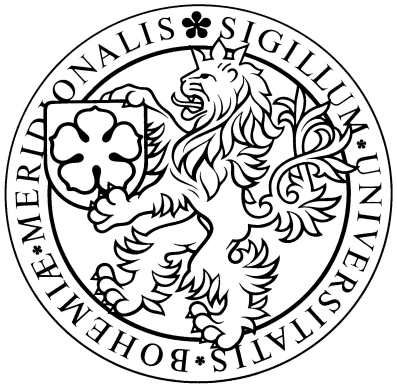


Jihočeská univerzita v Českých Budějovicích
Pedagogická fakulta

Matematika a šifrování

DIPLOMOVÁ PRÁCE

Pavla SIROTKOVÁ
České Budějovice, 22. dubna 2007

Prohlášení

Prohlašuji, že jsem na diplomové práci pracovala samostatně a že jsem uvedla veškerou literaturu, kterou jsem v této práci použila.

V Českých Budějovicích, dne 22. 4. 2007

.....
Pavla Sirotková

ANOTACE

Motivací, která vedla k vypracování této diplomové práce, byla snaha přiblížit matematické metody, které jsou používány k ochraně dat. V první části je uveden přehled šifrovacích metod od starověku po současnost. Dále jsou shrnuty matematické poznatky, které v dalších kapitolách slouží jednak k šifrování pomocí algoritmu RSA, ale také k různým druhům dešifrování.

ANNOTATION

The motivation which led to working out of this thesis was an effort to show mathematical methods which are used to data protection. In the first part of this thesis is introduced the outline of encrypting methods from ancient era until the present time. Further on there are summarized the mathematical findings which function not only to encryption by the help of algorithm RSA, but also to different kinds of decryption.

Poděkování

Děkuji panu doc. RNDr. Pavlu Tlustému, CSc., vedoucímu mé diplomové práce, za podnětné rady a připomínky, kterými mi pomohl při jejím vypracování.

Dále bych chtěla poděkovat své rodině a přátelům za podporu při studiu.

.....
Pavla Sirotková

Obsah

Kapitola 1

Úvod

Každý z nás si pod slovem šifra vybaví něco jiného. Někdo si vzpomene na dětské hry z letních táborů, jiný zase na populární knihy, kde hlavní hrdinové luští šifry a kódy, za kterými se skrývají největší tajemství všech dob.

Slovo šifra pravděpodobně pochází z arabského „as-sifr“, což je název pro číslici nula. Šifrování provází téměř celou lidskou historii. Způsoby, jak skrytě sdělit zprávu, vždy zajímaly nejen válečné stratégy, ale i diplomaty či obchodníky. Postupně se proto jednotlivé šifry komplikovaly a zdokonalovaly až k dnešní asymetrické kryptografii. Současně s uměním utajování zpráv (*kryptografií*) se rozvíjelo i umění nedovoleného dešifrování zpráv (*kryptoanalýza*). Stručný historický přehled nejznámějších a nejpoužívanějších šifrovacích metod od starověkého Řecka po druhou světovou válku je uveden ve druhé kapitole.

V druhé polovině dvacátého století byly mechanické šifrovací stroje nahrazeny počítači. Jejich rozšíření umožnilo použít moderní matematické teorie, které učinily dnešní šifrovací systémy bezpečnými. Přehledu matematických definic a vět, které se používají v moderních šifrovacích algoritmech je věnována třetí kapitola. V té jsou dále zavedeny základní pojmy z teorie pravděpo-

dobnosti a statistické matematiky a popsán problém faktorizace[1] složených čísel.

Jedním z hlavních cílů této diplomové práce je seznámení s algoritmem RSA. Tento algoritmus byl vybrán ze všech moderních šifrovacích systémů, které se používají v civilní kryptologii. Algoritmus RSA je základem dnešní kryptologie a byl první, který používal k zašifrování zprávy jiný klíč než k jejímu dešifrování. Dnes je jedním z prvků zabezpečení našich e-mailů, platebních karet, elektronických podpisů či různých přístupových hesel.

S rozvojem počítačů odpadl požadavek jednoduchých a pro člověka lehce pochopitelných algoritmů, čímž se minimalizovala naděje na průlom šifer. I přesto však pokračuje hledání nových ještě dokonalejších šifer. V tomto směru vkládají kryptologové naděje do vývoje kvantového počítače.

Jedním z nástrojů, které slouží k dešifrování zpráv, je frekvenční analýza. V páté kapitole blíže vysvětluji tuto dešifrovací metodu, která využívá matematickou statistiku a teorii pravděpodobnosti.

K diplomové práci jsou přiloženy některé jednoduché a zajímavé způsoby šifrování.

[1] Faktorizace je matematický problém rozložení čísla na součin menších čísel, v nejčastější podobě pak rozklad celého čísla na součin prvočísel.

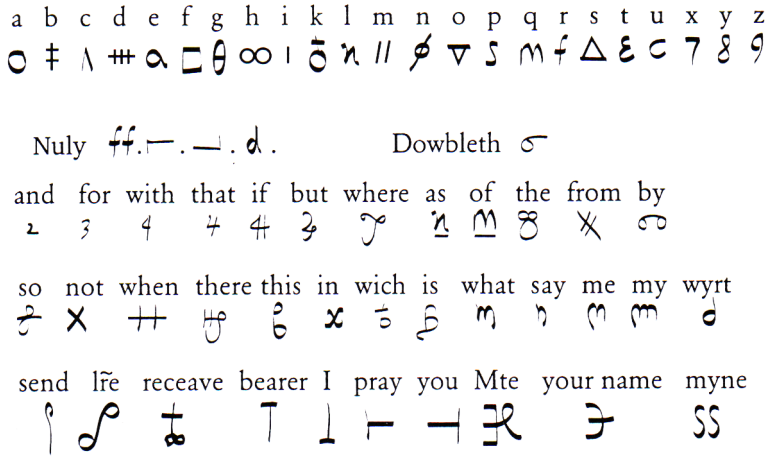
Kapitola 2

Historie šifrování

Po tisíce let spoléhali lidé, zvláště pak králové a generálové, na komunikační systémy, které jim umožňovaly předávání tajných informací. Problémem takovéto komunikace bylo riziko vyzrazení, které mohlo mít až tragické následky. Jedním z příkladů je šifra Marie Stuartovny, jejíž rozluštění znamenalo Mariinu popravu. Skotská královna Marie Stuartovna psala dopisy pomocí šifrové abecedy, ve které byl každému písmenu přiřazen určitý symbol (viz obrázek 2.1). Zašifrované texty byly pro zmatení doplněny i takzvanými nulami, symboly bez významu. Odpůrci Marie, tehdejší kryptoanalytikové, začali dešifrování identifikací nul, poté se jim podařilo uhádnout klíčová slova, až nakonec rozluštili celé dopisy.

Od doby této panovnice se mnohé změnilo, neustále však pokračuje boj mezi tvůrci a luštiteli kódů. Tvůrci šifer usilují o stále dokonalejší utajení komunikací, zatímco jejich luštitelé vyvíjejí ještě rafinovanější techniky útoku. Síla šifer je obrovská. Dokumentuje to i německý šifrovací přístroj Enigma, jehož rozluštění významně přispělo k vítězství spojenců ve 2. světové válce.

20. a 21. století přineslo největší změnu v kryptografii, kterou se stala převaha matematiků, a to jak v roli luštitelů, tak tvůrců šifer.

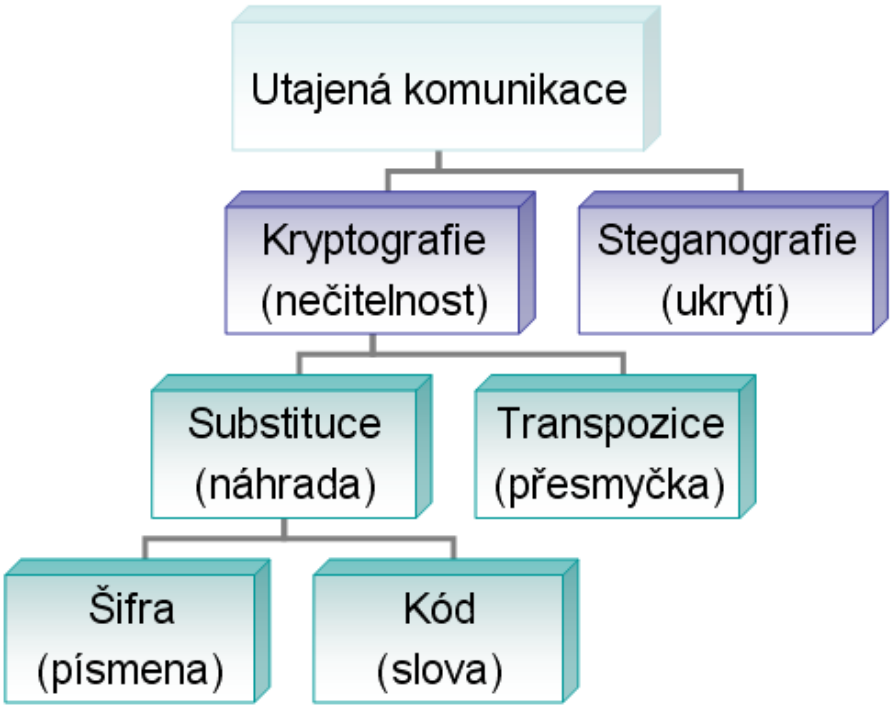


Obrázek 2.1: Šifra Marie Stuartovny [9]

V dnešní době provozují jednotlivé státy speciální šifrovací pracoviště, která jsou zodpovědná za bezpečnost komunikací, a kde se uvádějí do praxe nejlepší možné šifry. Kryptologové stáli také u vzniku moderních počítačů. Dnes přináší revoluci do kryptografie kvantová fyzika. Na obrázku 2.2 vidíte přehled šifrovacích systémů.

2.1 První tajná komunikace

Tajná sdělení existovala od počátku naší civilizace. Už otec historie Hérodotos z Halikárnassu (484 př. n. l. - 425 př. n. l.) ve své Historii píše o tajném písmu, které bylo použito v řecko-perských válkách. Do Řecka poslal tajnou zprávu Řek Demaratus, který žil ve vyhnanství v perském městě Susy. Demaratus seškrábal vosk ze dvou voskových psacích destiček, napsal zprávu přímo na dřevo a pak zprávu znovu zakryl voskem. Tato první tajná komunikace spočívala v prostém ukrytí zprávy. Jiný panovník této doby našel bezpečnější způsob ukrytí zprávy. Oholil hlavu svého posla, napsal zprávu na kůži jeho lebky a počkal, až poslovi znovu na-



Obrázek 2.2: Různé druhy utajování informací

rostou vlasy. Posel pak mohl cestovat bez potíží až do cíle své cesty, kde mu oholili hlavu a zprávu si přečetli.

V různých civilizacích existovalo mnoho druhů utajené komunikace pomocí ukrytí zprávy. Například staří Číňané psali zprávy na jemné hedvábí, které zmačkali do malé kuličky, zalili voskem a dali spolknout poslovi. Již z 1. století našeho letopočtu pochází neviditelný inkoust Plinia Staršího. Jako inkoust používal mléko pryšce, které je po zaschnutí zcela průhledné. Když se pak mléko lehce zahřeje, zhnědne. Tato tajná komunikace založená na ukrývání celých správ se nazývá *steganografie*, podle řeckých slov steganos (schovaný) a graphein (psát).

2.2 Počátky kryptografie

Souběžně se steganografií se začala rozvíjet i *kryptografie*. Kryptografie netají existenci zprávy, ale pomocí šifrování tají její význam. Zpráva se pozmění pravidly, která si mezi sebou předem dohodnou odesílatel a příjemce zprávy. Taková zpráva je pro nepřítele nečitelná. Tyto dvě techniky bylo výhodné kombinovat.[1]

2.3 Caesarova šifra

Důležité místo v historii kryptografie patří i římskému vojevůdci Juliu Caesarovi (100 př. n. l. - 44 př. n. l.). Caesar vymyslel první systém šifrování, který spočíval v tom, že každé písmeno zprávy nahradil písmenem, které bylo v abecedě o tři místa dále.

[1] Kombinovat steganografii s kryptografií dovedli k dokonalosti němečtí agenti za 2. světové války. Používali tzv. mikrotečku. Fotografickou cestou zmenšili rozsáhlý text do velikosti tečky o průměru menším než milimetr a tu pak umístili jako normální tečku za větou do nevinného dopisu.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Tabulka 2.1: Caesarova šifra

K	O	S	T	K	Y	J	S	O	U	V	R	Z	E	N	Y
N	R	V	W	N	B	M	V	R	X	Y	U	C	H	Q	B

Tabulka 2.2: Kostky jsou vrženy!

Pokud bychom ho chtěli napodobit (s použitím mezinárodní latinské abecedy), nahradili bychom písmena takto: písmeno *A* písmenem *D*, písmeno *B* písmenem *E*, a tak dále, písmeno *W* písmenem *Z*, potom písmeno *X* písmenem *A*, písmeno *Y* písmenem *B* a písmeno *Z* písmenem *C*. Tak jak je to vidět v tabulce 2.1. V druhé tabulce 2.2 je ukázka zašifrování zprávy touto šifrou.

Posun abecedy o několik pozic obecně nazýváme *monoalfabetickou substituční šifrou.*

2.4 Utajení zprávy podle Vigenère

Objev této šifry navázal na práce L. B. Albertiho, J. Trithemia a G. Porty. Ten, kdo ji však dopracoval do její konečné podoby byl francouzský diplomat Blaise de Vigenère (obrázek 2.3). Roku 1526 předvedl Viegenère novou šifru, která používá 26 různě posunutých abeced. Jednotlivé posunuté abecedy zapíšeme do tabulky, jak je vidět v tabulce 2.3.

Z tabulky je zřejmé, že první řádek odpovídá monoalfabetické substituční šifře s posunem 1. Kdybychom používali jen tento řádek, byl by výsledek pro kryptoanalytiky příliš jednoduchý. Trik je v tom, že každé písmeno můžeme zašifrovat kteroukoliv z 26 různě posunutých šifrovacích abeced. Systém, podle kterého bu-

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
1	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
2	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
3	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
4	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
5	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
6	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
7	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
8	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
9	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
10	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
11	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
12	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
13	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
14	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
15	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
16	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
17	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
18	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
19	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
20	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
21	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
22	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
23	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
24	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
25	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
26	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Tabulka 2.3: Vigenèrův čtverec

Obrázek 2.3: Blaise de Vigenère [24]

Klíčové slovo	t	u	z	k	a	t	u	z	k	a	t	u	z	k	a
Otevřený text	o	b	j	e	v	n	o	v	e	s	i	f	r	y	v
Zašifrovaný text	h	v	i	o	v	g	i	u	o	s	b	z	q	i	v

Tabulka 2.4: Šifrování zprávy podle Vigenèra

deme vybírat řádky pro šifrování jednotlivých písmen nám určí klíčové slovo.

Příklad zprávy zašifrované tímto způsobem můžeme vidět v tabulce 2.4. Do první řádky tabulky napíšeme klíčové slovo, v tomto případě *tužka*. V druhém řádku tabulky je napsán text, který chceme zašifrovat. Písmena textu budeme šifrovat postupně. První písmeno, v našem případě *O* šifrujeme pomocí abecedy (z Vigenèrova čtverce), která začíná písmenem *T*. Ve Vigenèrově čtverci najdeme v prvním řádku písmeno *O* a v prvním sloupci písmeno *T*. Na průsečíku těchto dvou linií leží písmeno *H*. Nyní se vrátíme k naší tabulce 2.4 a do posledního řádku zapíšeme první písmeno zašifrovaného textu *H*. Tímto způsobem šifrujeme celý otevřený text.

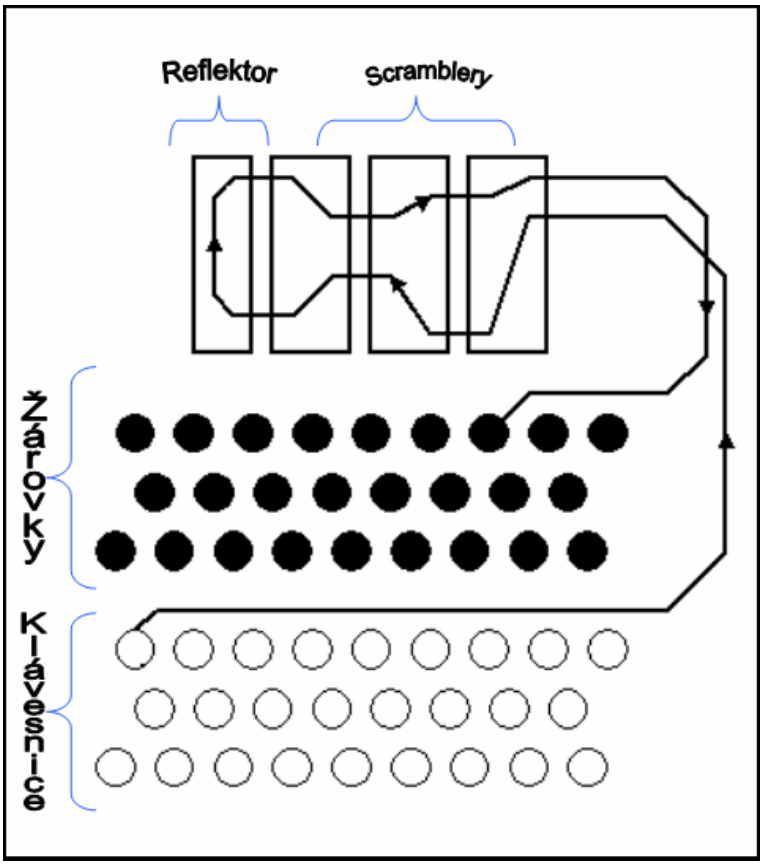
Obrázek 2.4: Šifrovací stroj Enigma [18]

2.5 Enigma

Během první světové války se používaly šifrovací systémy založené na „tužce a papíru“. Druhá světová válka si však žádala složitější systém utajování, a tak se ke slovu dostal šifrovací stroj Enigma. Její německý vynálezce Arthur Scherbius získal patent na Enigmu již v roce 1918. Dlouho se snažil prodat svůj vynález do obchodních kruhů, ale její vysoká cena obchodníky odradila. Až roku 1926 začala Enigma šifrovat komunikaci německé armády.

Na obrázku 2.4 vidíte na první pohled kufříkový psací stroj. Na druhý pohled pak složitý šifrovací stroj, hvězdu 2. světové války, Enigmu. Tento šifrovací aparát se skládá z klávesnice, propojovací desky, tří scramblerů, ovládacího kolečka, reflektoru a signální desky.

Zjednodušené schéma Enigmy na obrázku 2.5 obsahuje všechny hlavní části Enigmy. Obsluha stroje stiskne jedno písmeno na klávesnici, například *M*. Uvažujme abecedu s 26 písmeny. Signál z klá-



Obrázek 2.5: Zjednodušená verze Enigmy

vesnice pokračuje k prvnímu scrambleru, což je tlustý gumový kotouč protkaný dráty. Scrambler je zapojen tak, že každému písmenu přiřadí jiné písmeno abecedy a funguje vlastně jako monoalfabetická substituční šifra. Naše písmeno M se zašifruje na A. Trik je v tom, že scrambler se po zašifrování každého písmene pootočí o jednu dvacetišestinu otáčky. Pokud bychom tedy znova napsali písmeno M zašifrovalo by se na například na F.

Otáčející se scramblery jsou v Enigmě tři, zapojené za sebou. Druhý scrambler se otočí, až když první dokončí celou jednu otáčku. Třetí se otočí, až když druhý scrambler dokončí celou otáčku. Velmi důležitou součástí tohoto stroje je reflektor. Je to obdoba scrambleru, avšak neotáčí se. Reflektor přijme signál, který už prošel přes tři scramblery a vrátí ho přes scramblery jinou cestou zpátky. Přidáním reflektoru se mnohonásobně zvýší obtížnost dešifrování, ale reflektor bez scramblerů by opět šifroval pouze monoalfabetickou substituční šifrou. Obtížnost dále můžeme zvýšit tím, že před začátkem každého šifrování změníme zapojení mezi klávesnicí a prvním scramblerem. Nyní už signál prošel scramblery, reflektorem a opět scramblery. Výsledné zašifrované písmeno se rozsvítí na „žárovkové“ signální klávesnici.

A kolik je vlastně možných nastavení Enigmy, neboli možných šifrovacích klíčů? Každý ze 3 scramblerů může být nastaven do jedné z 26 výchozích pozic. Nastavení je tedy:

$$26 \cdot 26 \cdot 26 = 17\,576$$

Scramblery lze vyndat a vrátit je v jiném uspořádání. Možností uspořádání tří scramlerů je:

$$3! = 6$$

Ve skutečné Enigmě je možné pomocí kabelů zaměnit šest písmen za jiná. Děje se tak mezi klávesnicí a prvním scramblerem, na tzv.

propojovací desce. Prohodit lze šest párů písmen z celkového počtu 26 písmen. Počet způsobů prohození vypočítáme takto:

$$\binom{26}{2} \cdot \binom{24}{2} \cdot \binom{22}{2} \cdot \binom{20}{2} \cdot \binom{18}{2} \cdot \binom{16}{2} =$$

$$= \frac{26 \cdot 25 \cdot 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \cdot 15}{2^6} =$$

$$= 72\,282\,089\,880\,000 \doteq 72 \cdot 10^{12}$$

Počet všech možných klíčů Enigmy dostaneme jako součin výše uvedených čísel:

$$17\,576 \cdot 6 \cdot 72 \cdot 10^{12} \doteq 76 \cdot 10^{18}$$

K dešifrování zprávy použije příjemce svou vlastní Enigmu. Šifrování a dešifrování jsou vzájemně opačné postupy. Příjemce tedy nastaví scramblery a všechny součásti tak, jako je má nastaveny odesílatel.[2] Poté napíše zašifrovaný text na klávesnici Enigmy a na signální desce se mu objevuje původní otevřený text, písmeno po písmenu.

K utajení důležitých zpráv používaly různé státy různé šifry a šifrovací stroje. Němci šifrovali pomocí Enigmy, Američané používali šifrovací stoj M-2098(Hagelin), Japonci šifrovací stroj Purpur. Československí zpravodajci používali systém TTS (transpozice + transpozice + substituce).

[2] Němečtí vojáci měnili nastavení jednou denně, nastavení se říkalo denní klíč. Denní klíče byly domlouvány a distribuovány celé armádě vždy na začátku měsíce.

Kapitola 3

Matematické teorie užívané v šifrování

Od druhé poloviny 20. století se šifrování definitivně dostalo do rukou matematiků. K vysvětlení a pochopení dalších šifer budeme potřebovat některé poznatky z algebry, teorie pravděpodobnosti a matematické statistiky.

3.1 Teorie čísel

Teorie čísel je odvětví matematiky zabývající se vlastnostmi přirozených a celých čísel. Základy moderní teorie čísel položil C. F. Gauss (1777-1855). Obrázek 3.1.

Obrázek 3.1: C. F. Gauss [17]

Definice 3.1 *Přirozená čísla jsou čísla* $1, 2, 3, \ldots$, *množinu přirozených čísel značíme* $\mathbb{N}$.

$$\mathbb{N} = 1, 2, 3, ...$$

Definice 3.2 *Celá čísla jsou všechna přirozená čísla, čísla k nim opačná a nula. Množinu celých čísel značíme* $\mathbb{Z}$.

$$\mathbb{Z} = 0, \pm 1, \pm 2, \pm 3, ...$$

Definice 3.3 *Nechť* $a, b \in \mathbb{Z}$. *Řekneme, že* a ***dělí*** b, *značíme* $a \mid b$, *jestliže existuje* $k \in \mathbb{Z}$ *tak, že* $b = a \cdot k$. *V opačném případě říkáme, že* a ***nedělí*** b, $a \nmid b$. *Jestliže* $(a \mid b) \ \wedge (b \mid a)$, *říkáme, že čísla* a, b *jsou* ***asociována*** *a píšeme* $a \parallel b$, *pokud* a, b *nejsou asociována, píšeme* $a \nparallel b$. *Pokud* $a \mid b$, $a \neq \pm 1$, $a \neq b$, *pak řekneme, že* a *je* ***vlastním dělitelem*** *čísla* b. *Naopak, když bude platit* $a \parallel 1$ *nebo* $a \parallel b$, *řekneme, že* a *je* ***nevlastním dělitelem*** b.

Definice 3.4 *Celé číslo* $b > 1$ *se nazývá* ***prvočíslo***, *jestliže nemá vlastní dělitele. V opačném případě se nazývá číslo* ***složené***.

Následují definice a věty o dělení a násobení čísel.

Definice 3.5 *Nechť* $a_1, a_2, ..., a_n$ *jsou celá čísla,* $t \in \mathbb{Z}$ *nazveme* ***společným dělitelem*** *čísel* $a_1, a_2, ..., a_n$, *jestliže* $t \mid a_i$, $\forall i = 1, 2, ..., n$. *Číslo* $d \in \mathbb{Z}$ *nazveme* ***největším společným dělitelem*** *čísel* $a_1, a_2, ..., a_n$, *značíme* $D(a_1, ..., a_n)$, *jestliže* d *je společným dělitelem* $a_1, a_2, ..., a_n$ *a jestliže* t *je libovolný společný dělitel čísel* $a_1, a_2, ..., a_n$, *pak platí* $t \mid d$. *Pokud* $D(a_1, ..., a_n) = 1$, *řekneme, že čísla* $a_1, a_2, ..., a_n$ *jsou* ***nesoudělná***.

Věta 3.1 (O dělení se zbytkem) *Nechť* $a, \ b \neq 0$ *jsou celá čísla. Pak existují* $q, r \in \mathbb{Z}$ *tak, že*

$$a = bq + r, \ 0 < r < |b| \tag{1.1}$$

Věta 3.2 (Euklidův algoritmus) *Nechť $a, b \neq 0$ jsou dvě celá čísla. Pak existují celá čísla $q_0, q_1, q_2, ..., q_n$, $r_1, r_2, r_3, ..., r_n$ tak, že $r_n = D(a, b)$ a platí:*

$$a = bq_0 + r_1, 0 < r_1 < |b|$$
$$b = r_1q_1 + r_2, 0 < r_2 < r_1$$
$$\vdots$$
$$r_{i-1} = r_iq_i + r_{i+1}, 0 < r_{i+1} < r_i$$
$$\vdots$$
$$r_{n-2} = r_{n-1}q_{n-1} + r_n, 0 < r_n < r_{n-1}$$
$$r_{n-1} = r_nq_n$$

Definice 3.6 *Pro $0 \neq n \in \mathbb{N}$ označme $\varphi(n)$ počet všech čísel $k, 1 \leq k \leq n$ takových, že $D(k, n) = 1$. Funkce φ se nazývá **Eulerova funkce**.*

Definice 3.7 *Nechť $a_1, a_2, a_3, ..., a_n$ jsou celá čísla, $T \in \mathbb{Z}$ nazveme **společným násobkem** čísel $a_1, a_2, a_3, ..., a_n$, jestliže $a_i \mid T$, $\forall i = 1, 2, ..., n$. Číslo M nazveme **nejmenším společným násobkem** čísel $a_1, a_2, a_3, ..., a_n$, značíme $M = n(a_1, a_2, a_3, ..., a_n)$, jestliže M je společným násobkem čísel $a_1, a_2, a_3, ..., a_n$ a pro libovolný společný násobek T čísel $a_1, a_2, a_3, ..., a_n$ platí $M \mid T$.*

V šifrovacích algoritmech budeme potřebovat zbytek po dělení tzv. *modulo*, což je početní operace související s operací celočíselného dělení.

Definice 3.8 *Nechť $m \in \mathbb{Z}$. Jestliže $m \mid (a - b)$, říkáme, že **a je kongruentní s b podle modulu m** a píšeme symbolicky $a \equiv b \bmod m$. Pokud $m \nmid (a - b)$, říkáme, že a není kongruentní s b podle modulu m, a píšeme $a \not\equiv b \bmod m$.*

Věta 3.3 *Nechť $a, b, c \in \mathbb{Z}, m \in \mathbb{N}$. Pak platí:*

1.
$$a \equiv a \bmod m \tag{1.2}$$

2.
$$(a \equiv b \bmod m) \Rightarrow (b \equiv a \bmod m) \tag{1.3}$$

3.
$$[(a \equiv b \bmod m) \wedge (b \equiv c \bmod m)] \Rightarrow (a \equiv c \bmod m) \tag{1.4}$$

Věta 3.4 *Nechť $a, b, c, d \in \mathbb{Z}, m \in \mathbb{N}$.*
Pokud $a \equiv b \bmod m$ a $c \equiv d \bmod m$, pak platí:

1.
$$(a + c) \equiv (b + d) \bmod m \tag{1.5}$$

2.
$$(a - c) \equiv (b - d) \bmod m \tag{1.6}$$

3.
$$ac \equiv bd \bmod m \tag{1.7}$$

Věta 3.5 *Nechť $a, b, c \in \mathbb{Z}, m \in \mathbb{N}$. Nechť $D(c, m) = 1$, pak platí:*

$$ac \equiv bc \bmod m \Rightarrow a \equiv b \bmod m \tag{1.8}$$

Věta 3.6 *Nechť $a, b \in \mathbb{Z}, m, k \in \mathbb{N}$.*

$$a \equiv b \bmod m \Rightarrow a^k \equiv b^k \bmod m. \tag{1.9}$$

Věta 3.7 (Malá Fermatova) *Nechť $x \in \mathbb{Z}$, $p > 1$ je prvočíslo, $D(x, p) = 1$. Pak:*
$$x^{p-1} \equiv 1 \bmod p. \tag{1.10}$$

Věta 3.8 (Eulerova) *Nechť $x \in \mathbb{N}, p > 1, D(x,p) = 1$. Pak:*

$$x^{\varphi(p)} \equiv 1 \bmod p. \tag{1.11}$$

Věta 3.9 (Wilsonova) *Nechť p je prvočíslo. Pak:*

$$(p-1)! + 1 \equiv 0 \bmod p \tag{1.12}$$

Věta 3.10 (Čínská věta o zbytcích) *Nechť $N_1, N_2, ..., N_k \in \mathbb{N}$ jsou navzájem nesoudělná čísla, $N_i \geq 2$ pro $i = 1, 2 \ldots k$, potom každá soustava rovnic:*

$$x \equiv a_1 \bmod N_1 \tag{1.13}$$

$$x \equiv a_2 \bmod N_2 \tag{1.14}$$

$$x \equiv a_3 \bmod N_3 \tag{1.15}$$

má řešení x a toto řešení je určeno jednoznačně v modulo $N = N_1 \cdot N_2 \cdot \ldots \cdot N_k$.

Věta 3.11 (Gaussův algoritmus) *Řešení x rovnic z Čínské věty o zbytcích lze spočítat takto:*

$$x = \sum_{i=1}^{k} a_i n_i M_i \bmod N \tag{1.16}$$

kde $n_i = \frac{N}{N_i}$ a $M_i = {n_i}^{-1} \bmod N_i$.

3.2 Prvočísla

Prvočísla jsou základní kameny struktury přirozených čísel. Dosud největší známé prvočíslo $2^{232\,582\,657} - 1$ má 9 808 358 cifer. Mnoho matematiků, ale i laiků, se věnuje hledání provočísel. Na toto téma existuje nespočet internetových odkazů a různých programů.(Například programy na ověřování zda dané číslo je či není prvočíslo.)

Věta 3.12 (Základní věta aritmetiky) *Každé přirozené číslo lze rozložit na součin konečného počtu kladných prvočísel, a to až na pořadí jednoznačně.*

Věta 3.13 *Číslo $p \in \mathbb{N}$ je prvočíslo právě tehdy, když nemá žádného dělitele d takového, že $1 < d \leq \sqrt{p}$.*

Důkaz: Sporem. 1 *Nechť $n = n_1 \cdot n_2$. Předpokládejme, že $n_1 > \sqrt{n}$ a $n_2 > \sqrt{n}$. Pak ale*

$$n_1 \cdot n_2 > \sqrt{n} \cdot \sqrt{n} = n$$

*a to je **spor**. Tím je věta dokázána.*

Věta 3.14 *Prvočísel je nekonečně mnoho.*

Důkaz: Sporem. 2 *Nechť existuje jen konečně mnoho prvočísel. Označme je takto:*

$$p_1, p_2, \cdots, p_n.$$

Vytvoříme nové číslo x

$$x = p_1 \cdot p_2 \cdots p_n + 1$$

Číslo x je číslo složené.
Jelikož při dělení vždy dostaneme zbytek 1, číslo x není dělitelné žádným z prvočísel $p_1, \cdots, p_n$. Z toho vyplývá, že musí být dělitelné nějakým jiným prvočíslem.
*To ale znamená, že množina prvočísel z počátku důkazu nebyla úplná, což je **spor** dokazující platnost věty.*

Už řecký matematik Euklides (asi 365 př. n. l. – 300 př. n. l.) se zabýval prvočísly a věděl, že je jich nekonečně mnoho.

Další řecký matematik Eratosthenes z Kyrény (276 př. n. l. – 194 př. n. l.) používal k vyhledávání prvočísel voskové tabulky s napsanými přirozenými čísly většími než 1. Číslo 2 vynechal a jehlou

**	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
(**2**)	2	3		5		7		9		11		13		15		17		19		21		23		25	
(**3**)	2	3		5		7				11		13				17		19				23		25	
(**5**)	2	3		5		7				11		13				17		19				23			

Tabulka 3.1: Eratosthenovo síto

vypálil násobky 2. Pak stejně postupoval se 3 atd. Na konci výpočtu tabulka připomíná síto – tzv. *Eratosthenovo síto.*

V tabulce 3.1 hledáme posloupnost prvočísel od 2 do 26. V prvním řádku je zapsána celá tato posloupnost čísel. V druhém řádku jsme vynechali všechny násobky čísla 2, ale číslo 2 jsme nechali. První zbylé číslo po číslu 2 je číslo 3. V dalším řádku tedy vynecháme všechny násobky čísla 3, kromě čísla 3 samotného. Dále pokračujeme podle stejného pravidla a v posledním řádku vyškrtáme násobky čísla 5.
Obecně takto pokračujeme, dokud nenarazíme na prvočíslo p takové, že platí:

$$p^2 > n \tag{2.1}$$

Kde n je poslední číslo posloupnosti. Pak všechna zbylá čísla v posloupnosti jsou právě všechna hledaná prvočísla. V našem případě:

$$7^2 > 32$$

V posledním řádku tabulky 3.1 jsou zapsaná všechna prvočísla od 2 do 26.

Další nekonečné prvočíselné síto, jehož autorem je S. P. Sundaram, vidíme v tabulce 3.2. *Sundaramovo síto* je tvořeno nekonečným počtem aritmetických posloupností. Všimněme si, že každý člen první řady je zároveň prvním členem jedné z dalších řad. Diferencemi v jednotlivých řadách jsou postupně všechna lichá čísla počínaje trojkou.

Platí, že je-li číslo n obsaženo v této tabulce, pak je číslo $2n+1$ složené. Naopak není-li číslo n obsaženo v této tabulce, je číslo

4	7	10	13	16	19	22	$\cdots$
7	12	17	22	27	32	37	$\cdots$
10	17	24	31	38	45	52	$\cdots$
13	22	31	40	49	58	67	$\cdots$
16	27	38	49	60	71	82	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Tabulka 3.2: Sundaramovo síto

4	7	10	13	16	19	22	$\cdots$	$\Leftrightarrow$	$(1+3k) \wedge k = 1, 2 \ldots$
7	12	17	22	27	32	37	$\cdots$	$\Leftrightarrow$	$(2+5k) \wedge k = 1, 2 \ldots$
10	17	24	31	38	45	52	$\cdots$	$\Leftrightarrow$	$(3+7k) \wedge k = 1, 2 \ldots$
13	22	31	40	49	58	67	$\cdots$	$\Leftrightarrow$	$(4+9k) \wedge k = 1, 2 \ldots$
16	27	38	49	60	71	82	$\cdots$	$\Leftrightarrow$	$(5+11k) \wedge k = 1, 2 \ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		

Tabulka 3.3: Vyjádření aritmetických posloupností

$2n + 1$ prvočíslo.

Například číslo 7 se nachází v tabulce, platí tedy $2 \cdot 7 + 1$, tj. číslo 15 je číslo složené.

Číslo 5 se v tabulce nenachází, platí tedy $2 \cdot 5 + 1$, tj. číslo 11 je prvočíslo.

Důkaz 1 *Nyní dokážeme, že jde opravdu o nekonečné prvočíselné síto. Všechna čísla z tabulky i mimo ni vkládáme do vzorce* $2n+1$*, což je obecně předpis pro lichá čísla. Z toho vyplývá, že hledáme prvočísla pouze mezi lichými čísly. (Jediné sudé prvočíslo je číslo 2, ostatní sudá prvočísla jsou jím dělitelná.)*

Vraťme se k Sundaramově tabulce a zkusme obecně vyjádřit jednotlivé aritmetické posloupnosti. Vznikne nová upravená Sundaramova tabulka 3.3.

Nyní jednotlivá čísla z tabulky a obecné předpisy aritmetických posloupností vložíme do vzorce $2n + 1$*. Výsledek vidíme v další tabulce 3.4.*

9	15	21	27	$\cdots$	$\Leftrightarrow$	$2 \cdot (1 + 3k) + 1 = 3 + 6k = 3 \cdot (1 + 2k), k = 1, 2 \ldots$
15	25	35	45	$\cdots$	$\Leftrightarrow$	$2 \cdot (2 + 5k) + 1 = 5 + 10k = 5 \cdot (1 + 2k), k = 1, 2 \ldots$
21	35	49	63	$\cdots$	$\Leftrightarrow$	$2 \cdot (3 + 7k) + 1 = 7 + 14k = 7 \cdot (1 + 2k), k = 1, 2 \ldots$
27	45	63	81	$\cdots$	$\Leftrightarrow$	$2 \cdot (4 + 9k) + 1 = 9 + 18k = 9 \cdot (1 + 2k), k = 1, 2 \ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$		

Tabulka 3.4: Složená čísla vzniklá ze Sundaramova síta

V prvním řádku tabulky se nachází všechny liché násobky čísla 3, v druhém všechny liché násobky čísla 5, ve třetím všechny liché násobky čísla 7 atd. Uvědomme si, jaká čísla mohou vzniknout pomocí operace násobení.

$$\begin{array}{|c|}\hline liché \cdot liché \rightsquigarrow liché \\ liché \cdot sudé \rightsquigarrow sudé \\ sudé \cdot sudé \rightsquigarrow sudé \\ \hline\end{array}$$

Aplikací vzorce $2n + 1$ *na Sundaramovo síto jsme našli všechna čísla, která vznikla násobením dvou lichých čísel. To znamená, že jsme dostali všechna lichá složená čísla. Když si nyní vezmeme všechna lichá přirozená čísla, podobně jako jsme to dělali u Eratosthenova síta, a vyškrtáme všechna lichá složená čísla z našeho upraveného Sundaramova síta, pak zbydou pouze prvočísla. cbd.*

3	*5*	*7*	*9*	*11*	*13*	*15*	*17*	*19*	*21*	*23*	*25*	*27*	*29*	*31*
3	*5*	*7*		*11*	*13*		*17*	*19*		*23*			*29*	*31*

Mnoho matematiků se pokoušelo najít formuli, která by popisovala všechna prvočísla. V následujícím přehledu si uvedeme některé pokusy o nalezení takovéto funkce na množině $\mathbb{N}$, která by dávala pro libovolné $n \in \mathbb{N}$ prvočíslo.

Kvadratický polynom Například polynom tohoto tvaru:

$$P(n) = n^2 + n + 41 \tag{2.2}$$

dává prvočíslo pro všechna přirozená čísla menší než 40. Poslední prvočíslo, které pomocí toho polynomu získáme je 1601.

$$P(39) = 39^2 + 39 + 41 = 1\,601$$

Obecně žádná polynomická formule, která pro každé $n \in \mathbb{N}$ dává prvočíslo, nemůže existovat.

Fermatova prvočísla Tyto prvočísla studoval francouzský matematik Pierre de Fermat (1601 – 1665). Jeho domněnku, že všechna tato čísla jsou prvočísla, vyvrátil švýcarský matematik Leonard Euler (1707 - 1783). Fermatova prvočísla jsou čísla tvaru:

$$F(n) = 2^{2^n} + 1, \tag{2.3}$$

kde $n \in \mathbb{N}$.

$$F(4) = 2^{16} + 1 = 65\,537$$

Pro $n = 4$ dostaneme prvočíslo 65 537, avšak pro $n = 5$ dostaneme složené číslo $4\,294\,967\,297 = 641 \cdot 6\,700\,417$.

Mersennova prvočísla V současnosti je známo celkem 44 Mersennových prvočísel. Jedná se o prvočísla ve tvaru:

$$M(n) = 2^p - 1, \tag{2.4}$$

kde p je prvočíslo. Takto vypočítáme první Mersennovo prvočíslo:

$$M(2) = 2^2 \ - \ 1 = 3$$

Už víme, že největší známé prvočíslo je číslo $2^{232\,582\,657} - 1$, které má 9 808 358 cifer. Toto číslo je zároveň největším Mersennovým prvočíslem.[1]

[1] 44. Mersennovo prvočíslo bylo objeveno 4. září 2006 v rámci projektu GIMPS, The Great Internet Mersenne Prime Search. Do tohoto projektu se zapojují dobrovolníci z celého světa, kteří si do svého počítače nainstalují jednoduchý program na hledání těchto prvočísel.

3.3 Pravděpodobnost a matematická statistika

Pravděpodobnost náhodného jevu je číslo, které je mírou očekávatelnosti výskytu jevu. Pravděpodobnost události se obecně označuje reálným číslem od 0 do 1. Událost, která nemůže nastat, má pravděpodobnost 0, naopak jistá událost má pravděpodobnost 1.

Pravděpodobností jevu A pak nazveme číslo $P(A) = \frac{m}{n}$, kde n je počet všech výsledků náhodného pokusu a m je počet výsledků příznivých jevu A.

Definice 3.9 *Nechť A a B jsou náhodné jevy a platí $P(B) > 0$. **Podmíněnou pravděpodobnost** jevu A za podmínky, že nastal jev B, definujeme vztahem:*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{3.1}$$

Definice 3.10 *Nechť $A_1, A_2, ..., A_n$ jsou náhodné jevy, pravděpodobnost jejich průniku je:*

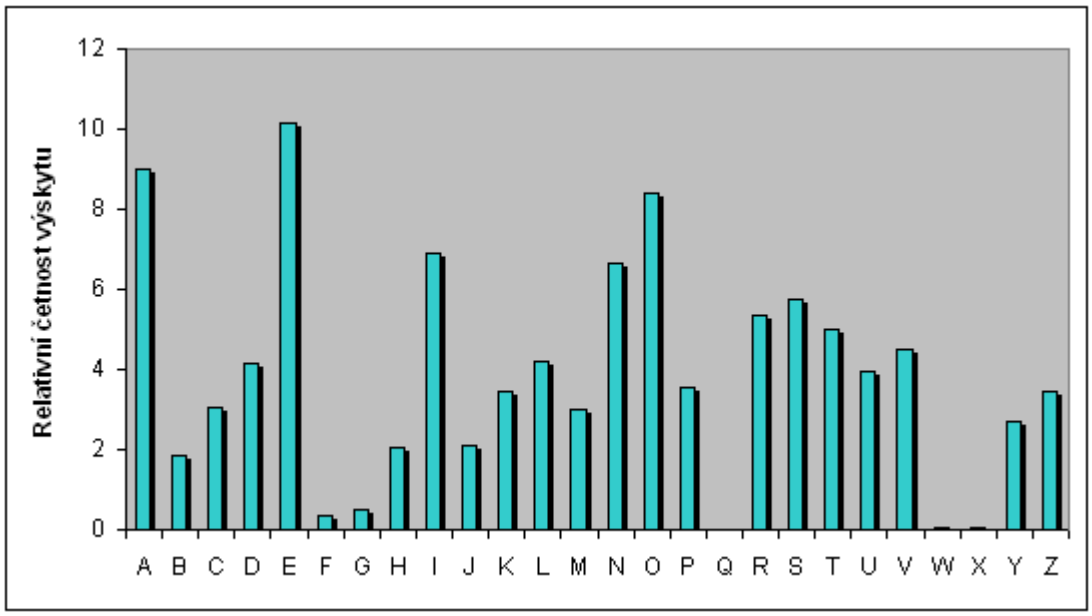
$$P\left(\cap_{i=1}^{n} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P\left(A_n| \cap_{i=1}^{n-1} A_i\right) \tag{3.2}$$

Definice 3.11 *Druhy četností:*

1. *Provedeme-li náhodný výběr o rozsahu n, mohou se některé hodnoty opakovat vícekrát. Počet výskytů n_i hodnoty x_i označujeme jako **absolutní četnost** hodnoty x_i.*

2. *Je-li n_i absolutní četnost hodnoty x_i a n je rozsah náhodného výběru, potom:*

   - $\frac{n_i}{n}$ *nazýváme **relativní četnost** hodnoty x_i.*
   - $100 \cdot \frac{n_i}{n}$ *nazýváme **procentní relativní četnost** hodnoty x_i.*



Obrázek 3.2: Histogram rozložení četností písmen v obecném českém textu

3. ***Komulativní absolutní četností*** *nazýváme součet všech četností, které nepřevyšují hodnotu* x_i.
   $\sum_{k=1}^{i} n_k$

4. ***Komulativní relativní četností*** *nazýváme součet všech relativních četností, které nepřevyšují hodnotu* x_i.
   $\sum_{k=1}^{i} \frac{n_k}{n}$

Definice 3.12 *Typy znázornění absolutních či relativních četností:*

Histogram *rozložení absolutních (relativních) četností je vlastně sloupcový graf, sestavíme ho tak, že na osu x vyneseme intervaly stejné délky a sestrojíme k nim pravoúhelníky, jejichž druhé rozměry jsou příslušné četnosti.*

Úsečkový diagram *sestrojíme podobně jako histogram. Nad středy jednotlivých intervalů sestrojíme úsečky kolmé k ose x o délce příslušné četnosti.*

Polygon *rozložení četností dostaneme, jestliže koncové body úsečkového diagramu spojíme úsečkami a vytvoříme tak lomenou čáru.*

Ogivní křivku *dostaneme, sestrojíme-li polygon komulativních relativních četností.*

Věta 3.15 ***Zákon velkých čísel*** *vyjadřuje představu, že při velkém počtu nezávislých pokusů je možné téměř jistě očekávat, že relativní četnost se bude blížit teoretické hodnotě pravděpodobnosti.*

Věta 3.16 (Bernoulliova věta) *n_i je počet nastoupení náhodného jevu A v n nezávislých opakováních pokusu a p je pravděpodobnost, že tento jev nastane v jednom pokusu, pak platí:*

$$\lim_{n\to\infty} P|(\frac{n_i}{n}) - p_i| < \varepsilon = 1 \tag{3.3}$$

Kapitola 4

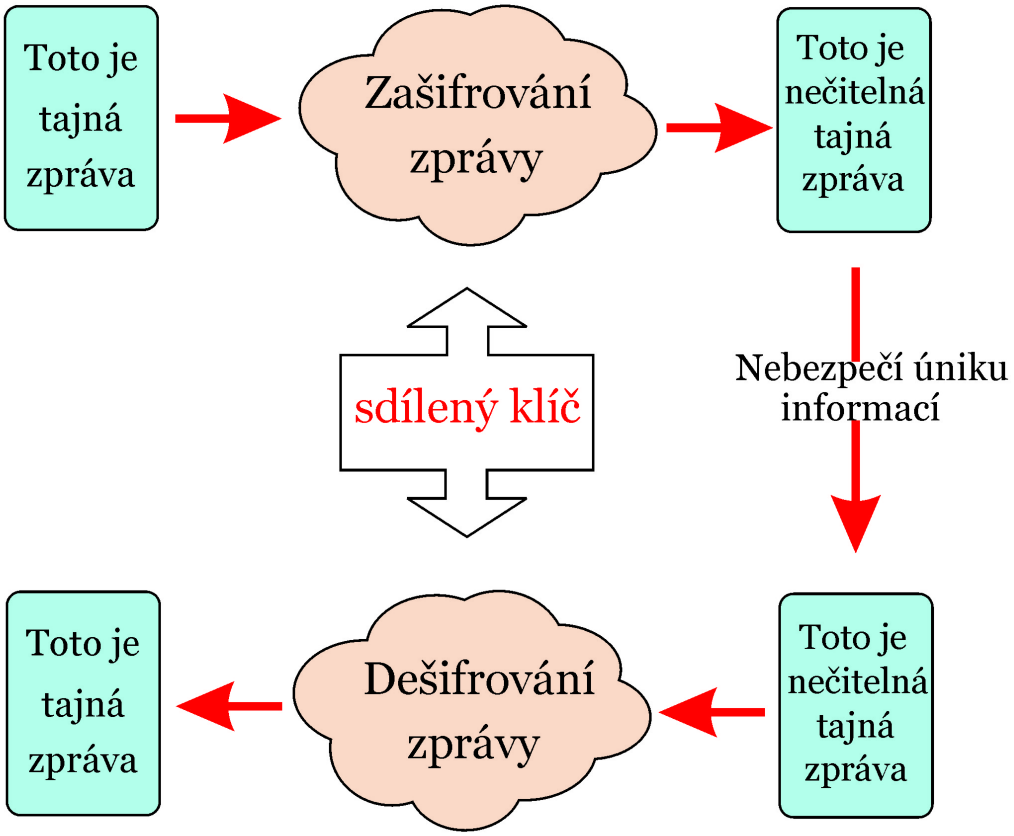
RSA

4.1 Symetrická šifra

Kryptoanalytici i po 2. světové válce pokračovali v rozvíjení a aplikaci počítačové technologie pro luštění šifer, čímž také přispěli k podobě dnešních moderních počítačů. Počítač lze například naprogramovat tak, aby napodobil činnost až stovek scramblerů z Enigmy (kapitola 2.5), z nichž některé se točí proti směru hodinových ručiček, některé zmizí po každém patnáctém písmenu a některé rotují rychleji než ostatní.

Ve všech předchozích kapitolách jsme popisovali pouze symetrické šifry, které při šifrování i dešifrování používají stejný postup (obrázek 4.1). Tzn. šifrujeme i dešifrujeme pomocí stejného klíče. I když může být toto šifrování velmi bezpečné a i obtížně odhalitelné, nastává problém s distribucí těchto klíčů. Toto je také nejslabší stránka symetrického šifrování. Pokud dojde k vyzrazení klíče, je celé šifrování zbytečné. Například pro Enigmu existovala kniha, kde bylo napsáno, jaké nastavení se používá ten který den. Toto nastavení bylo vlastně šifrovacím klíčem a pokud se kniha dostala do rukou nepřítele, byla celá tajná komunikace vyzrazená.

Ale opravdu potřebujeme jak pro šifrování, tak pro dešifrování



Obrázek 4.1: Symetrické šifrování

stejný klíč? Na tuto otázku se snažili matematici nalézt odpověď od konce 2. světové války. Jejich úsilí skončilo úspěšně. Bylo objeveno asymetrické šifrování.

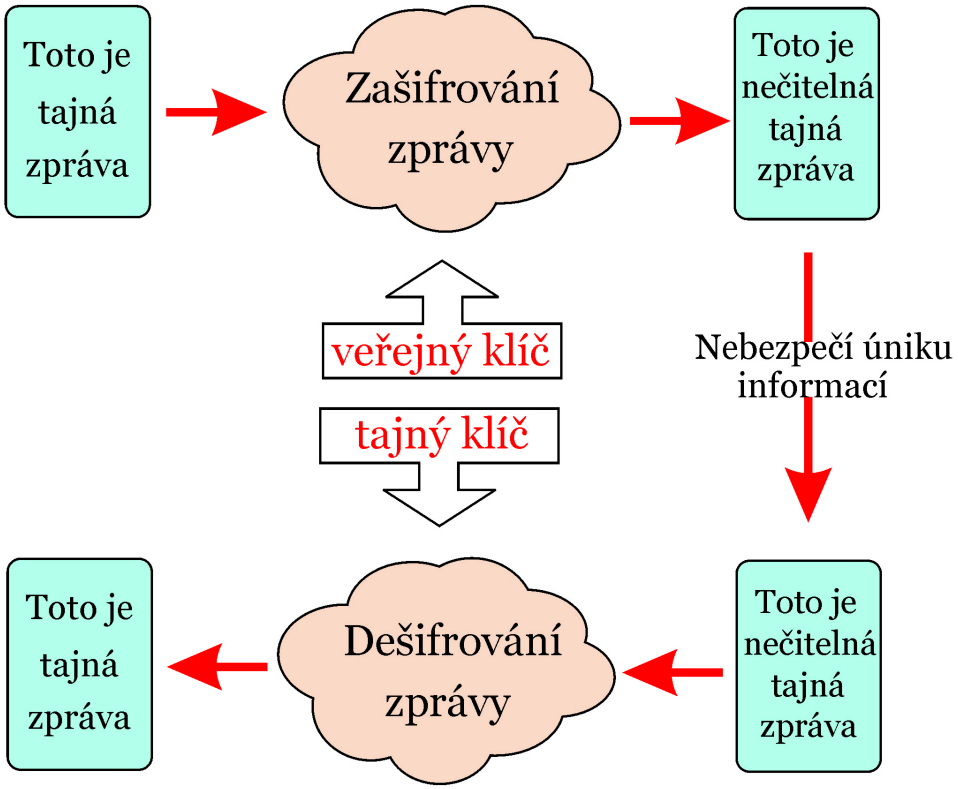
4.2 Asymetrická šifra

Asymetrická šifra je založena na páru klíčů, z nichž jeden umožňuje data zašifrovat a druhý dešifrovat. Podstatné je, že tyto klíče jsou různé a není možno jeden od druhého odvodit. Proto můžeme bez obav jeden klíč distribuovat (říkáme mu veřejný klíč) a ten druhý používat k opačnému postupu při šifrování (privátní klíč). To, jaký z klíčů (šifrovací nebo dešifrovací) si necháme, záleží na povaze úkonu, který se s nimi chystáme provádět.

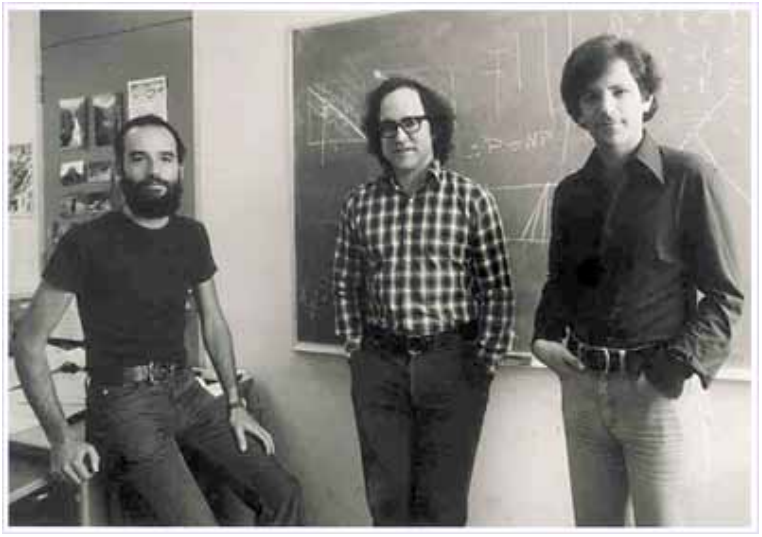
Šifrovací systém RSA získal svůj název dle počátečních písmen jmen svých objevitelů, kterými byli R. L. Rivest, A. Shamir a L. Adleman. Objev tohoto prvního, v praxi reálně použitelného, šifrovacího systému s veřejným klíčem, oznámili autoři 4. dubna 1977. Veřejný klíč je generován s použitím velkých prvočísel a celá bezpečnost RSA je založena na problému faktorizace velkých čísel na prvočísla. Základem je výběr dvou velkých (řádově stovky cifer) prvočísel, která se vynásobí. Na základě jejich součinu je vygenerován jak veřejný, tak i privátní klíč. Bez znalosti původních prvočísel je prakticky nemožné součin rozložit zpět na počáteční prvočísla. Algoritmus RSA se využívá například při digitálním podpisu.

4.3 Popis šifrování a dešifrování pomocí RSA

1. Nejprve náhodně vybereme dvě obrovská prvočísla p a q. Prvočísla musí být mimořádně velká. My si ale pro jednoduchost



Obrázek 4.2: Asymetrické šifrování



Obrázek 4.3: R. L. Rivest, A. Shamir a L. Adleman [27]

zvolíme $p = 61$, $q = 59$. Tyto hodnoty musí být uchovány v tajnosti.

2. Čísla p a q vynásobíme mezi sebou a dostaneme číslo N. V našem případě $N = 3\ 599$.

3. Náhodně vybereme číslo e, v tomto případě $e = 17$. e musí být nesoudělné s $\varphi(n)$.

$$\varphi(n) = (p-1)(q-1) = N + 1 - p - q \qquad (3.1)$$

4. Hodnoty e a N jsou veřejné. Vzhledem k tomu, že jsou tato dvě čísla pro šifrování nezbytná, musí být k dispozici každému, kdo chce zašifrovat zprávu. Dohromady se tato čísla nazývají veřejný klíč.

$$\text{veřejný klíč } (N,\ e)$$

5. Pro zašifrování musí být zpráva nejdříve převedena do čísla M. Toho se dá dosáhnout například tak, že se slovo převede

do ASCII binárních číslic, které lze pokládat za číslo v desítkové soustavě. M je potom zašifrováno tak, aby vytvořilo zašifrovaný text C podle vzorce:

$$C \equiv M^e \bmod N \tag{3.2}$$

Většinu výpočtů v této kapitole bychom jen těžko zvládli na běžné kalkulačce, proto použijeme matematický program Maple.(Kapitola 4.4.1)

6. V našem případě se pokusíme zašifrovat symbol pro podpis, tzv. zavináč: @. Jeho ASCII reprezentaci převedeme do desítkové soustavy a dostaneme číslo 64. Takže $M = 64$.

7. K zašifrování budeme potřebovat veřejný klíč, tedy čísla $N = 3599$ a $e = 17$. Potřebný vzorec pro zašifrování čísla $M = 64$ vypadá po dosazení takto:

$$C \equiv 64^{17} \bmod 3\ 599$$

8. Šifrový text zní $C = 3\ 164$.

9. Mocniny v modulární matematice jsou jednosměrné funkce, takže je velmi těžké postupovat nazpět a zprávu $C = 3164$ rozšifrovat na originální zprávu M.

10. Příjemci zprávy se to však podaří, protože zná hodnoty p a q. Vypočítá speciální číslo d neboli svůj privátní klíč.

privátní klíč (d)

Číslo d se počítá podle následujícího vzorce:

$$e \cdot d \equiv 1 \bmod (p-1)(q-1) \tag{3.3}$$
$$17 \cdot d \equiv 1 \bmod 60 \cdot 58$$

$$17 \cdot d \equiv 1 \bmod 3\,480$$
$$d = 1\,433$$

Výpočet hodnoty d není úplně samozřejmý, ale postup známý jako rozšířený Euklidův algoritmus nám umožní nalézt d rychle a jednoduše. K výpočtu použijeme program Matlab 6.5.(Kapitola 4.4.2)

11. K rozluštění zprávy slouží následující vzorec:

$$M = C^d \bmod N \tag{3.4}$$
$$M \equiv 3164^{1433} \bmod 3599$$
$$M \equiv 64$$

12. $M =$ @ v ASCII

    Pánové Rivest, Shamir a Adleman vytvořili speciální jednosměrnou funkci, kterou může invertovat pouze ten, kdo má přístup k důvěrným informacím, a to k hodnotám p a q. Funkce je určena zvolením čísel p a q, které po vynásobení dají N. Funkce umožní odesílateli zašifrovat zprávu pro příjemce. Odesílatel zná jen N, zatímco určený příjemce je jediný, kdo zná p a q, tudíž je jedinou osobou, která zná dešifrovací klíč d.

4.4 Využití matematických programů

4.4.1 Program Maple 9.5

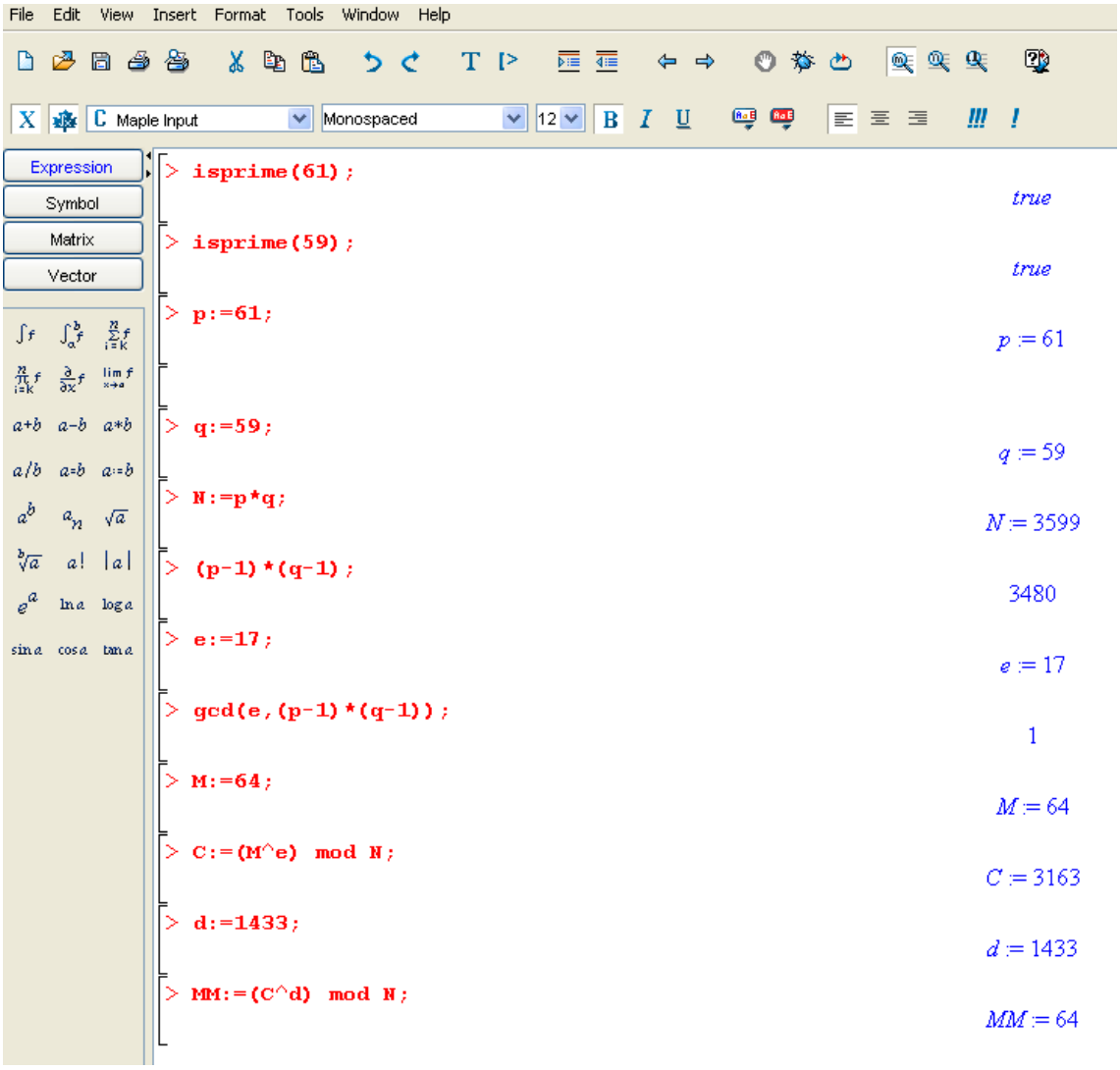
- Nejprve ověříme, že námi zvolená čísla p a q jsou prvočísla.

  $>$ **isprime(61);**

  true

  $>$ **isprime(59);**

  true



Obrázek 4.4: Program Maple

- Protože čísla čísla 59 a 61 jsou prvočísla, můžeme s nimi dále počítat.

  $>$ **p:=61;**

  $p := 61$

  $>$ **q:=59;**

  $q := 59$

- Vynásobení těchto dvou prvočísel dostaneme číslo N.

  $>$ **N:=p*q;**

  $N := 3599$

  $>$ **(p-1)*(q-1);**

  3480

- Nyní definujeme druhou část veřejného klíče, číslo e.

  $>$ **e:=17;**

  $e := 17$

- Čísla e a $(p-1) \cdot (q-1)$ musí být nesoudělná, aby nám všechny vzorce fungovaly. To znamená, že jejich největší společný dělitel má být roven jedné.

  $>$ **gcd(e,(p-1)*(q-1));**

  1

- Nyní zvolíme symbol, který chceme zašifrovat.

  $>$ **M:=64;**

  $M := 64$

- Symbol zašifrujeme pomocí této rovnice.

  $>$ **C:=(M^ e) mod N;**

  $C := 3163$

- Příjemce vytvoří dešifrovací klíč d. (Viz. kapitola Využití programu Matlab 4.4.2).

  $>$ **d:=1433;**

  $d := 1433$

- Tato rovnice nám umožní zprávu dešifrovat a získat tak zpět zašifrovaný znak.

  $>$ **MM:=(C^ d) mod N;**

  $MM := 64$

- Prostředí programu Maple 9.5 vypadá takto 4.4.

4.4.2 Program Matlab 6.5

- Program Matlab použijeme pro vyřešení této rovnice.

$$17 \cdot d \equiv 1 \bmod 60 \cdot 58$$

- Tato rovnice se řeší pomocí rozšířeného Eukleidova algoritmu. V programu vytvoříme funkci, nazveme ji rea. rea má tři neznámé:

  a je číslo, kterým násobíme d.
  b je zbytek po dělení.
  c je součin $(p-1) \cdot (q-1)$.

  $functionx = rea(a, b, c)$

- Nyní použijeme řetězec for. Zavedeme proměnnou d, která se mění od jedné do hodnoty c.

  $for \qquad d = 1 : c$

- Dále zavedeme proměnnou k.

  $k = a * d;$

- Funkce, do které dosazujeme vypadá takto:

  $x = mod(k, c)$;

- Pokud se hodnota x bude rovnat b, v našem případě to znamená, že zbytek po dělení bude 1, pak program vypíše hodnotu d.

  $if(x == b)$

  $vysledek = d$;

  $display(d)$;

  $break$;

  end

  end

- Takto lze funkci rea vyvolat pro různé hodnoty a, b, c.

  $>> rea(17, 1, 3480)$

  $d =$

  1433

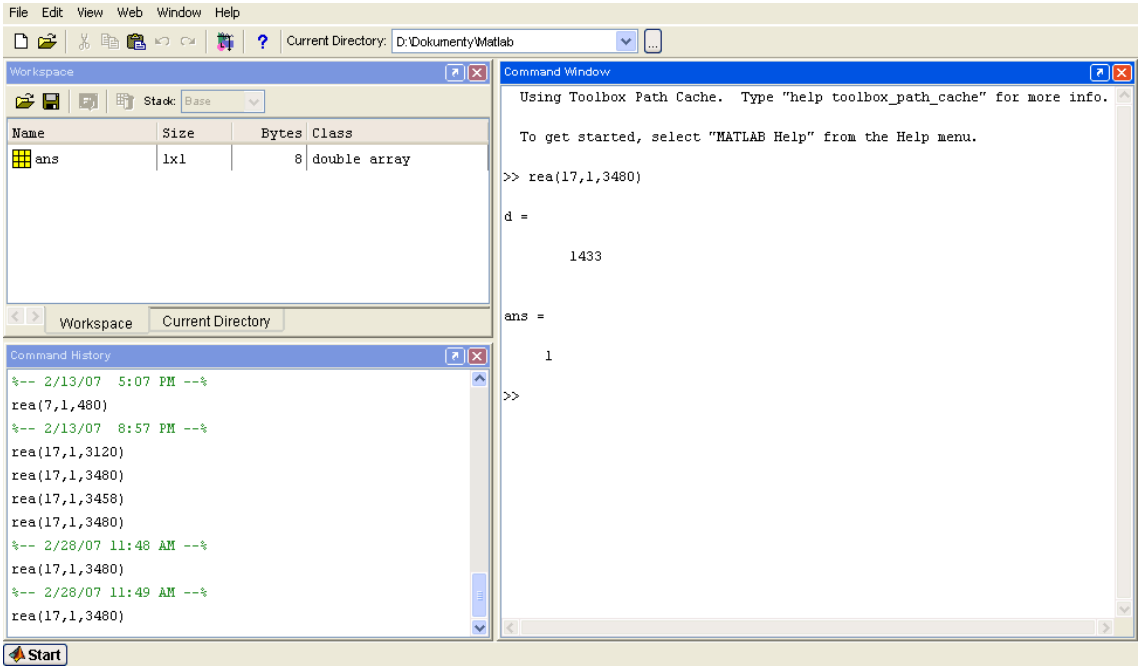
  $ans =$

  1

- Našli jsme hodnotu d a vyřešili rovnici.
- Na obrázku 4.5 vidíte prostředí programu Matlab 6.5.

4.5 Důkaz vzorce pro dešifrování

Nyní dokážeme vzorec z kapitoly 4.3, pomocí kterého jsme dešifrovali zprávu zašifrovanou pomocí algoritmu RSA.



Obrázek 4.5: Program Matlab

Příjemce zprávy vypočte svůj privátní klíč d, který nám po dosazení do uvedeného vzorce pomůže získat čitelný text.

$$M \equiv C^d \bmod N \tag{5.1}$$

Víme, že platí tato rovnice

$$e \cdot d \equiv 1 \bmod \varphi. \tag{5.2}$$

z toho vyplývá, že existuje $k \in \mathbb{Z}$, takové že platí:

$$e \cdot d = 1 + k \cdot \varphi \tag{5.3}$$

Pokud jsou m a p nesoudělná, použijeme malou Fermatovu větu:

$$M^{p-1} \equiv 1 \bmod p \tag{5.4}$$

Umocněním obou stran výrazem $k \cdot (q-1)$ dostaneme:

$$M^{k\cdot(p-1)\cdot(q-1)} \equiv 1 \bmod p. \tag{5.5}$$

Dále vynásobíme rovnici číslem M.

$$M^{1+k\cdot(p-1)\cdot(q-1)} \equiv M \bmod p. \tag{5.6}$$

Pokud je největší společný dělitel čísel M a p číslo p, pak to tato rovnice také platí, protože obě strany po dělení číslem p dávají zbytek 0.
Ve všech případech tedy podle věty 3.1.6 vychází:

$$M^{e\cdot d} \equiv M \bmod p. \tag{5.7}$$

Stejným způsobem se odůvodní i obdobná rovnice pro q

$$M^{e\cdot d} \equiv M \bmod q. \tag{5.8}$$

Protože p a q jsou prvočísla, můžeme tvrdit:

$$M^{e\cdot d} \equiv M \bmod N. \tag{5.9}$$

A odvodíme vzorec pro M, který jsme používali při dešifrování RSA.

$$C^d \equiv (M^e)^d \equiv M \bmod N. \tag{5.10}$$

4.6 Bezpečnost RSA

Tato kapitola je zaměřena na bezpečnost šifrování pomocí algoritmu RSA. Zabývá se jednak útoky, které jsou vedeny k prolomení tohoto šifrování, a také vlastnostmi algoritmu RSA, které útoky úspěšně odráží.

RSAP - RSA problém Zprávu zašifrovanou pomocí algoritmu RSA může rozšifrovat jen ten, kdo zná privátní klíč d. Opravdu ji však nemůžou dešifrovat ti, kteří znají jen veřejný klíč (N, e)? Žádný efektivní algoritmus, který by dokázal získat zpět čistý text, jen pomocí veřejného klíče, neexistuje. Takto lze formulovat RSA problém.

Jediný možný postup je, že útočník (člověk, který nezná privátní klíč, a přesto chce zprávu dešifrovat) dokáže číslo N rozložit na dvě prvočísla p a q. Poté vypočítá d, tak jako to dělá pravý příjemce zprávy, a pak už je schopen číst všechny zprávy určené příjemci.

Na druhou stranu, útočník může nejprve nějakým způsobem spočítat d, a pak už lehce rozloží N na prvočísla (faktorizuje N) a opět zprávu dešifruje.

To znamená, že problém výpočtu privátního klíče d z veřejného klíče (N, e) a problém faktorizace N jsou ekvivalentní. Když vytváříme RSA klíče je rozhodující vybrat p, q taková, aby faktorizace $N = p \cdot q$ byla v reálném čase neproveditelná.

Malé číslo e Aby se zlepšila efektivita šifrování, je výhodné vybrat malé číslo e, například $e = 3$. Velká skupina lidí může mít ve svém privátním klíči stejné e, nicméně každý z této skupiny lidí musí mít jiné číslo N. Představme si situaci, že chceme poslat stejnou zprávu třem přátelům z naší skupiny, všichni mají stejné číslo e, ale různé N. Zprávu M budeme šifrovat třikrát tímto způsobem:

$$C_i \equiv M^3 \bmod N_i \tag{6.1}$$

pro $i = 1, 2, 3$. Útočník, který naši komunikaci odposlouchává se dozví hodnoty C_1, C_2, C_3 a pomocí Gaussova algoritmu najde řešení X těchto tří rovnic:

$$X \equiv C_1 \bmod N_1 \tag{6.2}$$

$$X \equiv C_2 \bmod N_2 \tag{6.3}$$

$$X \equiv C_3 \bmod N_3 \tag{6.4}$$

Protože $M^3 < N_1 \cdot N_2 \cdot N_3$ platí podle Čínské věty o zbytcích rovnost:

$$M^3 = X \tag{6.5}$$

A proto by malé číslo e nemělo být používáno k šifrování stejných zpráv určených mnoha lidem. Jako prevence proti tomuto způsobu dešifrování lze použít tzv. „okořenění zprávy“. Vygenerujeme řetězec náhodných znaků vhodné délky, který vložíme na začátek zprávy. Každá zpráva nyní začíná řetězcem odlišných znaků a tudíž nelze použít pro dešifrování předchozí postup.

Útok vyhledáváním Je-li rozsah zprávy velmi malý nebo předvídatelný, pak protivník může text C dešifrovat tak, že postupně zašifruje všechny možné zprávy, až získá hledané C. Jednoduchý způsob, jak tento problém obejít, je popsán v předchozím odstavci, jde o takzvané "okořenění zprávy".

Malé číslo d Tak jako v případě, kdy jsme vyšetřovali malé číslo e, může být vybráno malé číslo d, aby se zlepšila efektivita šifrování. Uvažujme případ, kdy nejprve vybereme d a poté spočítáme e. Avšak, je-li nejmenší společný násobek $(p-1)$ a $(q-1)$ malý, což většinou bývá, a je-li d přibližně rovno $\frac{N}{4}$, pak je možné vypočítat d z veřejného klíče. Abychom se vyhnuli tomuto nebezpečí, mělo by být číslo d zhruba stejné velikosti jako číslo N.

Násobné vlastnosti M_1, M_2 jsou dva prosté texty, které budeme šifrovat a C_1, C_2 jsou tyto texty po zašifrování pomocí algoritmu RSA. Všimneme si, že platí:

$$(M_1 \cdot M_2)^e \equiv M_1^e \cdot M_2^e \equiv C_1 \cdot C_2 \bmod N \tag{6.6}$$

Jinými slovy, prostý text $M \equiv M_1 \cdot M_2 \bmod N$ náleží zašifrovanému textu $C \equiv C_1 \cdot C_2 \bmod N$. Toto je někdy popisováno

jako „homomorfní“ vlastnost algoritmu RSA. Tato skutečnost vede k možnému útoku proti algoritmu RSA.

Předpokládejme, že protivník se snaží dešifrovat konkrétní zašifrovaný text $C \equiv M^e \bmod N$, ten je ale určen jinému příjemci. Dále předpokládejme, že pravý příjemce je schopný dešifrovat libovolný text určený pro protivníka. Protivník se snaží utajit hodnotu C a proto zvolí náhodné číslo $x \in \mathbb{Z}$ a vypočte hodnotu $\overline{C}$.

$$\overline{C} \equiv C \cdot x^e \bmod N \tag{6.7}$$

Pravý příjemce obdrží hodnotu $\overline{C}$ a spočítá hodnotu $\overline{M}$.

$$\overline{M} \equiv \overline{C}^d \bmod N \tag{6.8}$$

Pokud se hodnota $\overline{M}$ dostane zpět k protivníkovi, může protivník spočítat původní M takto:

$$\overline{M} \equiv \overline{C}^d \equiv C^d \cdot x^{e \cdot d} \equiv (M \cdot x) \bmod N \tag{6.9}$$

$$M \equiv \overline{M} \cdot x^{-1} \bmod N \tag{6.10}$$

Tento útok můžeme v praxi přelstít omezením ve struktuře textu ještě před zašifrováním. Když pak příjemce obdrží zašifrovanou zprávu, která po dešifrování nemá danou strukturu, jde jistě o podvod. Toto omezení nám poskytne silnou technickou ochranu proti tomuto a jemu podobným útokům.

Problém s konstantou Opravdu podstatné je, aby si každý vybral své vlastní číslo N. Občas se doporučuje, aby nějaká důvěryhodná osoba vybrala jediné N a pak distibuovala různé (e_i, d_i) všem lidem v síti. Nicméně, jak už víme z této kapitoly, znalost (e_i, d_i) umožňuje faktorizaci N a tudíž každý z této skupiny může zjistit privátní klíče všech ostatních osob v síti. A také, jestliže byla zašifrována zpráva a poslána dvěma či více osobám v síti, pak ji mohla odposlouchávat

osoba mimo síť. Tento protivník může s velkou pravděpodobností dešifrovat zprávu a stačí mu k tomu pouze veřejně dostupné informace.

Cyklické útoky Opět budeme šifrovat tento text:

$$C \equiv M^e \bmod N \tag{6.11}$$

Číslo k je kladné celé číslo, takové že platí:

$$(C^e)^k \equiv C \bmod N \tag{6.12}$$

Dále platí také:

$$(C^e)^{k-1} \equiv M \bmod N \tag{6.13}$$

Tento postřeh vede k cyklickému útoku. Protivník bude dosazovat do rovnice 6.12 za $k = 0, 1, \ldots$. Když bude platit $(C^e)^k \equiv C \bmod N$, pak předchozí číslo v cyklu $(C^e)^{k-1} \equiv M \bmod N$ určuje původní text M.

Všeobecný cyklický útok spočívá v problému nalézt nejmenší kladné celé číslo u, takové že:

$$f = D((C^e)^u - C, N) > 1 \tag{6.14}$$

Když

$$(C^e)^u \equiv C \bmod p \wedge (C^e)^u \not\equiv C \bmod q \tag{6.15}$$

pak $f = p$. Obdobně, když

$$(C^e)^u \not\equiv C \bmod p \wedge (C^e)^u \equiv C \bmod q \tag{6.16}$$

pak $f = q$. V obou případech bylo N faktorizováno a protivník může získat d a následně i M. Na druhou stranu, když

$$(C^e)^u \equiv C \bmod p \wedge (C^e)^u \equiv C \bmod q \tag{6.17}$$

pak $f = N$ a $(C^e)^u \equiv C \bmod N$. Ve skutečnosti, u musí být nejmenší kladné číslo k pro které platí $(C^e)^k \equiv C \bmod N$. V tomto případě je útok úspěšný a M může být vypočítáno. Tento cyklický útok lze v podstatě považovat za algoritmus pro faktorizaci čísla N, většinou je však ukončený ještě dříve než vůbec nastane. Cyklické útoky nepředstavují v praxi hrozbu pro bezpečnost šifrování pomocí algoritmu RSA.

Utajené a neutajené zprávy M je text, který chceme zašifrovat, a pro který platí $0 \leq M \leq N - 1$. Může nastat zajímavá situace, kdy při použití šifrovacího algoritmu nedojde k zašifrování zprávy, například $M^e = M \bmod N$. Při použití algoritmu RSA vždy existují zprávy, které se nepodaří utajit. Například jde o $M = 0$, $M = 1$, a $M = N - 1$. Přesný počet neutajených zpráv pro konkrétní šifrování vypočítáme takto:

$$[1 + D(e - 1, p - 1)] \cdot [1 + D(e - 1, q - 1)] \tag{6.18}$$

Množství neutajených zpráv bývá minimálně 9. Jsou-li p a q náhodná prvočísla, a je-li e také vybráno náhodně nebo je použité malé e ($e = 3$ nebo $e = 2^{16} + 1 = 65\,567$), pak podíl zpráv, které jsou neutajené pomocí šifrování RSA, bude obecně zanedbatelně malý a tyto neutajené zprávy nebudou v praxi znamenat hrozbu pro bezpečnost algoritmu RSA.

4.7 RSA v praxi

RSA patent Šifrovací algoritmus RSA byl patentován v USA a v Kanadě. Mnoho světových organizací tvoří, nebo už vy-

tvořilo, šifrovací systémy, které používají RSA při šifrování, digitální podpisy a zprávu klíčů.

Jsou různé způsoby jak urychlit RSA šifrování a dešifrování, ať už jde o hardwarové či softwarové změny. I s takovými změnami je RSA stále pomalejší než běžně užívané algoritmy se symetrickými klíči. V praxi se RSA většinou používá pro bezpečný přenos klíčů pro symetrické šifrování a pro šifrování málo rozsáhlých zpráv.

Doporučená délka čísla N Doporučená délka čísla N je 512-bitů. Tato délka však už neposkytuje účinnou ochranu proti koncentrovanému útoku. V roce 1996 byly objeveny nové silné algoritmy na faktorizaci čísel, proto se dnes doporučuje délka pro N až 768-bitů. Pro dlouho trvající bezpečnost by délka měla být až 1024-bitů či delší.

Výběr prvočísel Jak už bylo řečeno, prvočísla p a q by měla být vybrána tak, aby faktorizace N byla počítačově velmi obtížná. Hlavní omezení pro p a q, abychom se vyhnuli silným algoritmům na faktorizaci čísel, je, že prvočísla p a q by měla být stejně dlouhá a dostatečně velká. Pokud například použijeme N o délce 1024-bitů, pak každé z prvočísel p a q musí být dlouhé 512-bitů.

Jiná podmínka pro prvočísla p a q je, že rozdíl $p - q$ by neměl být příliš malý. Je-li tento rozdíl malý, pak platí $p \approx q$ a $p \approx \sqrt{N}$. V tomto případě může být N jednoduše faktorizováno hledáním prvočísel blízkých $\sqrt{N}$. Jsou-li p a q zvoleny náhodně, pak $p - q$ je skoro jistě dostatečně velké.

Mnoho odborníků doporučuje, že p a q by měla být *silná prvočísla*. Prvočíslo p je silné prvočíslo, pokud splňuje následující podmínky:

(a) p je velmi velké prvočíslo.

(b) $p-1$ má velkého prvočíselného dělitele.
To znamená $p = x_1 \cdot y_1 + 1$, kde $x_1 \in \mathbb{Z}$ a y_1 je velké prvočíslo.

(c) $y_1 - 1$ má velkého prvočíselného dělitele.
To znamená $y_1 = x_2 \cdot y_2 + 1$, kde $x_2 \in \mathbb{Z}$ a y_2 je velké prvočíslo.

(d) $p+1$ má velkého prvočíselného dělitele.
To znamená $p = x_3 \cdot y_3 - 1$, kde $x_3 \in \mathbb{Z}$ a y_3 je velké prvočíslo.

Je-li p náhodně vybrané a přiměřeně dlouhé, pak můžeme očekávat velké prvočíselné dělitele čísel $p-1$ a $p+1$. Výše zmíněné podmínky maří silné algoritmy na faktorizaci čísel a řeší problém zacyklení z předchozí kapitoly. Jelikož současný stav znalostí faktorizačních algoritmů není příliš veliký, není nutné používání silných prvočísel při generování klíče RSA. Na druhou stranu nejsou silná prvočísla méně bezpečná, než náhodná prvočísla a vyžadují pouze minimální čas navíc při počítačovém zpracování. Není tedy důvod silná prvočísla nepoužít.

Malé číslo e Šifrování pomocí algoritmu RSA bývá urychlováno výběrem malého čísla e. V praxi je obvykle používáno $e = 3$, v tomto případě je třeba, aby platilo $3 \nmid (p-1) \wedge 3 \nmid (q-1)$. Velmi často se používá i $e = 2^{16} + 1 = 65\,567$.
Číslo $e = 65\,567$ navíc odolává útoku uvedenému v předchozí kapitole, kdy šifrujeme stejnou zprávu více příjemcům, kteří mají stejné číslo e, a proto je bezpečnější než $e = 3$.

Kapitola 5

Frekvenční analýza

Ve všech předchozích kapitolách jsme se snažili zašifrovat zprávy, aby si je mohli přečíst pouze lidé, kterým jsou určeny. Zabývali jsme se *kryptografií*. Avšak současně s kryptografií se rozvíjela i *kryptoanalýza*.[1] Což je proces luštění zašifrovaných zpráv bez dešifrovacího klíče. Základní metodou, která se využívá v kryptoanalýze, je *frekvenční analýza*. Jde o „útok hrubou silou", který využívá vlastnosti jazyka. Frekvenční analýzu budeme demonstrovat na příkladu.(Tabulka 5).

Jde o soutěžní úkol, který luštili čtenáři časopisu Crypto-World [16] v říjnu 2000. Ze zadání víme, že jde o jednoduchou záměnu, text je v češtině a bez mezer. Použitá je mezinárodní abeceda A-Z, tj. 26 znaků.

Každý jazyk má svou vlastní charakteristiku. Charakteristiky různých jazyků se získávají z velmi dlouhých textů (min. 10 000 znaků). V těchto textech spočítáme četnosti a relativní četnosti všech znaků dané abecedy. Charakteristiky českého, anglického, francouzského, německého, českého a slovenského obecného textu najdeme v tabulce 5.2 [30]. Histogram českého textu pak na obrázku 3.2.

[1] Kryptoanalýza je proces luštění původních zpráv ze zašifrovaného textu v případě, kdy není k dispozici příslušný klíč. Kryptologie je pak souhrnným označením pro kryptografii a kryptoanalýzu [7].

UFTAL	OTCSF	CILDO	TGLUL	JHSFN	PZIHF
NBGZU	FTALP	ZRZOB	NCHSF	NQBZA	ZFZGX
ZWOZG	OLPZX	AHBHU	FTALP	ZXIHJ	OTWZJ
HFAZD	NDTOS	BZLFN	WCHPR	ZPHCI	TUXHI
ZCITD	ZSAWT	BCHSF	NDNFT	ALPZG	ZGZPZ
WZIZD	NQAHS	WZOTP	TCOZJ	RZHWT	UBTPZ
HJOZW	TUBHB	LHJUB	ALOTP	ZWLUB	TOLXL
JZOZI	LADLP	TCPNG	SGDNU	ZOLOL	ULQIT
DHUBX	IHJOT	WZDHJ	HSRZG	OLPQH	SGZXI
HJOTW	ZRZXA	ZUBLA	ILOZD	CHJOL	QZUFZ
ASXAT	AHGQI	LJONW	CXAHW	ZUZWC	PHCHS
DGOTQ	LBTOZ	FZGXZ	WOZIL	BQNRL	QHOLX
ARZJO	ZSATO	BLQUZ	PHCHS	UBLBT	BGDRZ
JIZCH	SFNJA	SCHBO	ZRZJH	DLBNP	TDNRP
SBZXI	HJOTW	ZSQIL	JLPZJ	HHBZD	AZONW
CJNWC	LRTWT	WCHFL	ISOZF	HBDSG	LDAZO
NWCOZ	XAHXS	UBONW	CHFLI	ZWCUZ	XIHJO
TWZBG	DGLXL	ATGDI	LUBZO	ZDCHJ	OZRZS
IHGZO	TDXHI	NZBNI	ZOHDN	WCULW	WTWCO
ZRDCH	JOZRU	TPTHF	LINGS	UBLDL	RTBAL
JTWOZ	XIZBZ	OZQHU	TWQNO	ZRXHG	JZRTJ
ASCNJ	ZOXHU	FZASP	LRTFN	BCHSF	NGXAL
WHDLO	NEEEE				

Tabulka 5.1: Zašifrovaný text

Vraťme se k zadanému úkolu. Opět spočítáme četnosti jednotlivých znaků. (Tabulka 5.3). I když jde z hlediska frekvenční analýzy o velmi krátký text, pokusíme se tyto četnosti porovnat s četnostmi znaků v obecném českém textu. (Tabulka 5.2). Počítání četností zabere mnoho času. Pro urychlení naší práce můžeme použít programy volně dostupné na internetu.[2]

Písmeno, které se nejčastěji opakuje v našem zadaném textu, bude s vysokou pravděpodobností odpovídat písmenu, které se nejčastěji vyskytuje v obecném českém textu. Druhé písmeno z textu by mohlo odpovídat druhému nejčastějšímu písmenu z obecného textu atd. Nebude to však platit pro všechna písmena, protože náš text je velmi krátký, aby pokryl dostatečné zastoupení všech písmen. Dále může jít o tematicky zaměřený text. Například četnosti písmen v textu z lékařského prostředí nebudou odpovídat četnostem písmen v obecném českém textu.

[2] Na internetových stránkách časopisu Crypto-world [16] lze stáhnout dva použitelné programy JZ a VFQ. VFQ je navíc schopný dešifrovat samohlásky.

Písmeno	angl.	franc.	něm.	češ.	slov.
A	7,96	7,68	5,52	8,99	9,49
B	1,60	0,80	1,56	1,86	1,90
C	2,84	3,32	2,94	3,04	3,45
D	4,01	3,60	4,91	4,14	4,09
E	12,86	17,76	19,18	10,13	9,16
F	2,62	1,06	1,96	0,33	0,31
G	1,99	1,10	3,60	0,48	0,40
H	5,39	0,64	5,02	2,06	2,35
I	7,77	7,23	8,21	6,92	6,81
J	0,16	0,19	0,16	2,10	2,12
K	0,41	0,00	1,33	3,44	3,80
L	3,51	5,89	3,48	4,20	4,56
M	2,43	2,72	1,69	2,99	2,97
N	7,51	7,61	10,20	6,64	6,34
O	6,62	5,34	2,14	8,39	9,34
P	1,81	3,24	0,54	3,54	2,87
Q	0,17	1,34	0,01	0,00	0,00
R	6,83	6,81	7,01	5,33	5,12
S	6,62	8,23	7,07	5,74	5,94
T	9,72	7,30	5,86	4,98	5,06
U	2,48	6,05	4,22	3,94	3,70
V	1,15	1,27	0,84	4,50	4,85
W	1,80	0,00	1,38	0,06	0,06
X	0,17	0,54	0,00	0,04	0,03
Y	1,52	0,21	0,00	2,72	2,57
Z	0,05	0,07	1,17	3,44	2,72

Tabulka 5.2: Charakteristiky některých jazyků [30]

A	B	C	D	E	F	G	H	I	J	K	L	M
28	32	28	26	4	22	22	52	26	27	0	48	0
4,18	4,78	4,18	3,88	0,60	3,28	3,28	7,76	3,88	4,03	0,00	7,16	0,00
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
29	46	21	13	18	26	42	25	0	31	22	0	82
4,33	6,87	3,13	1,94	2,69	3,88	6,27	3,73	0,00	4,63	3,28	0,00	12,24

Tabulka 5.3: Četnost znaků v textu

Z	H	L	T	N	S	C
↓	↓	↓	↓	↓	↓	↓
E	O	A	I	Y	U	H

Tabulka 5.4: Dešifrování některých znaků pomocí frekvenční analýzy

V zadaném zašifrovaném textu se nejčastěji objevuje písmeno *Z*, písmeno *Z* tedy bude odpovídat písmenu *E* v čistém textu. Dále uplatňujeme stejné pravidlo a navíc použijeme program VFQ a identifikujeme samohlásky. Frekvenční analýza nám pomohla k dešifrování znaků zapsaných v tabulce 5.4.

Pomocí těchto sedmi znaků však ještě nejsme schopni rozluštit celý text. Dále se doporučuje postup, kdy nad písmena v zašifrovaném textu píšeme písmena čistého textu a snažíme se doplnit mezery. Vraťme se k zadání v tabulce 5. Několikrát se v tabulce vyskytuje bigram[3] *CS* (dešifrujeme na *HU*) a trigram *CHS* (dešifrujeme na *HOU*). Jelikož po *CS* i *CHS* následuje písmeno *F*, pokusíme se *F* dešifrovat jako *B*. Naše domněnka byla správná, protože z *CHSFN* nám vzniklo slovo *HOUBY*. V několika případech po *CHS* není *F*, zato je před ním bigram *PH*. V těchto případech bude *PHCHS* znamenat *MOHOU*.

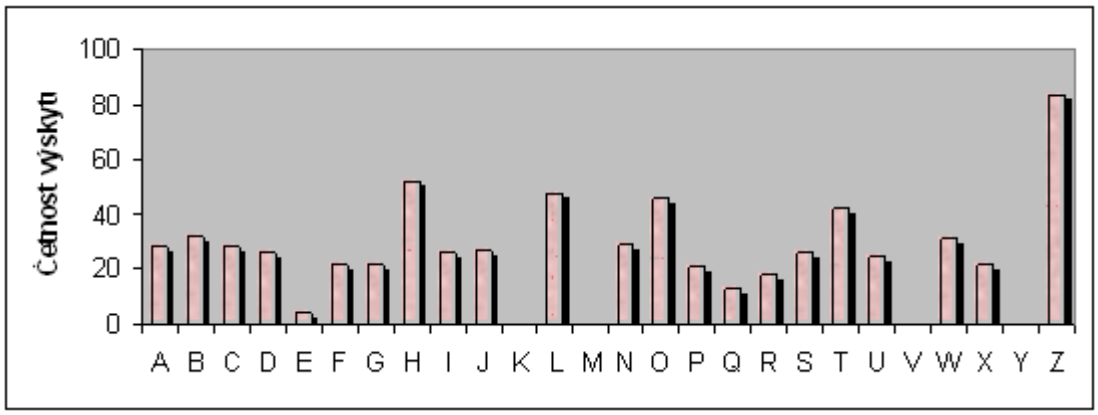
Podobně se snažíme uhádnout i dalších symboly a doplňujeme text tak, aby dával smysl. Vyluštění šifrovací abecedy je uvedeno v tabulce 5.5. Původní text vypadá takto:

Sbirani hub hlavni zasadou by melo byt ze sbirame jen ty houby ktere bezpecne zname proto sbirame plodnice dobre vyvinute abychom je mohli spolehlive urcit houby vybirame ze zeme cele vykroucenim ihned je ocistime od necistot a odstanime casti napadene larvami hmyzu zvysena nasaklivost plodnice vodou je znamkou ze plodnice je prestarla nevhodna ke sberu pri rozkladnych procesech mohou

[3] Bigramy jsou dvojice písmen. Tak jako jsme pro každý jazyk určovali relativní četnosti jednotlivých písmen, můžeme určovat i relativní četnosti bigramů a používat frekvenční analýzu. Trojice znaků jsou trigramy.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
R	T	H	V	X	B	Z	O	L	D	F	A	W	Y	N	M	K	J	U	I	S	G	C	P	Q	E

Tabulka 5.5: Převodová tabulka pro dešifrování



Obrázek 5.1: Četnost písmen v zašifrovaném textu

vznikat i nebezpecne latky jako napr jed neurin tak se mohou stat i tzv jedle houby druhotne jedovatymi vyjmute plodnice ukladame do otevrenych dychajicich obalu nebot v uzavrenych nepropustnych obalech se plodnice tzv zapari zvlaste nevhodne je ulozeni v polyetylenovych saccich nejvhodnejsimi obaly zustavaji tradicne pletene kosticky nejpozdeji druhy den po sberu maji byt houby zpracovany xxxx

V každém jazyce najdeme různé druhy opakování. Shody *typu A* jsou *polygramy*. Polygram je společný termín pro jednotlivé znaky, bigramy, trigramy atd. Kromě frekvenční analýzy můžeme při luštění šifer používat i teorii pravděpodobnosti. Například pravděpodobnost, že po souhlásce P bude následovat samohláska O bude určitě větší než, že po P následuje souhláska X.

Další metoda kryptoanalýzy vyšetřuje shody mezi různými texty, jde o *shody typu B*. Napíšeme-li dva dostatečně dlouhé texty pod

sebe, pak na určitých místech budou stát pod sebou totožné znaky. Tyto dvojice označíme za *shody*. Spočítáme-li počet shod mezi dvěma zašifrovanými texty mělo by se toto číslo rovnat počtu shod mezi dvěma otevřenými texty stejné délky v daném jazyce.

Při luštění zašifrovaných textů se dále používá test jednoabecednosti, test periodičnosti, charakteristika polohy znaku, počet chybějících písmen v textu a další.

Jak už bylo uvedeno v kapitole 2.5, československí zpravodajci za 2. světové války používali k šifrování tajných zpráv systém **TTS** (transpozice + transpozice + substituce). Bylo použito jedenáct hesel očíslovaných 1 až 0 a R. Například dne 12. listopadu byla denní depeše zašifrována nejprve pomocí hesla číslo 1 (ČESKOSLOVENSKÁ REPUBLIKA) a poté druhým transpozičním heslem (MISTR JAN HUS). Česká substituční abeceda měla celkem 45 znaků (26 písmen mezinárodní abecedy, písmena s háčky a interpunkci.). Později se jako hesla používali různé úryvky z knih.

J. Janeček [5] podrobil archivní materiály londýnské čs. vojenské rozvědky podrobné kryptoanalýze a dokázal, že:

> *„S pravděpodobností vyšší než 90 procent museli nepřátelští luštitelé takové zprávy rozluštit.“*

Přispělo k tomu mnoho faktorů. Jednak Češi používali všeobecně známé šifrovací systémy, které jen minimálně obměňovali a také často posílali depeše stejné délky a tak výrazně ulehčili práci německým dešifrátorům.

Kapitola 6

Závěr

Pro svou diplomovou práci jsem si zvolila velmi široké avšak z pohledu matematika velmi zajímavé téma. Cílem této práce bylo poukázat na možnosti využití matematických poznatků v kryptografii a podnítit zájem o tuto problematiku.

V dnešní době, aniž si to mnozí z nás uvědomují, se setkáváme se šifrovacími systémy na každém kroku (e-mail, PIN u platebních karet, elektronické podpisy). A do budoucnosti se počítá s ještě širším využitím šifrovacích systémů, které fungují na základě matematických postupů.

Tato práce se snaží inspirovat čtenáře, zvláště pak učitele matematiky a jejich žáky, a přispět tak k propojení učiva matematiky s praxí. Zařazení šifrování a dešifrování zpráv do výuky podporuje žákovu tvořivost a představivost. Hodiny matematiky se tak pro všechny zúčastněné stanou o něco zajímavější a zábavnější.

Literatura

[1] L. Bican: *Algebra (pro učitelské studium)*, Academia, 2004, ISBN 80-20008-60-8

[2] V. Borovička: *Přísně tajné šifry*, Naše vojsko, 1982

[3] O. Grošek, Š. Porubský: *Šifrovanie*, Grada a.s., 1992, ISBN 80-85424-62-2

[4] J. Janeček: *Odhalená tajemství šifrovacích klíčů minulosti*, Naše vojsko, 1994, ISBN 80-206-0462-6

[5] J. Janeček: *Válka šifer*, Votobia, 2001, ISBN 80-7198-505-8

[6] J.A. Menezes: *Handbook of applied cryptography*, CRC Press, 2001, ISBN 0-8493-8523-7

[7] F. Piper, S. Murphy: *Kryptografie*, Dokořán, 2006, ISBN 80-7363-074-5

[8] H. Riesel: *Prime numbers and computer methods for factorization*, Progress in Mathematics volume 123 Birkhäuser Boston, 1994, ISBN 978-0-8176-3743-9

[9] S. Singh: *Kniha kódů a šifer*, Dokořán, 2002, ISBN 80-86569-18-7

[10] P. Tlustý: *Obecná algebra pro učitele*, PF JCU České Budějovice, 2006, ISBN 80-7040-828-6

[11] P. Tlustý, V. Petrášková: *Úvod do počtu pravděpodobnosti*, PF JCU České Budějovice, 1992, ISBN 80-7040-058-7

[12] M. Vincencová: *Matematické metody ochrany dat - šifrování*, PF JCU České Budějovice, 2000, (Diplomová práce)

[13] P. Vondruška: *Kryptologie, šifrování a tajná písma*, Albatros, 2006, ISBN 80-00-01888-8

Použité internetové zdroje:

[14] http://cml.fsv.cvut.cz/ kupca/qc/node23.html

[15] http://crypto-world.info

[16] http://crypto-world.info/index2.php?vyber=soutez

[17] http://cs.wikipedia.org

[18] http://e.math.hr/enigma/index.html

[19] http://hem.passagen.se/tan01/poly.html

[20] http://kevingong.com/Math/PrimeTest.html

[21] http://reboot.cz/info/hacking/jak-funguje-rsa/articles.html?id=281

[22] http://rsa.navajo.cz

[23] http://sifry.sourceforge.net/

[24] http://www-ivs.cs.uni-magdeburg.de/bs/lehre/wise0102/progb/

[25] http://web.math.hr

[26] http://www.almaleh.com/es21.htm

[27] http://www.at-mix.de/rsa.htm

[28] http://www.buslab.org

[29] http://www.kai.vslib.cz/ kolar/rsa.html

[30] http://www.karlin.mff.cuni.cz/ tuma

[31] http://www.mersenne.org/prime.htm

[32] http://www.otr.com/ciphers.html

[33] http://www.root.cz/r/sifrovani

[34] http://www.shaman.cz/sifrovani

[35] http://www.scienceworld.cz

[36] http://www.wakan.cz/sifry/list.php

Kapitola 7

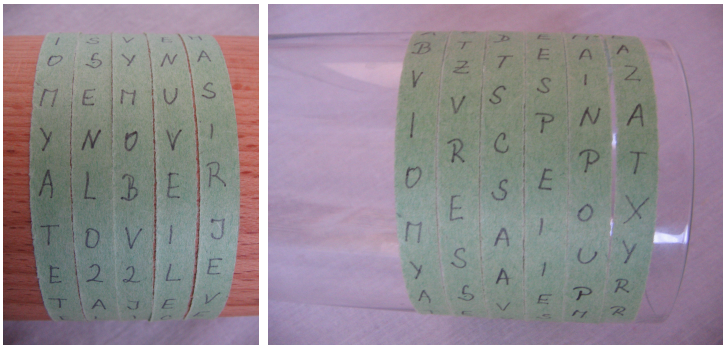
Přílohy

Skytale

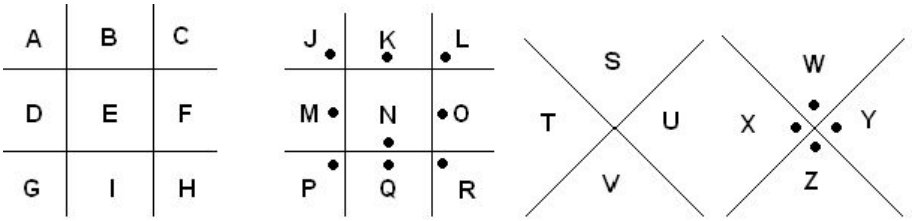
První šifrovací aparát, který sloužil pro vojenské účely byl velice jednoduchý. Scytale používali poslové, kteří nosili tajné zprávy z Persie pro Spartského krále Lysandra. Zpráva se v nečitelné formě ukrývala na opasku, který posel nosil. Odesílatel omotal opasek na dřevěnou tyč a napsal na něj zprávu, tak jak vidíme na obrázku 7.1 a). Když opasek z tyče sundal, byla zpráva nečitelná. Příjemce zprávy vlastnil tyč stejného průměru, poslův opasek na ni namotal a důležitou zprávu si lehce přečetl.

Názorná ukázka této jednoduché formy transpozice:

- Nejprve si tak jako staří Peršané připravíme pomůcky. V našem případě pásku papíru, psací náčiní a místo tyče nám poslouží kuchyňský váleček.
- Dále vybereme text, který chceme utajit před zvědavci a nepřáteli.
- Pásku namotáme kolem válečku, pečlivě přepíšeme zprávu. Poté pásku opět sejmeme.
- Na obrázku 7.1 b) se můžeme přesvědčit, že pokud pásku omotáme okolo válce jiného průměru, stane se nečitelnou.



Obrázek 7.1: a) Váleček b) Sklenice

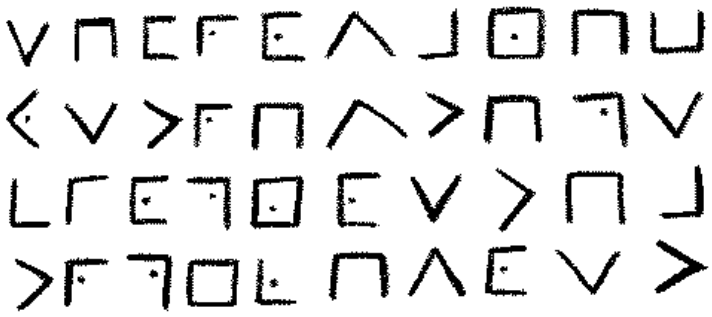


Obrázek 7.2: Šifra prasečích chlívků

Šifra prasečích chlívků

Šifru prasečích chlívků používali svobodní zednáři na počátku 18. století pro uchovávání svých tajných dokumentů. Dnes se s ní baví šikovné děti. Šifra každé písmeno textu nahradí symbolem podle následujícího vzoru na obrázku 7.2.

Pro zašifrování určitého písmene je třeba najít jeho pozici v jedné ze čtyř mřížek a načrtnout část mřížky, která písmeno představuje:
b=⊔
s=∨
Pokud znáte klíč, je snadné šifru prasečích chlívků rozluštit. Pokud ne, je jednoduché ji rozlomit pomocí frekvenční analýzy.



Obrázek 7.3: Ukázka šifry prasečích chlívků

Transpozice plukovníka Roche

Šifrovací systém plukovníka Roche je založen na vpisování otevřeného textu do bloků nestejné délky. Délky těchto bloků jsou $1, 2, 3, \ldots$. Délka celého textu T je pak součtem prvních n členů této posloupnosti a platí:

$$T = \frac{h}{2} \cdot (1 + h),$$

kde h je délka hesla.

Heslo *pondělí* nám stanoví velikost a pořadí jednotlivých bloků. Nyní se pokusíme zašifrovat Latinské přísloví: „*Dílo samo chválí svého tvůrce.*“ Písmenům z hesla přiřadíme čísla podle pořadí v mezinárodní abecedě.

P	O	N	D	Ě	L	Í
7	6	5	1	2	4	3

Tím jsme vytvořili šifrovací klíč a citát budeme vpisovat do bloků, jejichž délka je určena čísly z klíče. (Citát byl doplněn písmeny P, A a V.)

7	6	5	1	2	4	3
D	I	L	O	S	A	M
O	C	H		V	A	L
I	S	V			E	H
O	T	V			U	
R	C	E				
P	A					
V						

Zašifrovaný text čteme po sloupcích a zapisujeme do pětimístných skupin.

DOIOR PVICS TCALH VVEOS VAAEU MLH

Příjemce zprávy, který zná heslo, si spočte klíč a vytvoří bloky příslušné délky. Do bloků zapíše zašifrovaný text směrem shora dolů a po řádcích si přečte původní text.