UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

NA ČO SÚ DOBRÉ PRVOČÍSLA?

BAKALÁRSKA PRÁCA

2012 Kristína CIESAROVÁ

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

NA ČO SÚ DOBRÉ PRVOČÍSLA?

BAKALÁRSKA PRÁCA

Študijný program:	Ekonomická a finančná matematika
Študijný odbor:	1114 Aplikovaná matematika
Školiace pracovisko:	Katedra aplikovanej matematiky a štatistiky
Vedúci práce:	RNDr. Mária Trnovská, PhD.

Bratislava 2012 **Kristína CIESAROVÁ**

89792214

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Kristína Ciesarová
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský

Názov: Na čo sú dobré prvočísla?

Cieľ: Študovanie vlastností prvočísel a ich aplikácií v RSA kryptografii.

Vedúci: RNDr. Mária Trnovská, PhD.

Dátum zadania: 15.10.2011

Dátum schválenia: 27.10.2011

doc. RNDr. Margaréta Halická, CSc.
garant študijného programu

.....
študent

.....
vedúci

Poďakovanie

Týmto by som sa chcela poďakovať vedúcej mojej bakalárskej práce RNDr. Márii Trnovskej, PhD. za ochotu, odborné rady a pripomienky, ktoré mi pomohli pri písaní tejto práce. Ďakujem tiež svojej rodine a priateľovi za ich trpezlivosť a podporu.

Abstrakt

CIESAROVÁ, Kristína: Na čo sú dobré prvočísla? [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: RNDr. Mária Trnovská, PhD., Bratislava, 2012, 39s.

V tejto práci pozorujeme vlastnosti prvočísel a ich aplikáciu v kryptografii, konkrétne v RSA kryptografii, pričom cieľom našej práce je čo najjasnejšie vysvetliť danú problematiku. Uvádzame porovnanie kryptografie so symetrickým kľúčom a kryptografie s verejným kľúčom, kde verejný kľúč sa ukazuje ako v praxi efektívnejší. V rámci kryptografie s verejným kľúčom predstavuje naša práca RSA kryptosystém, v ktorom sú prvočísla veľmi účinným nástrojom na šifrovanie. Venujeme sa tiež testovaniu prvočíselnosti, čo je nevyhnutné pri zisťovaní základných parametrov v RSA kryptosystéme. Zaoberáme sa aj Euklidovým algoritmom, keďže je to algoritmus, ktorý je nevyhnutné poznať pri dešifrovaní správ šifrovaných pomocou RSA. Výsledkom našej práce je zhrnutie problematiky ilustrované aj na praktických príkladoch. V práci teda súhrnne opisujeme všetky aspekty RSA kryptosystému aj s algoritmami a vlastnosťami prvočísel, ktoré s ním súvisia.

Kľúčové slová: prvočísla, testovanie pseudoprvočíselnosti, RSA kryptosystém, Euklidov algoritmus

Abstract

CIESAROVÁ Kristína: Why are prime numbers useful? [Bachelor Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: RNDr. Mária Trnovská, PhD., Bratislava, 2012, 39p.

In this thesis we observe some prime numbers characteristics and their application in cryptography, namely in RSA cryptography, whereas our goal is to clearly explain given topic. We present a comparison between symmetric key cryptography and public key cryptography, where public key cryptography shows to be more effective in practice. Within public key cryptography our work introduces the RSA cryptosystem, where prime numbers are very efficacious encryption tool. We engage in primality testing as well, because this is inevitable to obtain fundamental parameters in RSA. Our thesis covers also Euclid's algorithm, since it is essential for decrypting messages in RSA. The outcome of our work is a summary of the topic illustrated by practical examples. We describe all the aspects of RSA cryptosystem with related algorithms and prime numbers characteristics.

Keywords: prime numbers, pseudoprimality testing, RSA cryptosystem, Euclid's algorithm

Obsah

Úvod

Už od začiatku písomnej formy komunikácie pochopili vojenskí velitelia a hlavy štátov, že je nutné nejakým spôsobom chrániť dôvernosť písomných informácií a mať nejaké prostriedky na odhaľovanie nepovolanej manipulácie so správami.

Toto sa dá vo všeobecnosti zabezpečiť dvoma spôsobmi a to pomocou steganografie alebo kryptografie. Úlohou steganografie je ukryť správu tak, aby si nezávislý pozorovateľ ani nevšimol, že komunikácia prebieha. Existuje mnoho rôznych metód ako pomocou steganografie ukryť správu (neviditeľný atrament, voskové tabuľky, správy ukryté v „neškodných“ písaných správach, ...), avšak tento spôsob má výraznú nevýhodu. Zachytenie tajnej správy už prakticky znamená jej prelomenie.

Tu sa však dostáva k slovu práve kryptografia, ktorej cieľom je zakódovanie tajnej správy tak, aby aj keď bude niekým zachytená, nebolo možné ju rozlúštiť, resp. aby to pre tretiu stranu bolo veľmi náročné. Preto môžeme sledovať dlhú históriu vývoja rôznych šifier od starovekého Egyptu, pokračujúc Cézarom a jeho šifrou, alebo neskôr počas druhej svetovej vojny, kedy sa dá povedať, že vznikla kryptografia na profesionálnej úrovni, až po dnešnú dobu. Práve rýchly rast a rozšírenie používania elektronického spracovania dát a elektronického podnikania cez internet v 21. storočí vyvolalo potrebu lepších metód ochrany počítačov a informácií, ktoré sú na nich uložené, spracovávané a prostredníctvom nich prenášané.

Moderná kryptografia sa venuje viacerým oblastiam, medzi inými kryptografii so symetrickým a asymetrickým kľúčom, kryptoanalýze (veda o tom, ako bez znalosti kľúča odvodiť správu zo šifrovaného textu (1)) a iným.

V tejto bakalárskej práci sa budeme zaoberať kryptografiou s verejným kľúčom, ktorú zaraďujeme do kryptografie s asymetrickým kľúčom, konkrétne tzv. RSA kryptosystémom. Práve tu sa ukázali prvočísla ako veľmi dôležité, pretože RSA kryptosystém na šifrovanie využíva rozklad veľkých celých čísel na súčin prvočísel, čo je výpočtovo náročný proces (2).

Cieľom tejto práce je oboznámiť čitateľa s niektorými vlastnosťami prvočísel, s fungovaním RSA kryptosystému a algoritmov, ktoré s ním úzko súvisia. Preto sme sa v práci venovali aj vybraným poznatkom z algebry a z teórie čísel, keďže úzko súvisia s procesom RSA šifrovania.

Na začiatku práce sa budeme venovať niektorým vlastnostiam prvočísel, ktoré sa aj v praxi využívajú (okrem iného aj v RSA kryptografii). V nasledujúcej kapitole v krátkosti uvedieme základné pojmy kryptografie aby sme čitateľa uviedli do danej problematiky. V ďalšej kapitole stručne vysvetlíme ako funguje kryptografia so symetrickým kľúčom a budeme sa podrobnejšie venovať kryptografii s verejným kľúčom. Poukážeme aj na širší význam tejto kryptografie – digitálny podpis. Potom predstavíme RSA kryptosystém a algoritmy, ktoré nám umožnia voľbu parametrov, vyskytujúcich sa v RSA. Podrobnejšie sa budeme venovať aj Euklidovmu algoritmu a jeho rozšíreniu, ktoré sa využíva pri šifrovaní pomocou RSA. Aby sme prehĺbili porozumenie čitateľa, ukážeme algoritmy aj na konkrétnych príkladoch.

Ohľadom našej témy existuje pomerne veľa odbornej literatúry, preto sme nemali problém s čerpaním informácií.

1. Vlastnosti prvočísel

V tejto kapitole predstavíme niektoré z vlastností prvočísel. Zdrojom nám bude najmä kniha (3), ale dá sa o nich dočítať viac aj v (4) alebo (5). Rozoberieme rozdelenie, asymptotické vlastnosti a niektoré ďalšie tvrdenia o prvočíslach. Začneme tým, že ukážeme, že prvočísel je nekonečne veľa. Predtým ako sa však k tomuto tvrdeniu dostaneme, musíme ešte predstaviť niekoľko pojmov.

Množinu všetkých prvočísel označíme ako P.

Uvedieme najprv základnú vetu, z ktorej bude naša teória vychádzať. Je to tzv. základná veta aritmetiky.

Veta 1 (Základná veta aritmetiky): *Každé prirodzené číslo väčšie ako 1 sa dá napísať ako súčin prvočísel.*

Dôkaz: Túto vetu dokážeme sporom. Predpokladajme, že existujú prirodzené čísla väčšie ako 1, ktoré sa nedajú zapísať ako súčin prvočísel. Vezmime najmenšie z nich a označme ho n. Toto číslo nemôže byť prvočíslo, keďže každé prvočíslo je súčinom jedného prvočísla, samého seba. Teda n musí byť zložené číslo, teda

$$n = a \cdot b,$$

kde a a b sú prirodzené čísla menšie ako n. Keďže n je najmenšie číslo, ktoré sa nedá zapísať ako súčin prvočísel, a a b sa takto dajú zapísať. Potom ale $n = a \cdot b$ tiež môže byť zapísané ako súčin prvočísel jednoducho tak, že skombinujeme rozklady čísel a a b, čo je spor. A teda všetky prirodzené čísla vieme napísať ako súčin prvočísel. ■

Tento súčin voláme **prvočíselný rozklad**.

Ďalej definujeme funkciu, ktorá bude udávať počet prvočísel menších ako nejaké prirodzené číslo.

Definícia 1: *Funkcia $\pi(x)$, kde x je kladné reálne číslo, označuje počet prvočísel menších alebo rovných ako x.*

Dostávame sa teda k tvrdeniu, že prvočísel je nekonečne veľa. Uvádzame aj dva rôzne dôkazy. Prvý je viac priamočiary, čiastočne sme ho čerpali z (4). Pochádza ale už od Euklida, ktorý ho vymyslel v roku 300 pred naším letopočtom. Druhý je však pre nás užitočnejší, čo uvidíme neskôr. Čerpali sme ho z knihy (3).

Veta 2: *Prvočísel je nekonečne veľa.*

Dôkaz: Predpokladajme, že prvočísel je konečne veľa. Vezmime potom prirodzené číslo

$$Q_n = p_1 p_2 \dots p_n + 1,$$

kde $n \geq 1$ a $p_1, p_2, \dots, p_n$ sú všetky prvočísla. Podľa základnej vety aritmetiky (**Veta 1**) sa dá každé prirodzené číslo napísať ako súčin prvočísel, teda:

$$Q_n = q_1 q_2 \dots q_m$$

(ak $m = 1$, tak Q_n je samo prvočíslo). Žiadne z prvočísel $p_1, p_2, \dots, p_n$ nemôže deliť Q_n, pretože $Q_n = 1 (\mathrm{mod}\, p_i)$ pre všetky i. Prvočísla $q_1, q_2, \dots, q_n$ musia byť potom iné prvočísla, ako $p_1, p_2, \dots, p_n$. Každá konečná množina prvočísel teda môže byť rozšírená na väčšiu konečnú množinu prvočísel; preto je prvočísel nekonečne veľa. ■

Alternatívny dôkaz: Tento dôkaz sme čerpali z knihy (3). Nech x je prirodzené číslo väčšie ako 1. Označme $\pi(x)$ počet všetkých prvočísel menších ako x (viď **Definícia 1**). Prvočísla menšie ako x označme $p_1, p_2, \dots, p_{\pi(x)}$. Každé číslo n, $1 \leq n \leq x$ vieme napísať v tvare $n = a^2 b$; $a, b \in \mathbb{N}$, kde $a \geq 1$ a b je číslo bez kvadratických deliteľov, teda nie je deliteľné druhou mocninou žiadneho prirodzeného čísla väčšieho ako 1. Keďže $n \leq x$, tak aj $a^2 \leq x$, resp. $a \leq \sqrt{x}$. Teda pre a nemáme viac ako $\lfloor \sqrt{x} \rfloor$ možností. Ďalej číslo b môžeme vyjadriť v tvare

$$b = p_1^{a_1} p_2^{a_2} \dots p_{\pi(x)}^{a_{\pi(x)}},$$

kde a_k $(k = 1, 2, \dots, \pi(x))$ je buď 0 alebo 1. Z toho vyplýva, že všetkých takých čísel b bez kvadratických deliteľov, pre ktoré platí $a^2 b \leq x$ pri nejakom $a \geq 1$, nie je viac ako $2^{\pi(x)}$. Keďže všetkých čísel n, $1 \leq n \leq x$ je práve x, dostávame nerovnosť

$$x \leq 2^{\pi(x)} \sqrt{x},$$

resp. $2^{\pi(x)} \geq \sqrt{x}$. Zlogaritmovaním dostávame

$$\pi(x) \geq \frac{\ln x}{2 \ln 2}.$$

Z toho vyplýva, že pre $x \to \infty$ sa $\lim_{x \to \infty} \pi(x) = +\infty$, čo znamená, že počet prvočísel sa blíži do nekonečna. ■

Tento alternatívny dôkaz je pre nás cennejší, pretože nám dáva istú informáciu o tom ako prvočísel pribúda v rastúcej postupnosti všetkých prirodzených čísel. Hovorí, že prvočísel menších ako x je „rádovo“ nie menej ako $\ln x$. Tým sa dostávame k jednému z najvýznamnejších výsledkov teórie čísel, tzv. prvočíselnej vete.

Veta 3 (Prvočíselná veta): $\lim_{n \to \infty} \frac{\pi(n)}{n / \ln n} = 1$.

Existuje viacero dôkazov tejto vety, ale všetky sú veľmi komplikované a presahujú rámec našej práce. Jeden z dôkazov je však možné nájsť v článku (6).

Bolo však ukázané, že existuje aj presnejší odhad hustoty prvočísel $li(n) = \int_2^n \frac{dt}{\ln t}$. Tiež to nebudeme dokazovať, ale v nasledujúcej tabuľke si môžeme pozrieť numerické porovnanie týchto dvoch odhadov. (4)

n	$\pi(n)$	$n/\ln n$	$li(n)$
10^3	168	144,8	178
10^5	9592	8685,9	9630
10^7	664579	620420,7	664918
10^9	50847534	48254942,4	50849235
10^{11}	4118054813	3948131663,7	4118156401
10^{13}	346065535898	334072678387,1	346065645810

Tabuľka 1

Keďže prvočísel je nekonečne veľa, môžeme ich zoradiť do nekonečnej postupnosti

$$p_1, p_2, \dots, p_k, \dots$$

To ale neznamená, že by sme poznali všetky prvočísla. Momentálne najväčšie známe prvočíslo je pravdepodobne $2^{43112609} - 1$, čo je 12978189-ciferné číslo. Nájdené bolo v roku 2008 a zatiaľ ešte väčšie prvočíslo nepoznáme. (7)

Ďalšou dôležitou vlastnosťou, ktorú spomenieme je asymptotická hustota množiny P. Najprv si ale tento pojem definujme.

Definícia 2: *Nech $A \subseteq \mathbb{N}$, $x \in \mathbb{N}$. Označme znakom $A(x)$ počet všetkých tých $a \in A$, pre ktoré $a \leq x$. Ak existuje*

$$\lim_{x\to\infty} \frac{A(x)}{x}$$

*nazývame toto číslo **asymptotickou hustotou množiny** A a označujeme ho $h(A)$.*

Môžeme teda sformulovať tvrdenie o asymptotickej hustote množiny prvočísel. Znie nasledovne.

Veta 4 (Asymptotická hustota prvočísel): *Platí $h(P) = 0$.*

Dôkaz: Nech $x \in \mathbb{N}$, $x \geq 2$. Potom $P(x) = \pi(x)$, a tak

$$\frac{P(x)}{x} = \frac{\pi(x)}{\left(\frac{x}{\ln x}\right)} \cdot \frac{1}{\ln x}$$

Prvý činiteľ na pravej strane rovnosti má podľa prvočíselnej vety (veta 2) pri $x \to \infty$ limitu 1. Druhý činiteľ má pri $x \to \infty$ limitu 0. Teda obe strany majú pri $x \to \infty$ limitu 0. ■

Ďalšie vlastnosti prvočísel sú uvedené v nasledovných tvrdeniach. Budeme ale potrebovať nasledovnú lemu.

Lema 1: *Ak* $n \in \mathbb{N}$*, tak platí*

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}.$$

Dôkaz: $(a-b)\sum_{k=0}^{n-1} a^k b^{n-1-k}$

$= \sum_{k=0}^{n-1} a^{k+1} b^{n-1-k} - \sum_{k=0}^{n-1} a^k b^{n-k}$

$= a^n + \sum_{k=1}^{n-1} a^k b^{n-k} - \sum_{k=0}^{n-1} a^k b^{n-k} - b^n$

$= a^n - b^n$. ■

Tvrdenie 1: *Ak* $2^n - 1$ *je prvočíslo, tak aj* n *je prvočíslo.*

Dôkaz: Predpokladajme, že tvrdenie neplatí, teda $2^n - 1$ je prvočíslo a n je zložené číslo. Nech teda $n = r \cdot s$, kde r, s sú prirodzené čísla a $r, s > 1$. Potom podľa **Lemy 1** máme

$$2^n - 1 = 2^{r \cdot s} - 1 = (2^s - 1)\left(2^{s(r-1)} + 2^{s(r-2)} + \cdots + 2^s + 1\right),$$

teda $(2^s - 1) | (2^n - 1)$ a zároveň $s > 1$, čo je spor s tým, že $2^n - 1$ je prvočíslo. Teda n je prvočíslo.

Poznámka 1: Prvočísla, ktoré sa dajú napísať v tvare $2^n - 1$ sa nazývajú **Mersennove prvočísla**. Môžeme si všimnúť, že aj zatiaľ najväčšie známe prvočíslo je Mersennovo prvočíslo.

Tvrdenie 2: *Ak* $2^n + 1$ *je prvočíslo, tak* $n = 2^m, m \in \mathbb{N}$.

Dôkaz: Ak n je prirodzené číslo, ale nie je mocninou 2, tak $n = r \cdot s$, kde $1 \leq r < n$ a $1 < s \leq n$, pričom s je nepárne. Podľa **Lemy 1** pre prirodzené číslo k platí

$$(a - b)|(a^k - b^k).$$

Substitúciou $a = 2^r$, $b = -1$ a $k = s$ a využitím toho, že s je nepárne dostávame

$$(2^r + 1)|(2^{r \cdot s} + 1),$$

a teda

$$(2^r + 1)|(2^n + 1).$$

Pretože $1 < 2^r + 1 < 2^n + 1$ dostávame, že $2^n + 1$ nie je prvočíslo, čo je spor. Teda $n = 2^m$. ■

Poznámka 2: Čísla v tvare $2^{2^n} + 1$ sa nazývajú **Frematove čísla**.

Posledná vlastnosť, ktorú spomenieme je obsahom tzv. malej Fermatovej vety.

Veta 5 (Malá Fermatova veta): *Ak* p *je prvočíslo, potom platí:*

$$\boldsymbol{a^{p-1} = 1 (\mathrm{mod}\ p)}, \qquad \forall \boldsymbol{a} \in \mathbb{Z}_{\boldsymbol{p}}^{+},$$

kde $\mathbb{Z}_p^+ = \{1, 2, \ldots, (p-1)\}$ *je množina nenulových zvyškov po delení čísla* p.

Aby sme dokázali túto vetu, budeme potrebovať nasledovnú lemu.

Lema 2: *Nech* $u, x, y \in \mathbb{Z}$ a $p \nmid u$. *Potom ak* $ux = uy \pmod p$, *tak* $x = y \pmod p$.

Dôkaz: Platí $ux = uy \pmod p$, čo znamená, že p delí $ux - uy = u(x - y)$. Keďže $p \nmid u$, tak p musí deliť $(x - y)$, teda $x = y \pmod p$. ■

Dôkaz Vety 5: Nech a je prirodzené číslo, ktoré nie je deliteľné číslom p (teda $a \in \mathbb{Z}_p^+$). Uvažujme postupnosť prvých $(p-1)$ násobkov čísla a

$$a, 2a, 3a, \ldots, (p-1)a. \qquad [1]$$

Podľa **Lemy 2** platí, že ak $ra = sa \pmod p$, tak $r = s \pmod p$. Teda tieto násobky musia byť rôzne a nenulové. Potom keď postupnosť [1] predelíme modulo p, výsledná postupnosť musí byť preusporiadaním postupnosti

$$1, 2, 3, \dots, (p-1). \qquad [2]$$

Teda ak medzi sebou vynásobíme čísla v jednotlivých postupnostiach, výsledky musia byť identické modulo p:

$$a \cdot 2a \cdot 3a \cdot \dots \cdot (p-1)a = \big(1 \cdot 2 \cdot 3 \cdot \dots \cdot (p-1)\big) \pmod p,$$

teda

$$a^{p-1}(p-1)! = (p-1)! \pmod p.$$

Nakoniec podľa **Lemy 2** môžeme vykrátiť $(p-1)!$ a dostávame

$$a^{p-1} = 1 \pmod p,$$

čo sme chceli dokázať. ■

Táto veta je veľmi užitočná v praxi, zakladajú sa na nej niektoré testy prvočíselnosti. Viac sa budeme testovaniu prvočísel venovať v kapitole 4.3.

2. Úvod do kryptografie

Cieľom tejto kapitoly je oboznámiť čitateľa so základnými pojmami, ktoré sa často vyskytujú v kryptografii a teda aj v našej práci. Preto je táto kapitola nevyhnutná na pochopenie zvyšných kapitol. Informácie v tejto kapitole budú čerpané najmä z publikácie (8).

Kryptografia je náuka o metódach posielania správ v šifrovanej podobe, ktorú môže odtajniť a prečítať len určený príjemca. Správa, ktorú posielame sa nazýva **prostý** (nešifrovaný) **text** a utajená správa sa nazýva **šifrovaný text**. Množinu všetkých prostých správ budeme označovať $\mathcal{P}$ a množinu všetkých šifrovaných správ označíme $\mathcal{C}$. Prostý aj šifrovaný text sú napísané v nejakej **abecede**, ktorá môže, ale nemusí byť pre obe rovnaká. Táto abeceda sa skladá z N znakov, čo môžu byť písmená, číslice, medzery, interpunkčné znamienka alebo iné znaky. Proces zmeny prostého textu na šifrovaný sa nazýva **šifrovanie** a opačný proces **dešifrovanie**.

Šifrovanie prebieha tak, že nezašifrovaná správa sa rozdelí na viacero častí, na ktoré sa aplikuje tzv. **šifrovacia transformácia** (ozn. $\boldsymbol{f}$). Šifrovacia transformácia je funkcia, ktorá zmení prostý text na šifrovaný. Inak povedané, je to funkcia z množiny $\mathcal{P}$ do množiny $\mathcal{C}$. Budeme predpokladať, že táto funkcia je bijekcia, teda prosté zobrazenie na množinu $\mathcal{C}$. Takto ku každej zašifrovanej správe existuje práve jedna prostá správa. **Dešifrovacia transformácia** je teda inverzná funkcia $\boldsymbol{f^{-1}}$, ktorá funguje spätne, teda zo šifrovaného textu dostane opäť prostý text. Schematicky môžeme toto zapísať diagramom

$$\mathcal{P} \xrightarrow{f} \mathcal{C} \xrightarrow{f^{-1}} \mathcal{P}.$$

Každý takýto systém šifrovacej a dešifrovacej transformácie nazývame **kryptosystém**.

3. Verejný kľúč

Už vieme, že kryptosystém pozostáva z nejakej šifrovacej transformácie f z množiny prostých správ $\mathcal{P}$ do množiny šifrovaných správ $\mathcal{C}$. V skutočnosti sa však pojem kryptosystém používa všeobecnejšie, a to na celú triedu nejakých šifrovacích transformácií, ktoré závisia od voľby istého počtu parametrov. Tieto parametre sa nazývajú **šifrovací kľúč** $\boldsymbol{K_E}$. V praxi môžeme predpokladať, že algoritmus šifrovania je verejne známy, čiže spôsob šifrovania nemôžeme utajiť. Avšak šifrovacie kľúče môžeme držať v tajnosti a prípadne by sa mali dať ľahko zmeniť.

Na dešifrovanie potrebujeme vedieť nielen algoritmus šifrovania, ale aj kľúč, teda potrebujeme zistiť dešifrovaciu transformáciu f^{-1}. Tento kľúč sa volá **dešifrovací kľúč** $\boldsymbol{K_D}$. V niektorých kryptosystémoch v podstate nie je potrebné bližšie špecifikovať dešifrovací kľúč, keď už poznáme šifrovací kľúč, pretože sa dá nejakým spôsobom zistiť (ak poznáme šifrovací algoritmus). To znamená, že ak je šifrovací kľúč prezradený tretej strane, tak je vlastne prezradený aj dešifrovací kľúč, a teda šifra už viac nie je bezpečná. Takéto kryptosystémy nazývame kryptosystémy so **symetrickým kľúčom** (niekedy sa nazývajú aj „klasické kryptosystémy“). V týchto kryptosystémoch ak už jeden vie ako šifrovať správu, tak ju vie aj dešifrovať. Preto je nutné držať v tajnosti oba kľúče a šifrovací kľúč poskytnúť len tomu, s kým chceme udržiavať zabezpečenú komunikáciu. Tu však vyvstáva viacero problémov, medzi inými tzv. **problém distribúcie kľúčov** (1), keďže jediným úplne dôveryhodným spôsobom doručenia šifrovacieho kľúča je osobné stretnutie, čo je však časovo náročné, hlavne ak sú osoby, ktoré chcú spolu komunikovať od seba veľmi vzdialené.

Tento problém kryptografov trápil po celé storočia, až kým neobjavili páni Diffie a Hellman v roku 1975 úplne nový typ kryptosystémov a to tzv. kryptosystém s **verejným kľúčom**. (8) Kryptosystém s verejným kľúčom má takú vlastnosť, že ten, kto správu šifruje pomocou šifrovacieho kľúča, nedokáže správu pomocou neho dešifrovať. Teda aj keď pozná šifrovací kľúč, nie je možné zistiť dešifrovací kľúč bez veľmi náročného a zdĺhavého výpočtu. Inými slovami, je ľahké spočítať šifrovaciu funkciu $f: \mathcal{P} \to \mathcal{C}$, keď poznáme

šifrovací kľúč, ale veľmi ťažké zistiť inverznú funkciu $f^{-1}: \mathcal{C} \to \mathcal{P}$. Takáto funkcia je teda prakticky neinvertovateľná bez doplnkovej informácie K_D. Funkcie s takouto vlastnosťou nazývame **jednosmerné** (trapdoor) **funkcie**.

Idea šifrovania pomocou verejného kľúča je znázornená na nasledovnom obrázku:

Obrázok 1: Kryptografia s verejným kľúčom (9)

Vďaka tomuto objavu by bola možná šifrovaná komunikácia na úplne inej úrovni ako dovtedy. Každý používateľ by mal dva kľúče – šifrovací kľúč, ktorý by mohol zverejniť každému, preto sa volá aj verejný kľúč a dešifrovací kľúč, ktorý by udržal v tajnosti, teda svoj súkromný kľúč. Takto by akúkoľvek správu zašifrovanú jeho šifrovacím kľúčom vedel dešifrovať len on sám.

Na nasledujúcom príklade (ktorého zadanie sme čerpali z (8)) si ukážeme o koľko je efektívnejší kryptosystém s verejným kľúčom oproti kryptosystému so symetrickým kľúčom ak sa pozeráme na množstvo potrebných kľúčov.

Príklad 1: Predpokladajme, že m účastníkov chce medzi sebou komunikovať pomocou klasického kryptosystému. Každý účastník trvá na tom, aby bol schopný komunikovať s každým ďalším účastníkom tak, aby ostatných $m-2$ nemohlo ich komunikáciu rozlúštiť. Koľko kľúčov $K = (K_E, K_D)$ musí byť vytvorených? Koľko kľúčov by potrebovali,

keby používali kryptosystém s verejným kľúčom? Koľko kľúčov je potrebných pre oba typy, ak $m = 1000$?

Riešenie: V klasickom kryptosystéme musí mať každá dvojica, ktorá chce spolu bezpečne komunikovať svoju vlastnú dvojicu kľúčov $K = (K_E, K_D)$. Teda počet kľúčov, ak je účastníkov m je $(m-1) + (m-2) + \cdots + 1 = \frac{m(m-1)}{2}$, resp. $\binom{m}{2}$. V kryptosystéme s verejným kľúčom má však každý účastník svoj vlastný kľúč, pomocou ktorého mu vie ktokoľvek zašifrovať správu tak, že ju vie rozlúštiť len príjemca sám. Teda počet kľúčov pre tento typ kryptosystémov je m.

Pre $m = 1000$ potom máme $\binom{1000}{2} = 499500$ kľúčov v klasickom kryptosystéme a len 1000 kľúčov, ak používame verejný kľúč.

Z príkladu vidíme, že z pohľadu počtu kľúčov potrebných na zabezpečenú komunikáciu je kryptosystém s verejným kľúčom menej náročný ako pri symetrickom kľúči. Verejný kľúč sa teda opäť ukazuje ako veľmi užitočný.

3.1. Digitálny podpis

V tejto podkapitole si predstavíme ďalšie využitie kryptografie s verejným kľúčom okrem samotného šifrovania správ. Zdroj týchto informácií je publikácia (8).

V písomnej komunikácii je jednou z najdôležitejších častí správy podpis. Podpis odosielateľa správy napísaný osobitým rukopisom, ktorý je ťažké nejakým spôsobom duplikovať umožní príjemcovi vedieť, že správa naozaj pochádza od osoby, ktorej meno je uvedené. Ak je správa mimoriadne dôležitá, je potrebné použiť aj iné spôsoby autentifikácie komunikácie. V minulosti sa na overenie vierohodnosti komunikácie používali rôzne prostriedky ako pečate, odtlačky a iné. V elektronickej komunikácii však nemôžeme použiť takúto fyzickú autentifikáciu, ani fyzický podpis, preto sa musíme spoliehať na iné metódy.

V kryptografii s verejným kľúčom existuje mimoriadne jednoduchá cesta na osobnú identifikáciu Nech teda A (Alica) a B (Bob) sú dvaja užívatelia systému. Ďalej nech f_A je šifrovacia transformácia, pomocou ktorej môže hocikto poslať správu Alici a f_B podobne pre Boba. Nech P je Alicin podpis (môže obsahovať nejaké identifikačné číslo, dátum kedy bola správa poslaná, atď.). Nestačí, keď Alica pošle Bobovi zašifrovanú správu $f_B(P)$, pretože to môže spraviť každý a teda nebolo by možné zistiť, či Alicin podpis nebol sfalšovaný. Namiesto toho vloží Alica na začiatok alebo na koniec správy $f_B f_A^{-1}(P)$. Keď potom Bob dešifruje celú správu pomocou f_B^{-1}, celá správa sa stane čitateľnou, okrem malej časti nezmyselného textu, čo je $f_A^{-1}(P)$. Keďže Bob vie, že správa má byť od Alici, tak aplikuje f_A (ktoré pozná, lebo Alicin šifrovací kľúč je verejný) a získa P. Keďže nikto okrem Alici nepozná f_A^{-1}, Bob vie, že správa je určite od Alici.

Vidíme teda, že kryptografia s verejným kľúčom by teoreticky mala bohaté využitie. Dlho však bolo problémom nájsť takú funkciu, ktorou by sa mohla dostať do praxe.

4. RSA

Kryptografovia dlho hľadali jednosmernú funkciu, ktorá by umožnila šifrovanie s verejným kľúčom, až kým v roku 1977 neprišli Rivest, Shamir a Adleman so svojou prevratnou myšlienkou. Bezpečnosť RSA kryptosystému je založená na výpočtovej náročnosti problému rozkladania celých čísel na súčin prvočísel. V dnešnej dobe patrí medzi najpoužívanejšie kryptosystémy s verejným kľúčom. (2)

Obrázok 2: Ron Rivest, Adi Shamir, Len Adleman (10)

Rivest, Shamir a Adleman (ktorých môžeme vidieť na obrázku 2) vytvorili špeciálnu jednosmernú funkciu, ktorú môže invertovať len ten, kto má prístup k dôverným informáciám, a to k hodnote dvoch čísel p a q. Funkcia je určená výberom p a q, ktoré po vynásobení dajú číslo N. Funkcia umožní odosielateľovi zašifrovať správu pre príjemcu a to tak, že odosielateľ pozná len N, zatiaľ čo príjemca je jedinou osobou, ktorá pozná p a q, a teda je jedinou osobou, ktorá pozná dešifrovací kľúč d. (1)

V nasledujúcich častiach našej práce sa budeme venovať tomuto spôsobu šifrovania, ako aj praktickým problémom, ktoré sa pri ňom vynárajú (problematika voľby parametrov p, q, výpočet d).

4.1. Matematika RSA

Fungovanie mechanizmu šifrovania a dešifrovania pomocou RSA si ukážeme na nasledovnom príklade. Čerpali sme z publikácie (1):

Alica[1] si najprv zvolí dve obrovské prvočísla p a q. Tieto prvočísla musia byť mimoriadne veľké, ale v našom príklade budeme pre jednoduchosť predpokladať, že $p = 11$ a $q = 17$. Tieto dve čísla uchová v tajnosti.

Následne Alica čísla vynásobí medzi sebou a dostane číslo N. V tomto prípade $N = 187$. Vypočíta $\varphi(N) = (p-1)\cdot(q-1) = N - p - q + 1$, teda $\varphi(187) = 10 \cdot 16 = 160$. Potom si zvolí číslo $e \in \mathbb{N}$ tak, aby e a $\varphi(N)$ nemali žiadneho spoločného deliteľa (t.j. $D\big(e;\varphi(N)\big) = 1$). Prečo musíme e voliť takto si vysvetlíme neskôr v podkapitole 5.2. V našom prípade $e = 7$.

Alica teraz zverejní čísla e a N v niečom, čo by sme mohli nazvať telefónnym zoznamom. Vzhľadom na to, že tieto dve čísla sú pre šifrovanie nevyhnutné, musia byť k dispozícii každému, kto chce zašifrovať nejakú správu pre Alicu. Tieto dve čísla spolu vytvárajú verejný kľúč. Číslo e môže byť súčasťou rôznych verejných kľúčov, každý kľúč ale musí mať rôznu hodnotu čísla N, ktoré závisia od voľby p a q.

Aby sme vôbec mohli šifrovať, musí byť správa najprv prevedená zo znakov (tu uvažujme znaky anglickej abecedy) do nejakého čísla P. To môžeme spraviť napríklad prevedením slov do ASCII binárnych číslic, ktoré už môžeme previesť na čísla v desiatkovej sústave. P potom môžeme zašifrovať tak, aby vytvorilo šifrovaný text C pomocou šifrovacej transformácie: $\boldsymbol{C = P^e (\mathbf{mod}\ N)}$.

Predstavme si teda, že Bob chce Alici poslať symbol bozku – písmeno X. Jeho reprezentácia v ASCII je 1011000, čo je v desiatkovej sústave číslo 88. Takže naše $P = 88$.

[1] Alica, Bob a Eva sú zaužívané mená archetypov v kryptografickej terminológii. Obvykle si Alica a Bob chcú poslať navzájom šifrovanú správu a Eva vystupuje ako tretia strana, ktorá sa snaží správu dešifrovať bez dešifrovacieho kľúča. (13)

Aby mohol Bob túto správu zašifrovať začne tým, že vyhľadá Alicin verejný kľúč a zistí, že $N = 187$ a $e = 7$. To mu poskytne potrebný šifrovací vzorec: $C = 88^7 (\text{mod } 187)$, teda $C = 11 \text{ (mod } 187)$. Bob teda môže Alici poslať šifrovaný text $C = 11$.

Keďže mocniny v modulárnej matematike sú v podstate jednosmerné funkcie, je veľmi ťažké postupovať naspäť od $C = 11$ a získať originálnu správu. Preto Eva nemôže správu rozšifrovať. Musela by mať dešifrovací kľúč $\boldsymbol{d = e^{-1} (\text{mod } \varphi(N))}$, čo je prevrátená hodnota čísla e modulo $\varphi(N)$. To však nedokáže, pretože nepozná p a q.

Alici sa to však podarí, pretože má zvláštnu informáciu – pozná hodnoty p a q. Vypočíta dešifrovací kľúč d, resp. jej súkromný kľúč. Máme teda $d = 7^{-1} (\text{mod } 160)$. Riešenie tejto rovnice vieme nájsť pomocou rozšíreného Euklidovho algoritmu, čo vysvetlíme v nasledujúcej kapitole. V našom prípade $d = 23$.

Aby Alica rozlúštila správu, použije dešifrovaciu transformáciu:

$$\boldsymbol{P = C^d \text{ (mod } N)}$$

V našom príklade: $P = 11^{23} (\text{mod } 187)$, teda $P = 88$, čo je písmeno X v ASCII.

4.2. Ako voliť parametre

V príklade v predošlej podkapitole mohla aj Eva pomerne ľahko zistiť utajené čísla p a q. Stačilo by, keby postupne skúšala aké prvočísla delia $N = 187$ bezo zvyšku. Pre takto veľké číslo by to netrvalo ani veľmi dlho, stačilo by jej vyskúšať len 5 prvočísel a zistila by, že $p = 11$ a $q = 17$. Preto treba tieto čísla voliť omnoho väčšie a to rádovo tak, aby N malo aspoň 309 cifier (1024 bitov)[2]. Dnes sa však používajú čísla až do veľkosti 2048 bitov, teda 617-ciferné číslo, aby sa zabezpečila maximálna ochrana šifrovanej správy. (11)

[2] 1024 bitov znamená, že číslo má 1024 cifier v binárnej číselnej sústave.

Rozložiť takéto číslo na súčin prvočísel je v praktickom časovom horizonte nemožné[3]. Avšak ako sa vyvíjajú technológie, tak sa počet postačujúcich miest stále zvyšuje.

Ďalšou podmienkou na voľbu týchto dvoch prvočísel je, že nemôžu byť veľmi blízko seba. Teda ich treba voliť tak, aby jedno bolo o niekoľko cifier dlhšie ako druhé. Takto bude mať $(p-1)$ a $(q-1)$ pomerne malého najväčšieho spoločného deliteľa a obe tieto čísla budú mať aspoň jedného veľkého prvočíselného deliteľa. Na nasledovnom príklade (zadanie opäť pochádza z publikácie (8)) si ukážeme, prečo je dôležité, aby neboli čísla p a q veľmi blízko pri sebe.

Príklad 2: Ukážeme, prečo je číslo 23360947609 obzvlášť zlá voľba na $N = p \cdot q$, vzhľadom na blízkosť p a q.

Riešenie: Keďže p a q sú príliš blízko, môže byť N ľahko rozložené pomocou tzv. Fermatovej faktorizácie nasledovne. Všimnime si, že ak $N = p \cdot q$ (nech $p > q$), tak $N = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2$. Ak sú teda p a q blízko pri sebe, tak $s = \frac{p-q}{2}$ je malé a $t = \frac{p+q}{2}$ je celé číslo len o niečo väčšie ako $\sqrt{N}$ s vlastnosťou, že $t^2 - N$ je úplný štvorec. Ak potom testujeme po sebe idúce čísla $t > \sqrt{N}$, tak rýchlo nájdeme také číslo, ktoré spĺňa $N = t^2 - s^2$, resp. také číslo t, pre ktoré je $s = \sqrt{(t^2 - N)}$ celé číslo. Odtiaľ potom vieme ľahko zistiť $p = t + s$ a $q = t - s$. V našom príklade teda skúšame celé čísla $t > \sqrt{23360947609} \cong 152842{,}89$. Začneme teda číslom 152843 a pokračujeme dovtedy, kým $\sqrt{(t^2 - N)}$ nebude celé číslo. Jednotlivé výpočty môžeme zapísať do tabuľky.

i	t_i	$s_i = \sqrt{(t_i^2 - N)}$
1	152843	187,1897
2	152844	583,7183
3	152845	804

Tabuľka 2

[3] V roku 1977 bol v časopise Scientific American predstavený RSA kryptosystém a spolu s ním bola zverejnená výzva na dešifrovanie správy, kde N bolo 129-miestne číslo. Jeho rozklad na dve prvočísla zisťoval 600-členný tím dobrovoľníkov z celého sveta a s výsledkom prišli až o 17 rokov, v roku 1994. (1)

Vidíme, že už t_3 nám dalo požadovaný výsledok. Dostali sme teda, že $t = 152845$ a $s = 804$. Odtiaľ máme $p = 152845 + 804 = 153649$ a $q = 152845 - 804 = 152041$. Vidíme teda, že naše číslo N je nevhodným verejným kľúčom, keďže na to, aby sme zistili jeho rozklad na prvočísla p a q nám stačilo vyskúšať iba tri čísla t.

Samozrejme, že ak niekto objaví metódu rozkladania čísel na prvočísla, ktorá funguje rýchlo za nejakých iných podmienok na p a q, tak sa budú používatelia musieť vyhýbať aj týmto podmienkam.

Tu sa však vynára otázka ako takéto veľké prvočísla nájsť. V nasledujúcej podkapitole si predstavíme ako sa takéto čísla v praxi hľadajú. Čerpať budeme najmä z (12).

4.3. Hľadanie veľkých prvočísel

V kapitole 1 sme si ukázali, že prvočísla nie sú až také zriedkavé. Preto je možné hľadať prvočísla tak, že testujeme náhodné prirodzené čísla príslušnej dĺžky až kým nenájdeme prvočíslo.

Proces náhodného výberu celého čísla n a testovania jeho prvočíselnosti môžeme brať ako Bernoulliho pokus. Z **Vety 3** máme, že pravdepodobnosť úspechu – teda pravdepodobnosť, že n je prvočíslo – je približne $1/\ln n$. Očakávaný počet pokusov kým nedostaneme prvočíslo dĺžky n je teda približne $\ln n$. To znamená, že ak hľadáme prvočíslo dĺžky 1024 bitov, potrebujeme vyskúšať približne $\ln 2^{1024} \cong 710$ náhodne zvolených 1024-bitových čísel. Toto číslo ešte môžeme zmenšiť na polovicu tak, že budeme vyberať iba nepárne čísla.

Posledné, čo nám ostáva, je zistiť, či je naše náhodne zvolené číslo prvočíslo alebo nie. Budeme predpokladať, že n má prvočíselný rozklad:

$$n = p_1^{e_1} \cdot p_2^{e_2} \cdot \ldots \cdot p_r^{e_r},$$

kde $r \geq 1$, $p_1, p_2 \dots p_r$ sú prvočíselné delitele n a $e_1, e_2 \dots e_r$ sú prirodzené čísla. Číslo n je prvočíslo práve vtedy, keď $r = 1$ a zároveň $e_1 = 1$.

Najjednoduchší spôsob ako zistiť, či je číslo n prvočíslo, je deliť ho postupne prvočíslami od 2 do $\lfloor \sqrt{n} \rfloor$. Ak ani jedno z nich nedelí n, tak n je prvočíslo. Táto metóda však funguje dobre len pre malé n, alebo vtedy, keď má n malého deliteľa. Inak je prakticky nepoužiteľná, pretože je enormne časovo náročná (jej náročnosť rastie exponenciálne s dĺžkou n). Síce nám navyše dáva informáciu o deliteľoch čísla n, táto informácia nás tu nezaujíma, keďže chceme iba vedieť, či je číslo prvočíslo, nie aké má delitele.

V nasledujúcej podkapitole si predstavíme spôsob testovania prvočíselnosti, ktorý „takmer funguje“ a je dobrý pre mnoho praktických aplikácií. Ide o testovanie tzv. pseudoprvočíselnosti.

4.3.1. Testovanie pseudoprvočíselnosti

Predtým ako si ukážeme fungovanie tejto metódy, musíme definovať, čo sú to pseudoprvočísla.

Definícia 3: *Hovoríme, že n je **pseudoprvočíslo o základe a**, ak n je zložené číslo a platí*

$$a^{n-1} = 1 (\mathrm{mod}\, n).$$

Aby sme ďalej mohli dobre definovať proces testovania pseudoprvočísel, musíme spomenúť ešte niekoľko pojmov:

- $\mathbb{Z}_n = \{0, 1, 2, \dots, (n-1)\}$ je množina všetkých zvyškov po delení čísla n,
- $\mathbb{Z}_n^+ = \{1, 2, \dots, (n-1)\}$ je množina všetkých nenulových zvyškov po delení čísla n,
- $\mathbb{Z}_n^*$ je množina takých zvyškov po delení čísla n, ktoré sú s n nesúdeliteľné.

Môžeme si všimnúť, že ak je číslo n prvočíslo, tak platí $\mathbb{Z}_n^+ = \mathbb{Z}_n^*$.

Testovanie pseudoprvočísel sa zakladá na princípe malej Fermatovej vety. Z Fermatovej vety (**Veta 5**) vyplýva, že ak n je prvočíslo, tak spĺňa rovnicu [3] pre každé a zo $\mathbb{Z}_n^+$. Teda

ak nájdeme aspoň jedno číslo $a \in \mathbb{Z}_n^+$, ktoré nespĺňa [3], tak n je s istotou zložené číslo. Je však prekvapivé, že opačné tvrdenie platí takmer vždy a preto je to vhodnou podmienkou na testovanie prvočíselnosti. Testujeme teda, či n spĺňa rovnicu [3] pre $a = 2$. Ak nie, tak n prehlásime na zložené číslo. Inak tvrdíme, že n je prvočíslo, aj keď de facto vieme iba to, že n je buď prvočíslo, alebo pseudoprvočíslo o základe 2. Algoritmus potom vyzerá nasledovne:

vstup: nepárne celé číslo $n > 2$
ak $2^{n-1} \neq 1 (\mathrm{mod}\, n)$
    n je zložené číslo (vieme s istotou)
inak n je prvočíslo (dúfame)

Tento postup robí chyby, ale iba jedného druhu. Keď povie o čísle n, že je zložené, tak má vždy pravdu. Ak povie o čísle, že je prvočíslo, robí chybu iba vtedy, keď je n pseudoprvočíslo o základe 2. Tieto čísla sú však dosť zriedkavé, napríklad existuje iba 22 čísel menších ako 10^4, pre ktoré by táto metóda zlyhala. Nebudeme to dokazovať, ale pravdepodobnosť, že tento algoritmus sa pomýli na náhodne vybratom čísle dĺžky β bitov sa blíži k nule s $\beta \to \infty$. (12) Teda ak sa snažíme nájsť veľké prvočíslo tak, že vyberáme náhodné čísla dovtedy, kým o jednom z nich tento algoritmus neprehlási, že je prvočíslo, takmer nikdy sa nepomýli. Ak však čísla nie sú vyberané náhodne, potrebujeme lepší prístup testovania prvočíselnosti.

Chyby ktoré vznikajú nemôžeme úplne odstrániť jednoducho tým, že rovnicu [3] budeme skúšať pre ďalší základ, napríklad pre $a = 3$, pretože existujú také zložené čísla, ktoré vyhovujú rovnici [3] pre všetky $a \in \mathbb{Z}_n^*$. Takéto čísla sa volajú **Carmichaelove čísla**. Carmichaelove čísla sú však extrémne zriedkavé menších ako 10^8 je ich iba 255 (prvé tri sú 561, 1105 a 1729). (4)

Aby sme sa úplne vyhli chybám, ktoré by záviseli od voľby čísla n existujú rôzne vylepšenia tejto metódy, napríklad Miller-Rabinov test prvočíselnosti. V tejto práci sa ním nebudeme viac zaoberať, ale dá sa o ňom viac dozvedieť napríklad v knihe (12).

4.4. Zhrnutie

Poďme si teraz zhrnúť ako funguje RSA. Každý, kto chce prijímať šifrované správy si vyberie dve veľké prvočísla p a q, ktoré môže nájsť pomocou spomenutých algoritmov, a nejaké číslo e, ktoré je nesúdeliteľné s $\varphi(N) = (p-1)\cdot(q-1)$. Potom vyráta $N = p \cdot q$ a prevrátenú hodnotu e modulo $\varphi(N)$: $\boldsymbol{d \overset{\text{def}}{=} e^{-1}(\mathbf{mod}\ \varphi(N))}$. Toto číslo d nemusí nutne existovať, jeho existencia úzko súvisí s tým, či $D(e, \varphi(N)) = 1$. Toto je presnejšie popísané v podkapitole 5.2. Následne zverejní $K_E = (N, e)$ a ponechá utajené $K_D = (\varphi(N), d)$. Šifrovacia transformácia je

$$\boldsymbol{f(P) \equiv P^e (\mathbf{mod}\ N)}$$

a dešifrovacia transformácia

$$\boldsymbol{f^{-1}(C) \equiv C^d (\mathbf{mod}\ N)}.$$

Tieto dve transformácie sú si navzájom inverzné vďaka voľbe čísla d. Malo by teda platiť, že

$$f^{-1}(f(P)) = P.$$

Vieme, že

$$f^{-1}(f(P)) = (P^e)^d (\text{mod } N) = P^{e\cdot d} (\text{mod } N).$$

Ďalej keďže d je prevrátená hodnota e modulo $\varphi(N)$, tak platí

$$d \cdot e = 1 (\text{mod } \varphi(N)) = 1.$$

Teda dostávame

$$f^{-1}(f(P)) = P^1 (\text{mod } N) = P.$$

Vidíme teda, že šifrovacia a dešifrovacia transformácia s takto zvolenými číslami e a d sú si navzájom inverzné.

Zhrňme si ešte informácie potrebné na šifrovanie do tabuľky:

Označenie	Popis
$\boldsymbol{p, q}$	veľké prvočísla
$\boldsymbol{N}$	$N = p \cdot q$
$\boldsymbol{\varphi(N)}$	$\varphi(N) = (p-1) \cdot (q-1)$
$\boldsymbol{e}$	e – nejaké číslo $\in \mathbb{N}$; $D\big(e, \varphi(N)\big) = 1$
$\boldsymbol{d}$	$d \coloneqq e^{-1}\big(\mathrm{mod}\ \varphi(N)\big)$
$\boldsymbol{f(P)}$	$f(P) \equiv P^e (\mathrm{mod}\ N)$
$\boldsymbol{f^{-1}(P)}$	$f^{-1}(C) \equiv C^d (\mathrm{mod}\ N)$

Tabuľka 3

5. Euklidov algoritmus

Euklidov algoritmus je algoritmus, ktorý nám umožňuje nájsť najväčšieho spoločného deliteľa dvoch čísel a a b, aj keď nepoznáme ich rozklad na prvočísla.

Euklidov algoritmus funguje nasledovne. Aby sme našli $D(a,b)$, kde $a \geq b$, najprv vydelíme číslo a (označíme r_0) číslom b (označíme r_1) so zvyškom, kvocient označíme q_1 a zvyšok r_2. Teda $r_0 = q_1 r_1 + r_2$. V ďalšom kroku bude číslo $b = r_1$ v úlohe čísla $a = r_0$ a zvyšok r_2 prevezme úlohu čísla b: $r_1 = q_2 r_2 + r_3$. Ďalej delíme r_2 číslom r_3: $r_2 = q_3 r_3 + r_4$. Takto pokračujeme až dovtedy, kým nebude zvyšok po delení rovný nule. Potom posledný nenulový zvyšok je hľadaný najväčší spoločný deliteľ čísel a a b. (8)

Zapísané schematicky:

Nech $r_0 = a$, $r_1 = b$, $k = 1$.

Pokiaľ $r_k > 0$, hľadáme r_{k+1}, q_k:

$$r_{k-1} = q_k r_k + r_{k+1},$$

$$0 \leq r_{k+1} < r_k.$$

Posledný nenulový zvyšok označíme r_n.

Na to aby sme dokázali, že r_n z Euklidovho algoritmu je najväčší spoločný deliteľ, čísel a a b budeme potrebovať nasledovnú lemu. Dôkaz sme čerpali z publikácie (4).

Lema 3: *Ak c a d sú celé čísla a $c = qd + r$, kde q a r sú celé čísla, tak $D(c,d) = D(d,r)$.*

Dôkaz: Nech celé číslo e delí aj c aj d. Potom e delí r, pretože $r = c - qd$. Ak teda $e|d$ a $e|r$, potom $e|c$, keďže $c = qd + r$. Vzhľadom na to, že spoločný deliteľ c a d je rovnaký ako spoločný deliteľ d a r, tak vidíme, že $D(c,d) = D(d,r)$. ■

Veta 6: $r_n = D(a,b)$. *Teda pomocou Euklidovho algoritmu dostaneme vždy najväčšieho spoločného deliteľa dvoch čísel.*

Dôkaz: Nech $a = r_0$ a $b = r_1$ sú kladné celé čísla, pričom $a \geq b$. Keď postupne aplikujeme delenie so zvyškom, zistíme, že:

$$r_0 = q_1 r_1 + r_2 \quad 0 \leq r_2 < r_1,$$

$$r_1 = q_2 r_2 + r_3 \quad 0 \leq r_3 < r_2,$$

$$\vdots$$

$$r_{n-2} = q_{n-1} r_{n-1} + r_n \quad 0 \leq r_n < r_{n-1},$$

$$r_{n-1} = q_n r_n.$$

Môžeme predpokladať, že napokon dostaneme zvyšok rovný nule, pretože postupnosť zvyškov $a = r_0 > r_1 > r_2 > \cdots \geq 0$ nemôže obsahovať viac ako a členov. Potom podľa **Lemy 3** vidíme, že $D(a,b) = D(r_0, r_1) = D(r_1, r_2) = \cdots = D(r_{n-1}, r_n) = D(r_n, 0) = r_n$. A teda platí, že $D(a,b) = r_n$, čo je posledný nenulový zvyšok. ■

Poďme si fungovanie Euklidovho algoritmu ukázať na príklade.

Príklad 3: Nájdite najväčšieho spoločného deliteľa čísel 1989 a 867.

Riešenie môžeme zapísať do tabuľky pre lepšiu zrozumiteľnosť.

k	q_{k-1}	$r_{k-2} = q_{k-1} r_{k-1} + r_k$
0		1989
1		867
2	2	$1989 = 2 \cdot 867 + 255$
3	3	$867 = 3 \cdot 255 + 102$
4	2	$255 = 2 \cdot 102 + 51$
5	2	$102 = 2 \cdot 51 + 0$

Tabuľka 4

Posledný nenulový zvyšok je 51, teda $D(1989, 867) = 51$.

V RSA sa však využíva tzv. rozšírený Euklidov algoritmus, o ktorom si povieme viac v nasledujúcej podkapitole, v ktorej budeme čerpať zo zdroja (12).

5.1. Rozšírený Euklidov algoritmus

Rozšírený Euklidov algoritmus je také rozšírenie Euklidovho algoritmu, ktoré okrem najväčšieho spoločného deliteľa dvoch čísel hľadá dvojicu čísel x a y, ktoré spĺňajú tzv. Bézoutovu identitu[4]: $ax + by = D(a, b)$. Rozšírený Euklidov algoritmus je obzvlášť užitočný, keď $D(a, b) = 1$ (teda a a b sú nesúdeliteľné čísla). Potom máme:

$$ax + by = 1.$$

To nám vraví, že $(ax - 1)$ je y-násobkom čísla b, teda že $b|(ax - 1)$. To znamená, že $ax = 1 \pmod b$, teda x je prevrátená hodnota a modulo b. Podobne môžeme postupovať s $(by - 1)$ a dostaneme, že y je prevrátená hodnota b modulo a. Keď sú teda a a b nesúdeliteľné, tak pomocou rozšíreného Euklidovho algoritmu vypočítame $x = a^{-1} \pmod b$ a $y = b^{-1} \pmod a$.

Tento algoritmus ráta výraz $r_k = ax_k + by_k$ pre zvyšok z každého kroku Euklidovho algoritmu. Z Euklidovho algoritmu máme $r_{k-2} = q_{k-1}r_{k-1} + r_k$, teda každý zvyšok sa dá vyjadriť pomocou predošlých dvoch zvyškov:

$$r_k = r_{k-2} - q_{k-1}r_{k-1}.$$

Po substitúcii r_{k-1} a r_{k-2} dostávame $r_k = (ax_{k-2} + by_{k-2}) - q_{k-1}(ax_{k-1} + by_{k-1})$, čo môžeme napísať ako:

$$r_k = a(x_{k-2} - q_{k-1}x_{k-1}) + b(y_{k-2} - q_{k-1}y_{k-1}).$$

Prvé dve hodnoty sú počiatočné argumenty a a b:

$$r_0 = a = a \cdot 1 + b \cdot 0$$

$$r_1 = b = a \cdot 0 + b \cdot 1.$$

[4] Táto identita je pomenovaná po francúzskom matematikovi Étiennovi Bézoutovi, ktorý dokázal, že koeficienty spĺňajúce Bézoutovu identitu existujú pre všetky nenulové celé čísla a, b.

Teda koeficienty sú na začiatku $x_0 = 1$, $y_0 = 0$, $x_1 = 0$ a $y_1 = 1$. Ostatné sú potom dané vzťahmi

$$x_k = x_{k-2} - q_{k-1}x_{k-1},$$

$$y_k = y_{k-2} - q_{k-1}x_{k-1}.$$

Výraz pre posledný nenulový zvyšok $r_n = a(x_{n-2} - q_{n-1}x_{n-1}) + b(y_{n-2} - q_{n-1}y_{n-1})$ nám dáva požadovaný výsledok, kde r_n je najväčším spoločným deliteľom čísel a a b.

Riešenie si ilustrujeme na nasledovnom jednoduchom príklade.

Príklad 4: Nech $a = 91$ a $b = 35$. Vypočítajte $D(a, b)$ a x,y, ktoré vyhovujú rovnici $ax + by = D(a, b)$.

Riešenie: Postup budeme zapisovať postupne do tabuľky. V prvom stĺpci máme číslo kroku, v druhom je kvocient, v treťom stĺpci je zvyšok po delení zapísaný v tvare $r_k = r_{k-2} - q_{k-1}r_{k-1}$. V štvrtom stĺpci je substitúcia v tvare $(ax_{k-2} + by_{k-2}) - q_{k-1}(ax_{k-1} + by_{k-1})$ a v poslednom je výstup rozšíreného Euklidovho algoritmu.

k	q_k	Zvyšok	Substitúcia	$r_k = ax_k + by_k$
0		91		$91 = 91 \cdot 1 + 35 \cdot 0$
1		35		$35 = 91 \cdot 0 + 35 \cdot 1$
2	2	$21 = 91 - 2 \cdot 35$	$91 \cdot 1 + 35 \cdot 0 - 2 \cdot (91 \cdot 0 + 35 \cdot 1)$	$21 = 91 \cdot 1 + 35 \cdot (-2)$
3	1	$14 = 35 - 1 \cdot 21$	$91 \cdot 0 + 35 \cdot 1 - 1 \cdot \big(91 \cdot 1 + 35 \cdot (-2)\big)$	$14 = 91 \cdot (-1) + 35 \cdot 3$
4	1	$7 = 21 - 1 \cdot 14$	$91 \cdot 1 + 35 \cdot (-2) - 1 \cdot (91 \cdot (-1) + 35 \cdot 3)$	$7 = 91 \cdot 2 + 35 \cdot (-5)$
5	2	0		koniec cyklu

Tabuľka 5

Teda dostávame: $D(91,35) = 7$, $x = 2$ a $y = -5$.

5.2. Využitie Euklidovho algoritmu v RSA

Aby sme v RSA kryptosystéme dostali dešifrovací kľúč, potrebujeme riešiť rovnicu $d = e^{-1}(\mathrm{mod}\ \varphi(N))$, resp. rovnicu

$$\boldsymbol{d \cdot e = 1(\mathrm{mod}\ \varphi(N))}. \qquad [4]$$

Keďže $d \cdot e$ po delení $\varphi(N)$ dáva zvyšok 1, tak $\varphi(N)|(d \cdot e - 1)$. To znamená, že $\varphi(N)$ je deliteľom čísla $(d \cdot e - 1)$, a teda platí, že $d \cdot e - 1 = q\varphi(N)$. Odtiaľ dostávame:

$$d \cdot e - q\varphi(N) = 1,$$

pričom e a $\varphi(N)$ poznáme. Môžeme si všimnúť, že presne takýto typ rovnice rieši rozšírený Euklidov algoritmus, pričom máme dané, že $D(e, \varphi(N)) = 1$. Môžeme teda aplikovať rozšírený Euklidov algoritmus a tak získame náš dešifrovací kľúč d.

Poďme sa teraz pozrieť, čo by sa stalo, ak $D(e, \varphi(N)) \neq 1$, teda tieto dve čísla by mali spoločného deliteľa $c \neq 1$. Platilo by potom:

$$d \cdot e - q\varphi(N) = c,$$

teda $(d \cdot e - c)$ je q-násobkom čísla $\varphi(N)$, resp. $\varphi(N)|(d \cdot e - c)$. To znamená, že

$$d \cdot e = c\ (\mathrm{mod}\ \varphi(N)). \qquad [5]$$

Číslo d je však definované ako prevrátená hodnota e modulo $\varphi(N)$, teda musí vyhovovať rovnici [4]. Keďže však $c \neq 1$, tak [4] $\neq$ [5], a dostávame sa k sporu, lebo $d \neq d$. Teda d by v takom prípade neexistovalo. Preto e a $\varphi(N)$ musia byť navzájom nesúdeliteľné.

Vráťme sa k nášmu príkladu s Alicou a Bobom z kapitoly 4. Alica potrebovala vyriešiť rovnicu

$$d = 7^{-1}(\mathrm{mod}\ 160).$$

Čiže $7d = 1(\mathrm{mod}\ 160)$, a teda $7d - 1 = 160q$, resp.

$$7d - 160q = 1.$$

Postup riešenia pomocou rozšíreného Euklidovho algoritmu zapíšeme opäť do tabuľky.

k	q_k	Zvyšok	Substitúcia	$r_k = \varphi(N)q_k + ed_k$
1		160		$160 = 160 \cdot 1 + 7 \cdot 0$
2		7		$7 = 160 \cdot 0 + 7 \cdot 1$
3	22	$6 = 160 - 22 \cdot 7$	$160 \cdot 1 + 7 \cdot 0 - 22 \cdot (160 \cdot 0 + 7 \cdot 1)$	$6 = 160 \cdot 1 + 7 \cdot (-22)$
4	1	$1 = 7 - 1 \cdot 6$	$160 \cdot 0 + 7 \cdot 1 - 1 \cdot \big(160 \cdot 1 + 7 \cdot (-22)\big)$	$1 = 160 \cdot (-1) + 7 \cdot 23$
5	6	0		koniec cyklu

Tabuľka 6

Dostávame, že $q = -1$ a $d = 23$, a teda Alicin dešifrovací kľúč je naozaj 23.

Takýmto spôsobom si teda môže každý používateľ RSA kryptosystému vypočítať svoj vlastný dešifrovací kľúč.

Záver

Cieľom tejto práce bolo čitateľovi predstaviť RSA kryptografiu a ukázať, ako sa v nej využívajú vlastnosti prvočísel. U čitateľa sme predpokladali základnú znalosť z teórie čísel, konkrétne vedomosti modulárnej aritmetiky. Na začiatku práce sme predstavili niekoľko dôležitých vlastností prvočísel, ktoré sa využívajú v RSA kryptografii a algoritmoch s ňou súvisiacich.

Z oblasti kryptografie sme nepredpokladali čitateľove znalosti, preto sme v kapitole 2 uviedli základné pojmy z tejto oblasti. Predstavili sme taktiež myšlienku kryptografie s verejným kľúčom a spomenutím jej aplikácie pri digitálnom podpise v podkapitole 3.1 sme zdôraznili jej praktický význam.

V ďalšej kapitole sme postupne vysvetlili ako funguje RSA šifrovanie, pričom sme najprv predstavili ideu, na ktorej sa tento kryptosystém zakladá a potom sme jednotlivé aspekty tejto problematiky vysvetľovali podrobne. V podkapitole 4.1 sme predstavili matematiku, ktorá sa v RSA využíva, avšak ešte bez hlbšieho náhľadu do danej problematiky. Cieľom tejto podkapitoly bolo čo najjasnejšie popísať ako toto šifrovanie funguje v praxi, preto sme sa vyhýbali zložitému vysvetľovaniu (najmä objasnenie existencie dešifrovacieho kľúča d). V nasledujúcich podkapitolách sme však už išli do hĺbky, venovali sme sa hľadaniu veľkých prvočísel p a q, ktoré sú základnými vstupmi v RSA kryptografii. V podkapitole 4.3.1 sme predstavili účinnú metódu na hľadanie prvočísel, ktorá síce nefunguje úplne vždy, ale pravdepodobnosť, že sa pomýli sa blíži k nule. Preto je v praxi použiteľná. Na koniec štvrtej kapitoly sme vložili zhrnutie, kde sme zosumarizovali všetky potrebné informácie, ktoré potrebuje vedieť používateľ RSA.

Kapitola o Euklidovom algoritme je zaradená na koniec práce, pretože sme v prvom rade chceli zabezpečiť dobrú čitateľnosť a prehľadnosť práce. Posledná kapitola je koncipovaná viac matematicky. Predstavili sme ako funguje Euklidov algoritmus a tzv. rozšírený Euklidov algoritmus, ktorý sa využíva na výpočet dešifrovacieho kľúča. V podkapitole 5.2 sme tiež rozobrali existenciu tohto čísla d (teda aké podmienky a prečo musia platiť, aby

toto číslo existovalo). Všetky algoritmy sme predstavili aj na príkladoch, kvôli lepšej názornosti a prehĺbeniu pochopenia čitateľa.

Táto práca je určená pre čitateľa so záujmom o kryptografiu. Pojmy sú však vysvetlené od základov, preto nie je nutné, aby mal čitateľ z oblasti kryptografie akékoľvek znalosti. Pre tých, ktorí sa kryptografii venujú môže byť táto práca užitočným zhrnutím. Práca môže byť však prínosom aj mimo oblasti kryptografie, pretože sú v nej predstavené algoritmy, ktoré majú aj oveľa širší význam (rozšírený Euklidov algoritmus 5.1, testovanie pseudoprvočíselnosti 4.3.1).

Naša práca bola prínosom aj pre samotného autora, keďže sa sám oboznámil s časťou kryptografie a jej históriou, spoznal matematické pozadie RSA šifrovania a sám si vyskúšal riešenie problémov pomocou spomenutých algoritmov.

Zoznam použitej literatúry

1. **Singh, S.** *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography.* New York : Doubleday, 1999.

2. **Mezenes, A., van Oorschot, P. a Vanstone, S.** *Handbook of Applied Cryptography.* s.l. : CRC Press, 1996. dostupné na internete: http://cacr.uwaterloo.ca/hac/.

3. **Šalát, T. a kol.** *Algebra a teoretická aritmetika II.* Bratislava : ALFA, 1986.

4. **Rosen, K.** *Elementary Number Theory and its Applications.* Reading, Massachusetts : Addison-Wesley, 1984.

5. **Koshy, T.** *Elementary Number Theory With Applications, 2nd edition.* New York : Elsevier, 2007.

6. **Newman, D.** *Simple analytic proof of the prime number theorem, str. 693-696.* Washington DC : American Mathematical Monthly, 1980. Zv. 87, dostupné na internete: http://www.jstor.org/stable/2321853.

7. **Ch., Caldwell.** *The Dictionary of Prime Number Trivia.* s.l. : CreateSpace, 2009.

8. **Koblitz, N.** *A Course in Number Theory and Cryptography.* New York : Springer-Verlag, 1988.

9. Microsoft MSDN. [Online] 30. 5 2012. http://i.msdn.microsoft.com/dynimg/IC155063.gif.

10. University of Southern California. [Online] 30. 5 2012. http://www.usc.edu/dept/molecular-science/RSA-2003.htm.

11. **Silverman, R.** RSA Laboratories. *Has the RSA algorithm been compromised as a result of Bernstein's Paper?* [Online] 8. Apríl 2002. http://www.rsa.com/rsalabs/node.asp?id=2007.

12. **Cormen, T. a kol.** *Introduction to Algorithms.* Massachusetts : MIT Press, 2001.

13. **Newton, D.** *Encyclopaedia of Cryptology.* s.l. : ABC-Clio, 1998.