}\big[(2^5)^{67} + (2^5)^{66} + \cdots + 2^5 + 1\big],$$

z čehož plyne, že 31 dělí $2^{340} - 1$, ale rovněž

$$2^{340} - 1 = (2^{10})^{34} - 1 = (2^{10} - 1)\big[(2^{10})^{33} + (2^{10})^{32} + \cdots + 2^{10} + 1\big],$$

odkud odvodíme, že $2^{10} - 1$ dělí $2^{340} - 1$. Přitom podle malé Fermatovy věty víme, že 11 dělí $2^{10} - 1$, takže dělí také $2^{340} - 1$. Protože 11 a 31 jsou navzájem nesoudělná a obě dělí $2^{340} - 1$, tak i jejich součin $31 \cdot 11 = 341$ dělí $2^{340} - 1$. Číslo $n = 341$ se tedy v rámci Fermatova testu vzhledem k $a = 2$ tváří jako prvočíslo.

Je-li n složené číslo a $a \in \mathbb{N}$, pak mohou nastat dvě situace:

- $a^{n-1} \not\equiv 1 \bmod n$ a číslo a se nazývá Fermatův svědek složenosti čísla n.
- $a^{n-1} \equiv 1 \bmod n$ a číslo n se nazývá pseudoprvočíslo vzhledem k bázi a.

Poznamenejme, že číslo $n = 341$ je nejmenší pseudoprvočíslo vzhledem k bázi 2. Už $a = 3$ je ale svědek složenosti čísla 341, protože platí $3^{340} \equiv 56 \bmod 341$.

Mezi $a \in \{1, 2, \ldots, n-1\}$ svědkem složenosti jistě nemůže být $a = 1$, ale ani $a = n - 1$. Platí totiž

$$(n-1)^{n-1} = n^{n-1} - \tbinom{n-1}{1} n^{n-2} + \tbinom{n-1}{2} n^{n-3} - \cdots + (-1)^{n-1} \equiv$$
$$\equiv (-1)^{n-1} \mod n,$$

takže n dělí $(n-1)^{n-1} - 1$. (Uvažujeme pouze liché n, protože pro sudá čísla test prvočíselnosti nemá smysl provádět.)

Fermatův test nám může možná rychle označit složené číslo za složené, existence pseudoprvočísel nicméně vypovídá o tom, že v opačném směru je to poněkud složitější. Abychom měli kritérium prvočíselnosti (tedy pravidlo „Pokud něco, pak n je prvočíslo, a pokud ne, pak n je složené.“) musíme vzít v úvahu následující větu.

Věta 3. *Přirozené číslo n je prvočíslo právě tehdy, když pro každé $a \in \{1, 2, \ldots, n-1\}$ platí $a^{n-1} \equiv 1 \mod n$.*[6)]

Implikace zleva doprava snadno plyne z malé Fermatovy věty 2. Čísla $a \in \{1, 2, \ldots, n-1\}$ jsou totiž všechna nesoudělná s prvočíslem n. Opačnou implikaci lze přeformulovat následovně: Je-li n složené, pak existuje $a \in \{1, 2, \ldots, n-1\}$ takové, že $a^{n-1} \not\equiv 1 \bmod n$. Není těžké ukázat, že takovým příkladem je každé číslo a, které má s n společného netriviálního dělitele. Fermatův test nám tedy označí složené číslo za složené, kdykoliv za a zvolíme číslo soudělné s n. Jaká je ale pravděpodobnost, že při náhodném výběru $a \in \{2, \ldots, n-2\}$ narazíme na číslo soudělné s n? To samozřejmě závisí na n. Protože ale charakter n dopředu neznáme, je nutno počítat s nejhorším případem.

Příklad. Je-li n součinem dvou velkých prvočísel, $n = pq$, pak počet čísel menších nebo rovných n soudělných s n je podle dříve spočítaného (1) roven $n - \varphi(n) = p + q - 1$. Pravděpodobnost volby soudělného a je tedy

$$\frac{n - \varphi(n)}{n} = \frac{p + q - 1}{pq}.$$

Pokud p, q jsou např. 100místná prvočísla, pak $p + q - 1 < 2 \cdot 10^{100}$, a přitom $pq > 10^{198}$, takže

$$\frac{p + q - 1}{pq} < \frac{2}{10^{98}}.$$

Chtěli-li bychom tedy použít Fermatův test, musíme doufat, že pro složená čísla n narazíme na svědka složenosti častěji než jen při volbě a soudělného s n. Pro $n = 77$ jsou například svědky složenosti všechna $a \in \{2, \ldots, 75\}$ kromě $a = 34$ a $a = 43$. Z možných 74 čísel a je tedy 72 svědků složenosti, přitom čísel soudělných s $n = 77$ je mezi nimi pouze 16.

[6)] Věta 3 charakterizuje prvočísla. čtenář si ale jistě všiml, že pokud bychom chtěli dokazovat prvočíselnost n jejím použitím, stálo by nás to více úsilí než prosté pokusné dělení.

I tato vlastnost ale není obecná a naše doufání v použitelnost Fermatova testu je marné. Existují totiž tzv. Carmichaelova čísla, což jsou ta složená čísla $n \in \mathbb{N}$, pro která $a^{n-1} \equiv 1 \bmod n$, jakmile je a nesoudělné s n. Carmichaelova čísla n jsou tedy pseudoprvočísla vzhledem ke všem bázím a nesoudělným s n a Fermatův test u nich selhává.

Příklad. Nejmenším Carmichaelovým číslem je $n = 561 = 3 \cdot 11 \cdot 17$. Číslo 561 dělí $a^{560} - 1$ pro každé a, které není dělitelné 3, 11 ani 17.

Dá se ovšem ukázat, že v případě, kdy n je složené číslo a není Carmichaelovo, je Fermatových svědků jeho složenosti dostatek (alespoň polovina).

O Carmichaelových číslech je známo mnoho dalších zajímavých věcí, například to, že nejsou dělitelná druhou mocninou žádného prvočísla, nebo to, že pokud prvočíslo p dělí Carmichaelovo číslo n, pak $p - 1$ dělí $n - 1$. Nicméně neexistuje žádný snadný (tj. rychlý) způsob, jak rozhodnout, zda dané číslo je nebo není Carmichaelovo, a to je hlavní překážka pro to, aby byl výše uvedený Fermatův test prvočíselnosti použitelný v praxi. I když šance, že náhodně vybrané přirozené číslo je zrovna Carmichaelovo, je malá, přesto není zanedbatelná pro účely tak důležité, jako je kryptografie.

Proto si zde uvedeme jiný test prvočíselnosti, který tuto slabinu Fermatova testu odstraňuje, a to tzv. Millerův–Rabinův test. Je to sice test pravděpodobnostní, což znamená, že prvočíselnost zvoleného p nikdy nevíme se 100% jistotou, ale protože pravděpodobnost mylného výsledku exponenciálně klesá s počtem průchodů testu, můžeme se na jeho výsledek spolehnout více než na nezávadnost hardwaru, který k testování používáme.

Myšlenka Millerova–Rabinova testu pouze rozvíjí test Fermatův; my ji nejprve ilustrujeme na příkladě:

Příklad. Předvedeme, jak Millerův–Rabinův test odhalí složenost Carmichaelova čísla $n = 561$. Zvolíme náhodné $a < 561$. Kdybychom měli štěstí a zvolili číslo soudělné s 561, pak 561 nedělí $a^{560} - 1$. Pokud ale sáhneme na jiné a, pak pak 561 dělí $a^{560} - 1$. Číslo $a^{560} - 1$ můžeme rozložit podle vzorce $A^2 - B^2 = (A + B)(A - B)$ postupně na součin

$$a^{560} - 1 = (a^{280} + 1)(a^{280} - 1) =$$
$$= (a^{280} + 1)(a^{140} + 1)(a^{140} - 1) = (a^{280} + 1)(a^{140} + 1)(a^{70} + 1)(a^{70} - 1) =$$
$$= (a^{280} + 1)(a^{140} + 1)(a^{70} + 1)(a^{35} + 1)(a^{35} - 1).$$

Bylo-li by číslo 561 prvočíslem, muselo by dělit alespoň jednu ze závorek v součinu.[7] Jakmile najdeme a, pro které 561 nedělí žádnou ze závorek, rozhodneme o tom, že číslo 561 musí být složené. To je pravda například už pro $a = 2$. Snadno totiž ověříme, že čísla $2^{280} + 1$, $2^{140} + 1$, $2^{70} + 1$, $2^{35} - 1$ nejsou dělitelná 3, a číslo $2^{35} + 1$ zase není dělitelné 17.

Algoritmus Millerova–Rabinova testu ilustrovaného na předchozím případě lze zapsat takto:

Vstup: liché přirozené číslo $n > 3$, parametr spolehlivosti testu k.
Postupným dělením čísla $n - 1$ dvěma najdi $m \in \mathbb{N}$ liché a $t \in \mathbb{N}$ tak, že $n - 1 = 2^t m$.
Opakuj k krát smyčku S:
    Vyber náhodné $a \in \{2, 3, \ldots, n - 2\}$ a polož $x := a^m \bmod n$.
    Když $x = 1$ nebo $x = n - 1$, začni novou iteraci smyčky S,
    jinak polož $i = 1$ a dokud $i < t$ prováděj
        $x := x^2 \bmod n$; $i := i + 1$;
        když $x = 1$, vrať „n je složené“ a konec;
        když $x = n - 1$, začni novou iteraci smyčky S.
    Vrať „n je složené“ a konec.
Vrať „n je pravděpodobně prvočíslo“.
Výstup: „n je složené“, je-li n složené, jinak „n je pravděpodobně prvočíslo“.

Všimněme si, že algoritmus je sestaven tak, aby byla minimalizovaná výpočetní náročnost. Proto se při ověřování dělitelnosti závorek v rozkladu

$$a^{n-1} - 1 = \left(a^{2^{t-1}m} + 1\right)\left(a^{2^{t-2}m} + 1\right) \cdots (a^m + 1)(a^m - 1)$$

postupuje od poslední závorky. Pokud $a^m \equiv 1 \bmod n$, znamená to, že $a^m - 1$ je dělitelné n. Pokud $a^m \equiv n - 1 \bmod n$, je $a^m + 1$ dělitelné n. V obou případech začínáme novou smyčku s jinou volbou a, protože n se vzhledem k tomuto a tváří jako prvočíslo. Pokud zbytek $a^m \bmod n$ není ani 1 ani $n - 1$, pokračujeme mocněním na druhou od a^m po $a^{2^t m}$ a zjišťujeme zbytek po dělení n. Jakmile narazíme na zbytek 1, nemá cenu v mocnění pokračovat, protože $a^{2^i m} \equiv 1 \bmod n$ implikuje $a^{2^j m} \equiv \equiv 1 \bmod n$ pro všechna $j \geq i$, a tudíž žádná ze závorek $(a^{2^j m} + 1)$ není

[7] Využíváme tuto vlastnost prvočísel: Pokud p dělí součin přirozených čísel, pak dělí alespoň jeden z činitelů.

dělitelná n. V tomto případě je číslo n složené. Pokud narazíme na zbytek $n-1$, narazili jsme na závorku, která je dělitelná n, a proto se opět n vzhledem k tomuto a tváří jako prvočíslo. Konečně pokud zbytky všech čísel $a^m, \ldots, a^{2^t m}$ jsou různé od 1 i $n-1$, pak žádná ze závorek není dělitelná n a a svědčí pro složenost čísla n.

Lze ukázat, že pro každé složené číslo je při Millerově–Rabinově testu více než 3/4 svědků složenosti. Proto při jednom průchodu smyčkou je pravděpodobnost lživé odpovědi „n je prvočíslo“ menší než 1/4. Pokud se n tváří jako prvočíslo pro k různých náhodně zvolených čísel $a \in \{2, 3, \ldots, n-2\}$, pak je pravděpodobnost chybného označení n za prvočíslo menší než $1/4^k$.

Pro toho, komu by nestačila ani pravděpodobnost $1 - 1/4^{100}$ dosažená při sto průchodech výše uvedeným testem, uveďme, že v roce 2002 byl sestaven algoritmus, který rozhodne se stoprocentní jistotou, zda n je prvočíslo, a to v čase polynomiálně závislém na počtu cifer čísla n. Přesto zůstává Millerův–Rabinův test díky své rychlosti nejpoužívanějším testem prvočíselnosti.

Literatura

[1] Agrawal, M., Kayal, N., Saxena, N.: PRIMES is in P. *Annals of Mathematics* **160**, 2 (2004), s. 781–793.

[2] Erdős, P., Surányi, J.: *Topics in the Theory of Numbers*. Springer-Verlag, Berlin, 2001.

[3] Herman, J., Kučera, R., Šimša, J.: *Equations and Inequalities: Elementary Problems and Theorems in Algebra and Number Theory*. CMS Books in Mathematics, Springer-Verlag, New York, 2003.

[4] Kučera, R.: *Algoritmy teorie čísel*, elektronické skriptum. `http://www.math.muni.cz/~kucera/texty/ATC2014.pdf`.

[5] Křížek, M., Somer, L., Šolcová, A.: *Kouzlo čísel – Od velkých objevů k aplikacím*. Academia, Praha, 2011.

[6] Lovász, L., Pelikán, J., Vesztergombi, K.: *Discrete mathematics: Elementary and Beyond*. Graduate Texts in Mathematics, Springer-Verlag, New York, 2003.

[7] Rivest, R., Shamir, A., Adleman, L.: A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. *Communications of the ACM* **21**, 2 (1978), s. 120–126.