Pedagogická fakulta
Faculty of Education
Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

Jihočeská univerzita v Českých Budějovicích

Pedagogická fakulta

Katedra matematiky

Bakalářská práce

Vybrané šifrovací metody

Vypracoval: Eva Milichovská

Vedoucí práce: prof. RNDr. Pavel Tlustý, CSc.

České Budějovice 2016

Prohlášení

Prohlašuji, že svoji bakalářskou práci na téma Vybrané metody šifrování jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě, elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích ..... ……

Poděkování

Na tomto místě bych ráda poděkovala prof. RNDr. Pavlu Tlustému, CSc., vedoucímu mé bakalářské práce, za jeho odborné vedení práce, věcné připomínky, dobré rady a vstřícnost při konzultacích.

Anotace

Cílem této práce je uvedení do problematiky klasické kryptologie od jejích počátků ve starověku až do poloviny 20. století. Dále je zde uvedeno rozdělení základních šifer na substituční a transpoziční. Z velkého množství šifer byly vybrány jen ty nejdůležitější a byly zde uvedeny také názorné ukázky šifrování. Následně je zde předveden způsob dešifrování Vigenèrovy „nerozluštitelné“ šifry.

Klíčová slova: kryptologie, substituční šifry, transpoziční šifry, frekvenční analýza, steganografie, kryptoanalýza.

Annotation

The aim of this thesis is introduction into issues of classical cryptology from its beginnings in the antiquity to the half of the 20th century. Further, the division of basic ciphers on substitutional and transpositional ones is stated. Only the most important ciphers were chosen from the large number and demonstrations of decipherment were stated here, too. Then, the way of decipherment of Vigenèr´s „indecipherable“ cipher is showed here.

Key words: cryptology, substitutional ciphers, transpositional ciphers, frequency analysis, steganography, cryptoanalysis.

OBSAH:

1 Úvod

Klasickou kryptologii lze označit jako vědu o utajení obsahu zpráv a také o luštění zpráv. Spadá do ní kryptografie, kryptoanalýza a také steganografie. Kryptologie provází lidskou společnost už od samého vzniku písma, a proto má velmi zajímavou a dlouhou historii.

Dané téma jsem si zvolila především proto, že jsem chtěla propojit své studijní obory, a to matematiku a historii. Z tohoto důvodu bude část práce zaměřena na historii tvoření šifer, tedy na kryptografii, a dále i na dějiny luštění šifer, tedy na kryptoanalýzu.

Ve své práci uvedu stručný výčet několika nejzajímavějších šifer od antiky po polovinu 20. století, neboť do této doby byly šifry výhradně lidskou doménou a ne doménou výpočetní technologie.

Uvedu i matematické definice a věty, které poskytují základní znalosti v oblasti modulární aritmetiky pro pochopení základů šifrování.

Hlavním cílem této práce je popis dešifrování Vigenèrovy šifry bez znalosti klíče. Tento úkol se nepodařilo vyřešit 300 let, tedy až do poloviny 19. století, kdy byl objeven tzv. Kasiského test, který nám umožňuje zjistit délku cyklicky se opakujícího klíče v polyalfabetické šifře.

2 Základní principy šifrování

Tato kapitola bude věnována základní charakteristice šifrování, terminologii, která se při šifrování používá, a základnímu rozdělení šifer, přičemž ukázky různých druhů šifer jsou uvedeny ve třetí části této práce.

2.1 Kryptologie

Všechny následně popsané vědy patří pod vědní obor známý jako kryptologie, což je věda, která se zabývá různými způsoby, jak skrýt komunikaci a následně jak se dostat k podstatě zprávy.

2.1.1 Kryptografie

Slovo kryptografie pochází z řeckých slov kryptós - skrytý a gráphein - psát. Hlavním úkolem kryptografie je zamaskovat zprávu, aniž by nepovolaná osoba byla schopná jí rozumět, tedy vymýšlet matematické metody, které zajistí, že nikdo kromě odesilatele a adresáta se nedostane k podstatě sdělení.

Jako vědní disciplína se kryptografie vyvíjela již od samého vzniku písma, kdy bylo potřeba bezpečně poslat sdělení o vojenském tažení či o intrikách.

2.1.2 Kryptoanalýza

Vědou s opačným cílem než kryptografie je kryptoanalýza, která se zabývá metodami luštění šifer. Tato věda tedy zjišťuje bezpečnost těchto šifer a pravděpodobnost jejich prolomení.

2.1.3 Steganografie

Další důležitou vědou je steganografie, která se zabývá utajením komunikace a ukrýváním zpráv.

Steganografie se vyvíjela především ve starověku, následně se však nepovažovalo

za dostatečně bezpečné zprávu před potencionálními nepřáteli pouze schovat a bylo nutné informace ve zprávě i šifrovat.

Ve starověké Číně k utajení zprávy stačilo, aby byla napsaná, ale pro úplnou ochranu zprávy se text napsal na hedvábný šátek, který se následně smotal do malé kuličky, ta se poté dala jednoduše schovat v šatech či jiných předmětech. Dalším způsobem, jak Číňané schovávali své zprávy, jsou tzv. měsíční koláčky. Jedná se o dutou kuličku z těsta.

Ze starověkého Řecka máme informaci o schování šifry Řekem jménem Demaratus. Tento muž žijící v Persii se dozvěděl o přípravách na válku krále Xerxa[1] s jeho rodným Řeckem. Aby svou domovinu varoval, vyřezal do dřeva zprávu a poté ji zalil voskem.

Důvodů, proč bychom chtěli utajit naši komunikaci, je celá řada, zde jsme uvedli jenom několik základních:

- Důvěrnost - jak už bylo řečeno, chceme zamezit, aby se k obsahu dat někdo dostal.
- Ověření - díky šifrovacímu klíči známému pouze nám je šifrovaná zpráva pro nás důkazem, že pochází od námi očekávané osoby. Jedná se tedy o potvrzení identity odesílatele.
- Integrita - díky neznalosti klíče třetí osobou se zprávou mohl manipulovat pouze původní odesílatel, a tak je zajištěno její původní znění.

2.2 Základní terminologie

Na začátku je nutné definovala pojmy, které se zde budou velmi často objevovat, tedy základní terminologii používanou při šifrování.

Informace, které chceme zabezpečit před nechtěným přečtením, se většinou označují jako *otevřený text.* V této práci budeme pracovat pouze s českým textem a veškerý otevřený text bude psán bez použití diakritiky. Taky české písmeno

[1] Xerxes - perský král, 550-486 př. n. l.

ch budeme brát jako složeninu písmen c a h a ne jako samostatné písmeno. Abychom tento otevřený text zabezpečili a mohli ho odeslat bez obav, že si někdo neoprávněný zprávu přečte, použijeme proces, při kterém zakryjeme podstatu pravé informace. Tento proces se nazývá *šifrování*. Takto vzniklý text nazýváme *šifrovaný text* nebo *kryptogram*. Sada pravidel, které použijeme k vytvoření kryptogramu, se pak nazývá *šifrovací algoritmus*. Ten ve spojení s *šifrovacím klíčem*, který lze definovat jako určitý parametr šifrovacího algoritmu, jež ovlivní podobu výsledného šifrovaného textu, převede otevřený text na kryptogram.

Pokud jako příjemce chceme zprávu rozluštit a dostat se tak k otevřenému textu, musíme použít *dešifrovací algoritmus* ve spojení s *dešifrovacím klíčem*, který nám převede kryptogram na otevřený text.

2.3 Definice šifrování

Pro definování šifrování potřebujeme zavést definici abecedy, která nám umožní manipulaci s písmeny.

Budeme tedy využívat značení

$$Z_N = \{1, 2, \ldots, N - 1\}.$$

Pod textem nad abecedou Z_N budeme tedy rozumět libovolnou n-tici

$$x = (x_1, x_2, \ldots, x_n), \text{ kde } x \in Z_N \text{ pro } 1 \leq i \leq n.$$

Tuto základní abecedu Z_N a množinu všech n-tic nad Z_N budeme značit jako $Z_{N,n}$.[2]

Pod šifrováním, nebo tedy kryptografickou transformací, budeme rozumět libovolné prosté zobrazení

$$T : Z_{N,n} \rightarrow Z_{N,n}.$$

Pod šifrovacím systémem budeme potom rozumět systém

$$T^K = \{T_k : k \in K\},$$

kryptografických transformací T_k, které jsou specifikované parametrem

[2] Grošek, 1992, s. 47.

k . Parametr k budeme nazývat *klíčem* a množinu K *prostorem klíčů*.

V kryptografickém systému T^K je tedy každá n-tice

$$(x_1, x_2, \ldots, x_n) \in Z_{N,n}$$

znaků $x_i \in Z_N$, $1 \leq i \leq n$ přímého textu přetransformovaná použitím zobrazení $T_k \in T^K$ na n-tici

$$(y_1, y_2, \ldots, y_n) \in Z_{N,n},$$

kde $y_i \in Z_N$, $1 \leq i \leq n$, přičemž

$$(y_1, y_2, \ldots, y_n) = T_k\,(x_1, x_2, \ldots, x_n).$$ [3]

2.4 Definice dešifrování

Stejně jako jsme definovali šifrování, musíme taky definovat jeho opačný proces, a to dešifrování.

Ke každému prostému zobrazení

$$T_k : Z_{N,n} \rightarrow Z_{N,n}$$

existuje inverzní zobrazení

$$(T_k)^{-1} : Z_{N,n} \rightarrow Z_{N,n}$$

takové, že pokud je n-tice $(x_1, x_2, \ldots, x_n) \in Z_{N,n}$ znaků $x_i \in Z_N$, $1 \leq i \leq n$ přímého textu zašifrovaná na n-tici

$$(y_1, y_2, \ldots, y_n) \in Z_{N,n}$$

kde $y_i \in Z_N$, $1 \leq i \leq n$, tak

$$(y_1, y_2, \ldots, y_n) = (T_k)^{-1}\,(x_1, x_2, \ldots, x_n).$$ [4]

Dešifrování je tedy hledání inverzní funkce $(T_k)^{-1}$.

[3] Grošek, 1992, s. 49.
[4] Tamtéž, s. 78.

2.5 Typy klíčů

Jak jsme již zmínili, klíč je jistý parametr šifrování, jenž nám ovlivní šifrovaný text. Můžeme využít jednotný způsob šifrování, ale při použití jiného klíče nám výsledné kryptogramy vyjdou rozdílné.

Mezi základní typy klíčových, šifrovacích algoritmů patří symetrický a asymetrický klíč.

2.5.1 Symetrický (soukromý) klíč

Symetrický klíč je v podstatě algoritmus, kde šifrovací klíč může být vypočítaný z dešifrovacího klíče. Abychom zajistili, že naše korespondence bude bezpečná, je nutné, abychom uchránili šifrovací klíč před potencionálními narušiteli.

2.5.2 Asymetrický (veřejný) klíč

Asymetrický klíč je typ algoritmu, který se liší od symetrického klíče v tom, že šifrovací a dešifrovací klíče nejsou totožné. Šifrovací klíč je veřejný, třetí osoba se o něm tedy dozví, ale bohužel kvůli neznalosti dešifrovacího klíče, který se nazývá soukromý klíč, nemůže zprávu rozluštit. Tento klíč se používá především pro veřejné počítačové sítě. Šifrování pomocí veřejného klíče je dozajista mnohem bezpečnější způsob utajení informací než využití symetrického klíče, neboť odstraňuje jednu z nejnebezpečnějších částí šifrování, a to je doprava klíče od odesílatele k příjemci.

Dalším problémem při použití veřejného klíče je i samotný počet klíčů. Pro síť obsahující *n* uživatelů by se celkový počet klíčů rovnal $\binom{n}{2}$. Například pro síť, která má 10 uživatelů, bychom potřebovali 45 klíčů, v případě sítě se 100 uživateli počet klíčů vzroste na 4950.

V následující části je popsáno základní rozdělení šifrovacích systémů, a to na substituční a transpoziční systém.

2.6 Substituční systém

Substituce se v kryptografii používá asi nejčastěji. Díky obrovskému množství různých možností se jedná o mnohem bezpečnější cestu, jak skrýt informaci, než systém transpoziční či obyčejná steganografie. Při využití substitučního systému se každý znak v otevřeném textu nahradí jiným, a tak vznikne šifrovaný text.

Substituční šifry mají čtyři základní typy.

2.6.1 Monoalfabetická šifra

Monoalfabetická šifra, u tohoto druhu šifer je po celou dobu šifrování pevně daný klíč, který se během šifrování nemění. Můžeme to zapsat ve tvaru $f(a) = (a + k) \bmod n$, kde n je velikost abecedy, a je prvek otevřeného textu, k je hodnota posunu abecedy. K prolomení této šifry se využívá frekvenční analýzy jazyka, která zjišťuje častost písmen v daném jazyce.

Pokud má tedy abeceda n znaků, získáváme tak $n!$ možných šifrovaných abeced. Pro naši upravenou českou abecedu s 26 znaky se tedy jedná o 26! různých možností jak šifrovat text, což je přibližně $4 \cdot 10^{26}$ možností. Pokud by byl počítač schopen provést miliardu výpočtů za vteřinu, zabralo by mu to přibližně 12,8 miliard let.

2.6.2 Homofonní šifra

Homofonní šifra, při použití této šifry se každý znak otevřeného textu může nahradit různými znaky. V praxi to znamená, že písmeno s největší četností v daném jazyce se zašifruje nejvíce znaky. V angličtině je nejčastější písmeno E, které představuje přibližně 10 % textu, tak se toto písmeno zašifruje celkově deseti znaky. V českém jazyce je také nejčastější písmeno E a i jeho procentuální zastoupení je stejné jako v angličtině, a to 10 % textu.

2.6.3 Polygrafická šifra

Polygrafická šifra využívá jak homofonní, tak i monoalfabetickou šifru. Při využití této šifry se skupina znaků zašifruje jako další skupina znaků. Při využití této

šifry je nutné znát pojem bigram, tedy dvojice písmen, trigram tedy trojice písmen atd. Mezi polygrafické šifry řadíme především šifru Playfair a Hillovu šifru.

2.6.4 Polyalfabetická šifra

Polyalfabetická šifra, v průběhu šifrování pomocí této metody se klíč podle daných pravidel mění a ztěžuje prolomení šifry. Každý znak otevřeného textu je tedy zašifrovaný podle vlastního klíče. Jednou z nejznámějších šifer pracujících na tomto principu je tzv. Vigenèrova šifra, o které se důkladněji zmíním ve čtvrté kapitole.

2.6.5 Index koincidence

Při luštění nám neznámého textu je jednou z prvních úloh poznat, jakým druhem šifry byl daný text zašifrován. Pro zjištění, jestli byl využit monoalfabetický či polyalfabetický substituční systém, nám poslouží tzv. index koincidence, dále značený už jen jako IC. Tento test vymyslel William Frederick Friedman[5], který působil jako kryptolog pro americkou armádu za druhé světové války. Jeho teorie se poprvé objevila v knize *The index of coincidence and its applications in cryptographic analysis.*

Tento výpočet pracuje s relativní četností písmen v šifrovaném textu, který následně porovnáme s relativní četností písmen běžného textu. Výpočet můžeme napsat následovně:

$$IC = \sum_{i=a}^{z} \frac{f_i \cdot (f_i - 1)}{n \cdot (n - 1)}$$

Přičemž n je počet všech znaků ve zprávě a f_i je počet písmen jednoho druhu, následně se hodnoty budou sčítat přes všechna písmena abecedy, tedy od a do z. Pokud se IC výrazně liší od IC jazyka, ve kterém by podle nás měla být zpráva napsaná, bude se jednat o polyalfabetickou šifru, jež mění výrazně tuto hodnotu. Čím je větší odchylka od IC jazyka, ve kterém je zpráva podle předpokladu napsána, tím větší můžeme předpokládat délku klíče.

[5] W. F. Friedman - americký kryptolog, 1891-1969

Jazyk	IC
Čeština	0,06027
Angličtina	0,06689
Dánština	0,07073
Finština	0,07380
Francouzština	0,07460
Holandština	0,07981
Němčina	0,07667
Italština	0,07329
Ruština	0,05607
Španělština	0,07661

Tabulka 1. Index koincidence evropských jazyků[6]

2.7 Transpoziční systém

Při použití transpoziční šifry se písmena nemění na jiná, jako tomu bylo u substituce. V tomto případě se mění pouze pořadí písmen v otevřeném textu. Ke správnému zašifrování se může používat určitý druh geometrického útvaru. Pro znázornění dané šifry použijeme slovo JAZYKOLAM. Pokud toto slovo přepíšeme po sloupcích do tabulky 3×3, vznikne:

J	Y	L
A	K	A
Z	O	M

Tabulka 2. Transpoziční systém v praxi

Pokud následně přepíšeme řádky, vznikne kryptogram JYLAKAZOM. Pro dešifrování je tedy nutné použít stejný postup jako pro šifrování, a to přepsat šifrovaný text do tabulky nám známé podoby.

2.8 Frekvenční analýza

Jedním z řešení jednodušších substitučních šifer mimo homofonní šifry je frekvenční analýza. Tato lingvistická disciplína nám ukazuje, jak často se v daném jazyku objevují jednotlivá písmena, slova atd. Mezi její zakladatele patří nejspíše

[6] *Katedra inženýrské informatiky* [online]. [cit. 2016-04-13]. Dostupné z: http://kix.fsv.cvut.cz/~vanicek/vyuka_l01/kos3.htm.

Arabové, kteří již od 7. století našeho letopočtu byli na takové úrovni v matematice a lingvistice, že mohli pracovat na podobné vědecké disciplíně. Její využití je podmíněno tím, že známe jazyk, ve kterém je otevřený text napsán.

Poznatky z této vědní disciplíny budeme využívat v kapitole o řešení Vigenèrovy šifry.

Obrázek 1. Frekvenční analýza českého jazyka[7]

Tato tabulka nám ukazuje frekvenční analýzu obecné češtiny, tedy češtiny jak odborné, administrativní, tak umělecké. Ovšem v každém z těchto stylů je frekvence trochu jiná. V přílohách je uvedena tabulka srovnávající obecnou a odbornou češtinu. Odchylky jsou téměř zanedbatelné, rozdíly se pohybují v desetinách procent.

[7] KRÁLÍK, Jan. Statistika českých grafémů s využitím moderní výpočetní techniky, Slovo a slovesnost XLIV, 1983, s. 295-304.

3 Základy modulární algebry

Jak jsme poznamenali na začátku této práce, metody, které kryptologie využívá, jsou především matematické. Proto bych chtěla v této kapitole definovat základy z teorie čísel a algebry.

Teorie čísel je část matematiky, jejímž základním objektem zkoumání jsou vlastnosti přirozených a celých čísel. *Přirozená čísla* jsou tedy prvky množiny

$$N = \{1, 2, 3, \ldots\}$$

a *celá čísla* jsou prvky množiny

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

V rozšířeném smyslu slova lze algebru definovat jako část matematiky, která se zabývá vyšetřováním vlastností množin, jejich prvků a manipulace s nimi.

Šifrovací metody, které v této práci budu popisovat, využívají především dělitelnost. Proto bych zde uvedla definice dělitelnosti, definice kongruence a její základní vlastnosti.

Definice 2.1.

Nechť $a, b \in Z$. Řekneme, že *a dělí b*, značíme $a \mid b$, jestliže existuje $k \in Z$ tak, že $b = a \cdot k$. V opačném případě říkáme, že *a nedělí b* a píšeme $a \nmid b$. Jestliže $a \mid b$ a současně $b \mid a$, říkáme, že čísla *a, b* jsou *asociovaná* a píšeme $a \parallel b$; pokud *a, b nejsou asociovaná*, 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 $a \nparallel 1$ a $a \nparallel b$, řekneme, že *a je nevlastním dělitelem čísla b*.

Pokud jsme definovali dělitelnost, je nutno zde uvést i definici prvočísla.

Definice 2.2.

Přirozené číslo p, které má právě dva různé dělitele v množině přirozených čísel, a to číslo 1 a samo sebe, se nazývá *prvočíslo*. V opačném případě se p nazývá *složené číslo.*

Význam prvočísel podtrhuje tzv. „Základní věta aritmetiky“.

Věta 2.3.

Každé přirozené číslo lze rozložit v součin konečného počtu prvočísel, a to až na pořadí jednoznačně.

Definice 2.4.

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

S pojmem největší společný dělitel souvisí i nejmenší společný násobek.

Definice 2.5.

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

V typech šifrování, o kterém tato práce hovoří, se bude především využívat zbytek po dělení tzv. *modulo.*

Definice 2.6.

Nechť $m \in Z$. Jestliže $m \mid (a - b)$, říkáme, že *a je kongruentní s b* podle modulu m a píšeme

$$a \equiv b \bmod m.$$

Kongruence má mnoho užitečných vlastností. Uvádím zde jen ty, které jsou pro šifrování nejdůležitější.

Věta 2.7.

Nechť $a, b, c \in Z, m \in N$. Pak platí:

1. $a \equiv a \bmod m$,
2. jestliže $a \equiv b \bmod m$, pak také $b \equiv a \bmod m$,
3. pokud $a \equiv b \bmod m$ a zároveň $b \equiv c \bmod m$, pak také $a \equiv c \bmod m$.

Věta 2.8.

Nechť $a, b, c \in Z, m \in N$. Pokud $a \equiv b \bmod m$, pak

1. $(a + c) \equiv (b + c) \bmod m$,
2. $(a - c) \equiv (b - c) \bmod m$,
3. $ac \equiv bc \bmod m$.

Věta 2.9.

Nechť $a, b, c, d \in Z, m \in N$. Pokud $a \equiv b \bmod m$ a $c \equiv d \bmod m$, pak platí:

1. $(a + c) \equiv (b + d) \bmod m$,

2. $(a - c) \equiv (b - d) \bmod m$,

3. $ac \equiv bd \bmod m$.

Věta 2.10.

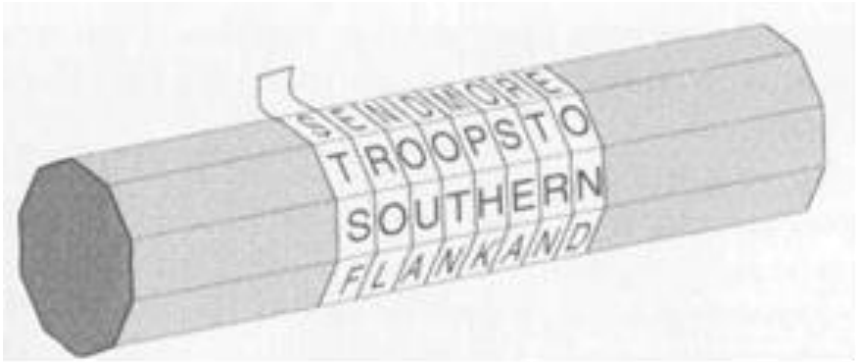
Nechť $a, b, c \in Z, m \in N$. Nechť $D(c, m) = 1$, Pokud $ac \equiv bc \bmod m$, pak také $a \equiv b \bmod m$.

4 Historie kryptografie

Tato kapitola pojednává jen o nejzákladnějších a nejzajímavějších šifrách, které kdy byly popsány. Jedná se především o šifry z doby před objevením počítačů, tedy o šifry, kdy pro šifrování a dešifrování byl použit jen lidský um a mozek.

4.1 Skytale

Jednou z prvních nám známých šifer byla šifra používaná Sparťany v 7. století před naším letopočtem. O této šifře víme díky řeckému filosofovi a historikovi Plútarchovi[8]. Doslova to označovalo palici, kterou měli všichni spartští stratégové u sebe při svých vojenských výpravách. Zpráva byla napsaná na kusu kůže a omotaná kolem této palice. Zpráva byla čitelná pouze tehdy, pokud měla palice správný poloměr a písmena na pásku kůže se seskupila do srozumitelného textu. Šifra *skytale* je tedy jednoduchá transpoziční šifra.



Obrázek 2. Spartská šifra skytale[9]

4.2 Caesarova šifra

Jedna z nejznámějších starověkých šifer se nazývá *Caesarova*, podle římského diktátora Julia Caesara[10]. O této šifře se dozvídáme z knihy *Zápisky o válce galské* od

[8] Plútarchos - asi 40-125 n. l.
[9] Singh, 2003, s. 16.
[10] Gaius Julius Caesar - 100-44 př. n. l.

samotného Caesara a také od Gaia S. T. Suetonia[11] z jeho díla *De vita caesarum* (český překlad *Životopisy dvanácti císařů*). Pro své vojenské účely Julius Caesar používal jednoduchou substituční šifru, kterou bychom mohli definovat následovně:

Jedná se o kryptografický systém $Z_{N,n}$, který lze zapsat ve tvaru

$$T_k : Z_{N,n} \to Z_{N,n},\ T_k(n) \equiv (n + k) \bmod N,$$

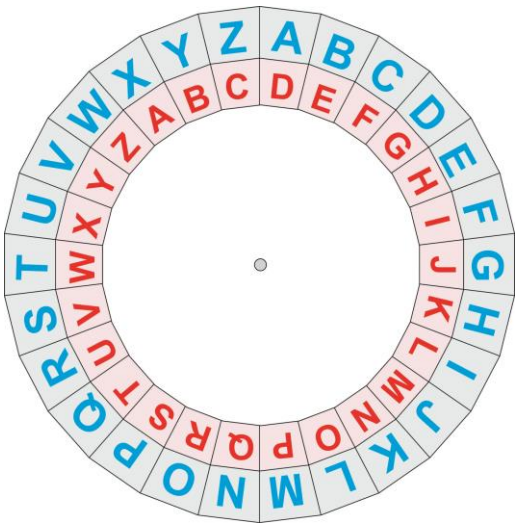
kde je k libovolné číslo a N velikost abecedy.[12]

Jedná se tady o posun písmen v abecedě o předem daný počet. Caesar standardně posouval písmena v abecedě o 3. Jeho šifru lze tedy přesně matematicky definovat jako:

$$f(n) \equiv (n + 3) \bmod 26$$

kde $f(n)$ je šifrovaný text a n je text otevřený.

Pro ukázku šifrování tímto algoritmem použijeme slovo JABLKO jako náš otevřený text. Každému písmenu otevřeného textu tedy přiřadíme jeho číselný ekvivalent 9-0-1-11-10-14. Pro přičtení $k = 3$ a výpočtu modulo 26 dostaneme čísla 12-3-4-14-13-17 a těmto číslům odpovídají písmena M D E O N R.



Obrázek 3. Kotouč s posunem písmen

[11] Suetonius – římský historik, 69-140 n. l.
[12] Grošek, 1992, s. 95.

Kvůli zdlouhavému procesu výpočtu každého písmena používal Caesar dva kotouče přeložené přes sebe. Na každém byla vypsaná po obvodu abeceda. Otáčením do správné polohy tak získal šifrovanou abecedu.

Ve své době byla Caesarova šifra v podstatě nerozluštitelná, až do té doby, než celý postup šifrování vyzradil Caesarovým nepřátelům Cicero[13]. Tuto šifru využíval i Gaius Octavianus Augustus[14], on ovšem posouval pouze o hodnotu $k = 1$.

Obdoba Caesarovy šifry se objevuje i před Caesarem, a to přesně ve Starém zákoně. Jedná se hebrejskou šifru *atbaš*. Při šifrování touto metodou se první písmenko abecedy zašifruje jako poslední, druhé písmeno jako předposlední atd. Název této šifry je odvozen od hebrejských písmen.

4.3 Afinní šifra

O něco složitější než Caesarova šifra je *afinní šifra*. Tato šifra vychází stejně jako Caesarova šifra z posunu písmen v abecedě o předem známou hodnotu. Ale do procesu výpočtu kryptografického znaku je přidán prvek a, kterým se násobí prvek otevřené abecedy n. Celý proces tedy můžeme definovat jako:

Kryptografický systém nad abecedou $Z_{N,n}$, který lze napsat ve tvaru:

$$\mathrm{T}_{a,k} : Z_{N,n} \rightarrow Z_{N,n},\ T_{a,k}(n) \equiv (a \cdot n + k) \bmod N,$$

kde a, k jsou celá čísla taková, že a je nesoudělné s N.[15]

Pro příklad vezmeme slovo DVEŘE a jeho podoba pro šifrování bude DVERE, díky naší úmluvě, že nebudeme používat diakritiku. Pokud každému písmenu přiřadíme jeho číselný ekvivalent, vyjdou nám čísla 3-21-4-17-4. Zvolíme $a = 5$ a $k = 3$ stejně jako u Caesarovy šifry. Po vynásobení prvkem a, přičtení prvku k, výpočtem modulo získáváme číslice 18-4-23-10-23. Těmto číslům poté odpovídají písmena S E X K X.

[13] Marcus Tullius Cicero - římský politik, právník a řečník, 106-43 př. n. l.
[14] Gaius Octavianus Augustus - první římský císař, 63 př. n. l.-14 n. l.
[15] Grošek, 1992. s 103.

4.4 Polybiův čtverec

Mezi další antickou šifrovací metodu patří i *Polybiův*[16] *čtverec*. Jedná se opět o jednoduchou substituční metodu, která využívá tabulku 6×6, kdy se písmena abecedy vypíší do této tabulky, a to po řádcích. Naše abeceda ovšem obsahuje 26 znaků, a proto se pro tento způsob šifrování využívá taková abeceda, ve které jsou písmena J a I totožná. Sloupce i řádky očíslujeme.

Tabulka tedy vypadá následovně:

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Tabulka 3. Polybiův čtverec

Každému písmenu odpovídají dvě číslice, a to číslo řádku a následně sloupce. Slovo POLYBIUS tedy zašifrujeme jako sled čísel 35-34-31-54-12-24-45-43.

Tato šifra se využívala i jako tzv. telegrafický kód. Doručování této zprávy tak mohlo probíhat i na dálku, aniž by se musela zpráva fyzicky poslat. Poslání takové zprávy probíhalo za pomoci dvou panelů s vlajkami. Čísla řádků a sloupců se určovala z počtu vlajek na panelech.

4.5 Vigenèrova šifra

Tato šifra byla vytvořena francouzským diplomatem Blaisem de Vigenèrem[17] v 16. století. Jeho metoda byla po zhruba 300 let velmi hojně používaná, neboť se o ní tvrdilo, že je nerozluštitelná.

[16] Polybios - helenistický historik, 210-120 př. n. l.
[17] Blaise de Vigenère - 1523-1596

Jedná se o složitou polyalfabetickou substituční šifru. Pokud je klíč stejně dlouhý jako otevřený text, je nemožné tuto šifru vyluštit, neboť v tomto případě nejde využít tzv. Kasiského test, který se pro luštění polyalfabetické šifry využívá.

Přiřaďme stejně jako u Caesarovy šifry každému písmenku abecedy číslo od 0 do 25, tedy A = 0,…, Z = 25. Jako klíč použijme slovo nebo i celou větu, na které jsme se domluvili předem s adresátem. Každé písmenko v otevřeném textu zašifrujeme podle daného písmenka v klíči, který se neustále opakuje.

Například pokud chceme zašifrovat zprávu SEDM a použijeme klíč PES, platí stejný princip jako u Caesarovy šifry, avšak posun nebude o jednotnou hodnotu *k*, ale o hodnotu *k*, která se rovná odpovídajícímu písmenku v klíči. Tedy první písmeno zašifrujeme jako $f(n) \equiv (18 + 15) \bmod 26$, kde se $f(n) = 7$, tedy písmeno H. Celá zpráva tedy bude znít HIVB.

Šifrování i následné dešifrování je zdlouhavější, a proto se kvůli zjednodušení procesu používá tzv. *Vigenèrův čtverec*.

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

Tabulka 4. Vigenèrův čtverec

4.6 Šifra Marie Stuartovny

Abychom zde uvedli i substituční šifru, která nevyužívá latinku jako šifrovací abecedu, uvedeme i velmi známou šifru skotské královny Marie Stuartovny[18]. Za spiknutím, které nakonec stálo Marii hlavu, stála skupinka katolických šlechticů, která chtěla svrhnout protestantku Alžbětu[19] a dosadit na anglický trůn katoličku Marii. Korespondence, kterou Marie s těmito muži vedla, byla složená ze symbolů, jež nezasvěcenému člověku připadaly nesmyslné. Na anglickém dvoře ovšem působil Thomas Phelippes[20], který dokázal tyto symboly rozluštit. Marie díky tomu stanula před soudem a byla obviněna z velezrady a nakonec popravena. V přílohách je přiložena

[18] Marie Stuartovna - 1542- 1587
[19] Alžběta - anglická královna, 1533-1603
[20] Thomas Phelippes - 1556-1625

ukázka, jak vypadala Mariina šifrovací abeceda.

4.7 Šifra Playfair

Šifra Playfair je polygrafická substituční šifra pojmenovaná podle anglického vědce jménem Lyon Playfair[21]. Šifra však byla ve skutečnosti vytvořena v roce 1854 Playfairovým přítelem Charlesem Wheatstonem[22], významným britským fyzikem, a byla využívána Brity během válek proti Búrům[23] a také během první světové války.

Šifra Playfair je založena na substituci bigramů, tedy dvou znaků v otevřeném textu za sebou. Klíčem je jedno slovo, ve kterém se neopakují písmena. Toto slovo se vepíše do tabulky 5×5 po řádcích. Stejně jako u Polybiova čtverce písmena J a I jsou sjednocena. Do zbývajících polí v tabulce se vypíší zbývající písmena z abecedy. Pokud jednu dvojici v bigramu tvoří dvě stejná písmenka, vložíme mezi ně X, pokud zpráva obsahuje lichý počet znaků, vložíme nakonec opět písmenko X.

Otevřený text následně budeme šifrovat podle tří následně uvedených pravidel.

1. Pokud se obě písmena nacházejí v tabulce na stejném řádku, každé se nahradí písmenem ležícím o jedno napravo. Pokud je písmeno úplně vpravo, nahradí se prvním písmenem na tomto řádku.
2. Pokud se obě písmena nacházejí v tabulce ve stejném sloupci, každé se nahradí písmenem ležícím o jedno dolů. Pokud je písmeno úplně dole, nahradí se prvním písmenem v tomto sloupci.
3. Pokud se písmena nacházejí v různých řádcích i sloupcích, pak se každé nahradí písmenem, které leží na stejném řádku, avšak ve sloupci, kde leží druhé písmenko z dvojice.

Pro lepší ujasnění si zašifrujeme slovo RHODODENDRON. Slovo rozdělíme na části o délce dvou hlásek, takže RH OD OD EN DR ON. Jako klíč použijeme slovo

[21] Lyon Playfair - 1818-1898
[22] Charles Wheatstone - 1802-1874
[23] Búrské války - osvobozenecké války jihoafrických osadníků na Britském impériu, 1899-1902

PETRKLIC. Vznikne nám tedy šifrovací tabulka:

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

Tabulka 5. Šifra Playfair

Podle pravidel výše vypsaných zašifrujeme otevřený text a vznikne kryptogram ve znění AS NF NF PO HP QO.

Podobnou bigramovou substituci jako je Playfair využívali i Němci, a to v průběhu druhé světové války. U této šifry však nevyužívali jednu, ale hned dvě šifrovací tabulky. Tyto tabulky byly tištěny na každý den jiné, a proto nemusel být klíč srozumitelné slovo. Pro každý měsíc byly vydávány šifrovací brožury s předem danými tabulkami. U každé této tabulky bylo připsáno i číslo *n*, které označovalo počet písmen, která se zapisovala do jednoho řádku zprávy. Způsob vybírání bigramů se liší od Playfair tak, že nevybíráme písmena po sobě následující, nýbrž pod sebou. Následně se první písmeno vyhledá v první tabulce a druhé písmeno v druhé tabulce. Princip šifrování je stejný jako u Playifair, pokud jsou písmena na stejném řádku, posuneme o jedno písmeno, ale první znak z bigramu se zašifruje podle druhé tabulky a druhé písmeno podle první tabulky. Stejné je to i v případě, že se písmena nacházejí ve stejném sloupci. Pokud leží písmena v různých sloupcích a řádcích, vytvoříme stejně jako u Playfair obdélník, jehož rohy jsou pro nás oba známé znaky, zbývající rohy odpovídají kryptogamu, opět se ale budou používat ty rohy, které leží v druhé tabulce, než je písmeno, které šifrujeme.

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

Tabulka 6. Německá šifrovací tabulka

Pro lepší ilustraci budeme šifrovat slovo MINULOST. Rozdělíme ho do řádků po $n = 4$.

M I N U

L O S T

A následně budeme postupovat podle pokynů, které jsme uvedli výše. Jako první tedy zašifrujeme písmena ML. Obě se nacházejí na stejném řádku, a tak písmeno M zašifrujeme jako následující písmeno v druhé tabulce, tedy V, a L zašifrujeme jako R. Následně provedeme šifrování u ostatních bigramů a vyjde nám kryptogram VR VC KY QE.

4.8 Hillova šifra

Následující šifru publikoval v roce 1929 Lester S. Hill[24]. Šifra byla používána americkým námořnictvem během druhé světové války.

Hillova šifra je polygrafická šifra, tedy šifra, kdy se šifrovaní provádí po skupinách *n* znaků. Princip šifrování je založen na násobení matic.

Otevřený text si rozdělíme do bloku o *n* prvcích, každému písmenu přiřadíme jeho číselný ekvivalent a zapíšeme je ve formě vektoru. Klíčem pro zašifrování i následnou dešifraci bude matice A stupně *n*. Blok zprávy zašifrujeme tak, že vezmeme jeho vektor a ten vynásobíme maticí A. Každý prvek výsledné matice modulo 26 nám určí kryptogram.

Pro příklad zašifrujeme slovo MATICE. Slovo si rozdělíme do dvou bloků o délce $n = 3$ a klíčem pro nás bude matice v tomto tvaru:

[24] Lester S. Hill - 1891-1961.

Matice A:

$$\begin{pmatrix} 2 & 4 & 1 \\ 5 & 7 & 3 \\ 2 & -3 & -2 \end{pmatrix}$$

Prvnímu bloku písmen odpovídají čísla (12, 0, 19), po vynásobení matic vzniknou čísla (43, 117, -14). K zápornému číslu připočteme 26. Výsledná čísla jsou tedy (17, 13, 12), čemuž odpovídají písmena RNM. Druhý blok zašifrujeme totožně, vyjdou nám tedy písmena COC.

Náš otevřený text MATICE zašifrujeme tedy jako RNM COC.

5 Vývoj kryptoanalýzy

Stejně důležité jako vytváření šifer je i jejich řešení. Tím se zabývá věda, která se nazývá kryptoanalýza.

Za otce evropské kryptologie je považován Leon Battista Alberti[25], italský renesanční architekt, malíř, skladatel a filozof, který se začal věnovat tomuto vědnímu oboru na popud papežského sekretáře. Jako první v západní Evropě publikoval v roce 1466 práci o kryptoanalýze. Jako první si uvědomil, že frekvenční analýzu lze využít při luštění monoalfabetických šifer. Díky jeho poznatku tak mohly vzniknout šifry polyalfabetické. Přestože byla jeho myšlenka revoluční, již v ní nepokračoval. A tak výrazné objevy na poli polyalfabetických šifer vznikly až díky Němci Johannesi Trithemiovi[26], Italovi Giovannimu Portovi[27] a samozřejmě díky Blaisovi de Vigenèrovi[28].

Jedním z nejznámějších kryptoanalatiků 16. století je známý francouzský právník a matematik Francois Viète[29]. V roce 1589 byl Viète poradcem hugenotského krále Jindřicha IV.[30], který vedl krvavé boje ve Francii s katolíky. Viètovi se podařilo nejen rozluštit šifru, kterou katolíci používali ve svém boji se svými hugenotskými protivníky, ale i další šifrovací systémy, které používali např. Španělé a Belgičané. Díky tomu se z něj stává jeden z prvních významných matematiků ve službách kryptologie.

Dalším pozoruhodným Francouzem je Antoine Rossignol[31], který naopak napomáhal katolíkům v letech 1628 rozluštit zprávu hugenotských vojáků z městečka Realmonte, které bylo obleženo. Díky práci Rossignola, který byl známý tím, že se ve volném čase věnuje luštění šifer, byla hugenotská zpráva rychle rozluštěna a poslána zpátky do města. Realmonte se poté vzdal katolickému vojsku.

Velmi pozoruhodný je příběh francouzského kryptoanalytika Ètienna Bazeriese[32].

[25] Leon Battista Alberti - 1404-1742
[26] Johannes Trithemius - 1462-1516
[27] Giambattista della Porta - 1535-1615
[28] Blaise de Vigenère - 1523-1596
[29] Francois Viète - 1540-1603
[30] Jindřich IV. - 1553-1610
[31] Antoine Rossignol - 1600-1682
[32] Ètienne Bazeries - 1846-1931

Jeho přítel a vojenský historik objevil během výzkumu v archivu velmi zajímavý šifrovaný spis francouzského krále Ludvíka XIV.[33] Autorem šifry byl již nám známý Rossignol, který se díky svým znalostem dostal až ke královskému dvoru a pro krále vytvořil tzv. Velkou šifru.

Při studiu zašifrovaných materiálů Bazeriese zjistil, že je text složen z tisíců čísel, díky různorodosti čísel bylo jasné, že se nejedná o monoalfabetickou substituci, neboť čísel bylo více, než je 26 písmen abecedy. Jeho výzkum, zda se jedná o homofonní substituci, nebyl úspěšný.

Po vyluštění této šifry se Bazeriesovi dostaly do rukou důkazy o aféře tzv. Muže se železnou maskou. Podle předchozích domněnek se jednalo o dvojče francouzského krále, který ho uvěznil a zakryl jeho tvář, aby si nikdy nemohl nárokovat trůn. Podle důkazů byl tímto vězněm francouzský vojenský velitel, který byl kvůli neuposlechnutí rozkazu a uprchnutí z bojiště potrestán tímto krutým způsobem.

5.1 Kryptoanalýza Vigenèrovy šifry

Vigenèrova šifra byla, jak jsme již poznamenali, považována za nerozluštitelnou, a to po dobu 300 let. Za její rozluštění můžeme poděkovat jednomu z nejlepších kryptoanalytiků 19. století. Tímto mužem je Charles Babbage[34].

Charles byl dozajista všestranně nadán, zajímal se jak o techniku, tak o botaniku a matematiku. Zlomovým se pro jeho práci stal okamžik, kdy v roce 1821 našel v tabulkách, jež se používaly pro astronomické a navigační výpočty, velké množství chyb. Mnohé ztroskotání lodě se přičítalo právě těmto chybám v navigačních materiálech. Tyto tabulky se psaly ručně, a tak chyby v nich obsažené udělal člověk. Charles se tedy chtěl vyvarovat dalších chyb způsobených selháním lidského faktoru a chtěl vytvořit stroj, jenž by dokázal tyto tabulky sestavit bez chyb. Díky finanční pomoci britské vlády tak mohl již v roce 1823 navrhnout svůj tzv. Difference engine no.1, kalkulátor. I když byl jeho vynález geniální, nikdy ho Babbage nedokázal zprovoznit. Po deseti letech se pokusil o sestrojení vylepšeného prototypu Difference engine no.2, kvůli ukončení dotování jeho práce ze strany britské vlády nikdy

[33] Ludvík XIV. - 1643-1715
[34] Charles Babbage - 1791-1871

nedokončil ani druhý prototyp. Jenom náklady na jeho první stroj přesáhly 17 000 liber, za něž mohla vláda postavit dvě bitevní lodě. I přes to, že svůj stroj nedokázal nikdy sestrojit, byla jeho práce využita za druhé světové války, kdy byl prototyp jeho přístroje využit při kryptoanalýze nepřátelských zpráv.

Jeho největším počinem v oblasti kryptoanalýzy je ale bezesporu rozluštění Vigenèrovy šifry. Jeho záliba v kryptoanalýze se projevila již v dětství. Jako mladý hoch dokázal již po přečtení několika řádků zašifrovaných zpráv svých spolužáků rozluštit obsah celého sdělení. Díky tomu za školních let zažil jistě i perné chvíle, ale láska k tomuto vědnímu oboru ho nikdy neopustila. Díky jeho věhlasu se na něj obracely osoby, převážně historikové, jež díky nerozluštěným materiálům byly v koncích se svým bádáním. Dokázal vyluštit šifru Henrietty Marie, jež byla manželkou anglického krále Karla I., který vládl Anglii na počátku 17. století. Díky všem poznámkám z dešifrování sepsal knihu The Philosophy of Decyphering.

Inspirací pro zdolání nevyluštitelné šifry byla korespondence s jistým bristolským zubařem, jenž tvrdil, že vytvořil zcela novou šifru, ale ve skutečnosti se jednalo pouze o ekvivalent Vigenèrovy šifry. Babbage byl vyzván, aby tuto šifru prolomil. Tento úkol se mu podařil v roce 1854, ale o svém úspěchu nevydal svědectví. Přičítá se to jeho zvyku nechat práci rozdělanou a nepublikovat výsledky. Jeho úspěch je naštěstí zaznamenán o několik let později, díky Kasiskému[35].

Friedrich W. Kasiski, pruský major 33. pluku, ve své knize *Die Geheimschriften und die Dechiffrir-kunst* (česky *Tajné šifry a luštitelské umění*), vydané roku 1863, zveřejnil postup, jak dešifrovat polyalfabetickou šifru s opakujícím se klíčem.

Přestože tento objev patří Babbageovi, celý tento postup se jmenuje po Kasiském, a to Kasiského test.

5.2 Kasiského test

Kasiského test je metoda, jak odhadnout délku cyklicky se opakujícího klíče využívaného u polyalfabetické šifry.

Prvním krokem u luštění takové šifry je hledání sekvence písmen, jež se v textu

[35] Friedrich W. Kasiski - 1805-1881

opakují. Toto opakování vzniká dvěma způsoby. Jednou z možností je, že byla tatáž písmena v otevřeném textu zašifrována touž částí klíčového slova. Ale existuje ještě jedna možnost, která je ovšem méně pravděpodobná, a to ta, že rozdílná písmena v otevřeném textu byla zašifrována rozdílnými částmi klíčového slova a náhodou vznikl stejný výsledek jako v prvním případě.

Pro lepší ilustraci celého procesu budu dešifrovat zprávu, jež má následující znění:

NYXOT MINUY FLTMJ ZVAYD NTLDE VLRYR DTXES FDPCI ARZFA XIVVI XMFJE WYEFY KONST VNJJD ZSTPR JVLMI COVVI XELLY XHZWZ VJTCT DLTJE IADOK JRPCP JNOON XEMED ZBPJP ZCYKJ ZNFDN ZAMIC COXEC CRLXI GIDSF MOGKC DKWSC KRPNP JTPXC DOYKL IIXSN VRFCI OEWSM.

Abychom mohli zjistit délku klíče, musíme zjistit vzdálenost dvou takových opakujících se sekvencí, tedy dvou podobných částí kryptogramu, budeme využívat části, které jsou větší než dva znaky.

Dělitelé	VVI (2) - 35	VIX (2) - 35	MIC (2) - 64	VVIX (2) - 35
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				

Tabulka 7. Možné délky klíče Vigenèrovy šifry

Předcházející tabulka označuje různé sekvence, které se v textu objevují. Čísla v závorkách určují počet těchto opakování. Následující číslo označuje jejich vzdálenost. Vyznačené body v tabulce nám ukazují dělitele těchto vzdáleností, což je pro nás mnohem důležitější než vzdálenost samotná. V tabulce chybí řádek s číslem 1, neboť tam by se jednalo o posun o jednu konstantní hodnotu a šifra by v podstatě byla monoalfabetická.

Nejčastěji se zde v tabulce objevuje číslo 5 a 7, což tedy budeme považovat za délku potencionálního klíče. Kvůli zdlouhavé práci s ověřováním obou variant zvolíme první z nich a budeme předpokládat, že náš klíč má pouze 5 znaků.

Tímto jsme na náš kryptogram použili Kasiského test. Dále se budeme zabývat otázkou, jak zprávu úplně dešifrovat.

5.3 Aplikace frekvenční analýzy

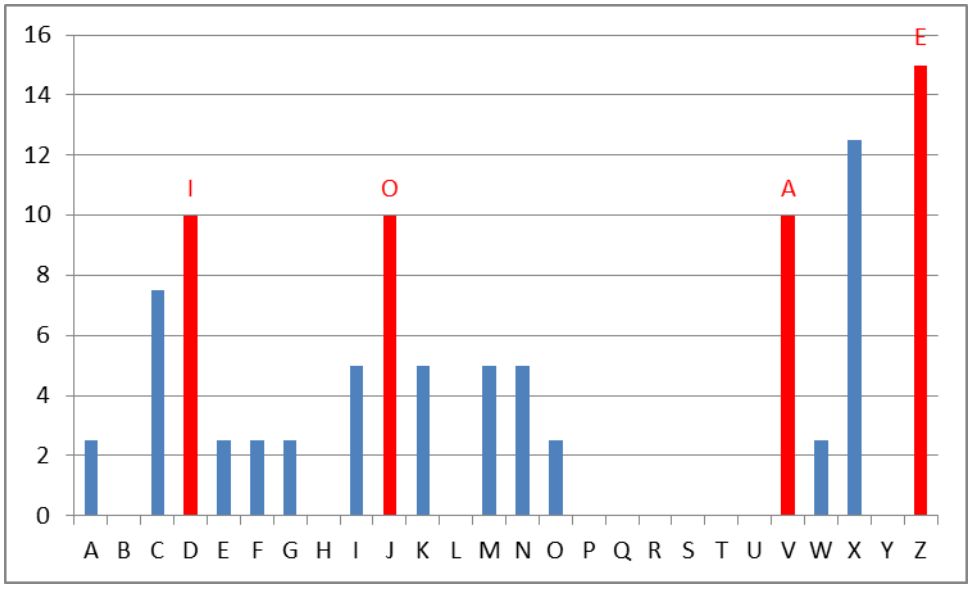
Každé písmeno našeho klíče označíme písmenem K, takže náš klíč má prozatím podobu K_1, K_2, K_3, K_4, K_5.

Znak K_1 nám tedy určuje, jak bylo zašifrováno první písmeno zprávy. Vytváří nám první šifrovací monoalfabetickou šifru, druhý znak K_2 nám šifruje druhý znak a tak to jde dále až k K_5. Jedná se tedy o to, že každý znak z našeho klíče K_1, K_2, K_3, K_4, K_5 nám dešifruje jednu pětinu zprávy. Abychom tedy vyluštili zprávu, každému znaku K_1 až K_5 z našeho klíčového slova přiřadíme písmena z šifrovaného textu, která jsme tímto klíčem zašifrovali. Tedy najdeme písmena z šifry, která se šifrovala např. znakem K_3. Následně využijeme frekvenční analýzu k prolomení této jednoduché substituční šifry, po aplikaci na všechny části šifrovaného textu jsme rozluštili „nerozluštitelnou“ Vigenèrovu šifru.

Pokud tedy budeme pokračovat s naší šifrou, prvnímu prvku klíče K_1 vypíšeme všechny jemu odpovídající znaky. Jedná se o písmena N M E Z N V D F A X X W K V Z J C X X V D I J J X Z Z Z Z C C G M D K J D I V O.

Abychom mohli poznat, podle kterého písmena byl zašifrován první oddíl textu, budeme se všímat základních vlastností, které má klasický český text. Mezi nejfrekventovanější písmena v češtině patří E A O I. Abychom tohoto poznatku mohli

využít, najdeme mezery mezi těmito písmeny. Mezi písmeny A a E je mezera 4 písmena, další mezera 4 a další 6. V tabulce s četností písmen v kryptogramu tedy nalezneme několik písmen, která se se opakují nejčastěji a pokusíme se najít tyto mezery, které by poukazovaly na často užívaná písmena.



Obrázek 4. Četnost znaků u K_1

V kryptogramu se nejčastěji objevují písmena D J V X Z. Z tohoto poznatku budeme dále vycházet. Můžeme si všimnout, že jediné mezery velikosti 4 u často vyskytovaných písmen jsou u dvojic VZ a ZD. Dále budeme hledat, jestli se ve vzdálenosti 8 od počátečního písmene nenachází další frekventované písmeno, u dvojice VZ je to písmeno D, u druhé dvojice tomuto místu odpovídá písmeno H, které se ovšem v kryptogramu nevyskytuje vůbec. Dalším krokem je hledání posledního frekventovaného písmene, a to ve vzdálenosti 14 od prvního písmene. V našem jediném případě tomuto místu odpovídá písmeno J. Jako klíč pro tento oddíl šifry tedy budeme považovat písmeno V, které odpovídá posunu písmene A.

Podobný proces budeme opakovat u dalších čtyř oddílů naší šifry. Pro K_2 nám vyjde písmeno A. Pro K_3 máme dva možné výsledky, a to písmeno P, které je pravděpodobnější, a písmeno L. Pro K_4 máme jediný výsledek, a to písmeno K, a pro K_5 nám vychází výsledek E.

Dostaneme tedy dvě podoby možného klíče, a to slova VAPKE a VALKE. Obě tyto podoby jsou zvláštní, a proto zkusíme dešifrovat první část našeho kryptogramu v délce 10 písmen a zkusíme zjistit, z kterého klíče nám vyjde srozumitelnější text.

Pro klíčové slovo VAPKE nám vyjde SYLEPRIYKU a pro klíčové slovo VALKE nám vyjde slovo SYMEPRICKU. Díky těmto ne zrovna srozumitelným podobám otevřeného textu jsme zjistili, že ani jeden klíč nebude zcela správně. Proto budeme muset najít příhodnější tvar klíčového slova.

Díky dvěma samohláskám u sebe v prvním slově, a to písmen I a Y, což se v češtině zpravidla nevyskytuje, budeme předpokládat, že písmeno P v prvním klíči není správně, a tak dále budeme pracovat s druhou verzí našeho klíče jako se správnější variantou.

Nejméně pravděpodobné písmeno v našem klíči bude nejspíše poslední písmeno E. Z frekvenční analýzy českého jazyka vyplývá fakt, že posledním písmenem jakéhokoliv slova bývá ze 70 % u slovníkových tvarů a ze 73 % u obecné češtiny samohláska. Jako poslední písmeno bude tedy pravděpodobně A, I, Y, O, U. Písmeno I můžeme rovnou zavrhnout, neboť v českém textu se I zpravidla nepíše za písmenem K.

Nejpravděpodobnějším klíčem K_5 bude tedy písmeno A.

Za naše klíčové slovo tedy budeme považovat slovo VALKA. Pro tento klíč nám tedy vyjde otevřený text jako SYMETRICKY, což je naprosto srozumitelné.

Po použití tohoto klíče tedy dešifrujeme celý šifrovaný text a jako otevřený text vyjde definice symetrického klíče, kterou jsme uvedli na začátku této práce.

6 Závěr

Šifrování má velmi zajímavou a bohatou historii, přesto jsem ve své práci uvedla pouze nejznámější z šifer, které během více jak dvoutisícileté historie tohoto vědního oboru vznikly. Záměrně jsem neuváděla šifrovací systémy, které byly objeveny v období po druhé světové válce, tedy v druhé polovině 20. století.

V své práci jsem se věnovala rozdělení šifer na substituční a transpoziční. Také jsem uvedla příklady každé této šifry v kryptologické praxi s příkladem konkrétně zašifrovaného slova. Byly zde uvedeny příklady monoalfabetické, polygrafické, polyalfabetické i jednoduché transpoziční šifry.

Dále byla v této práci věnovaná celá kapitola matematickým definicím z oboru modulární aritmetiky, které se využívají pro tvoření i luštění šifer.

V poslední části své práce jsem chtěla poukázat na složitost luštění Vigenèrovy šifry bez znalosti klíče a proč trvalo 300 let, než se podařilo prolomit tuto šifru. Uvedla jsem na příkladu zašifrovaného textu použití tzv. Kasiského testu a následně použití frekvenční analýzy pro dešifrování textu.

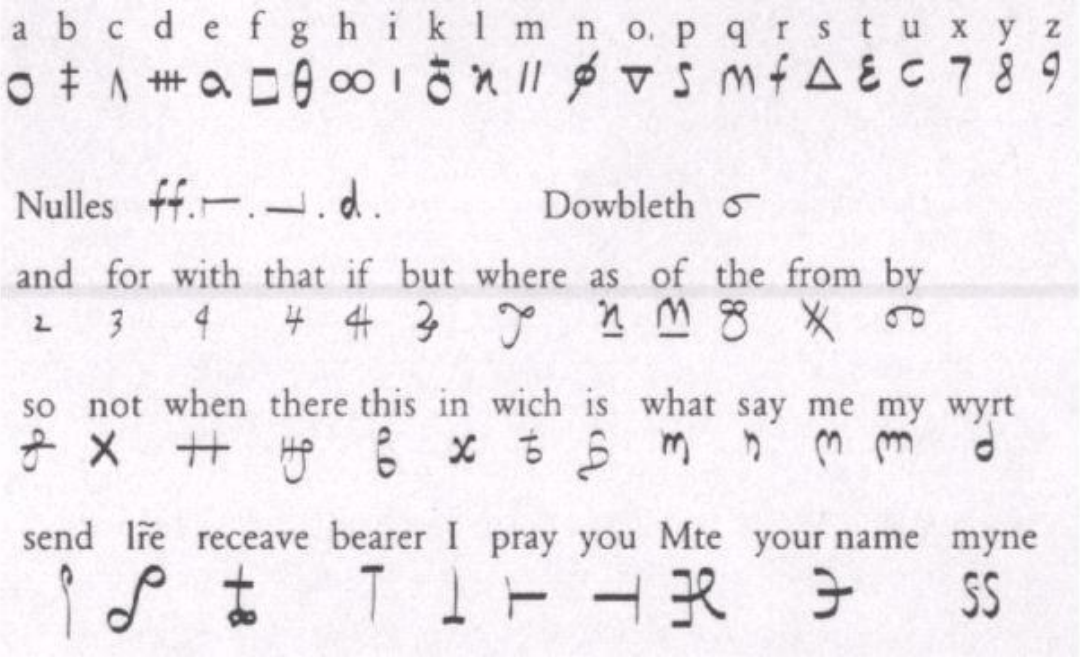
7 Použité zdroje

1. GROŠEK, Otokar a PORUBSKÝ, Štefan. *Šifrovanie - algoritmy, metódy, prax*. Praha: Grada, 1992.
2. JANEČEK, Jiří. *Rozluštěná tajemství: luštitelé, dešifranti, kódy a odhalení*. Praha: XYZ, 2008.
3. KAHN, David. *The Codebreakers: The Story of Secret Writing*. New York: Macmillan, 1968.
4. *Katedra inženýrské informatiky* [online]. [cit. 2016-04-13]. Dostupné z: http://kix.fsv.cvut.cz/~vanicek/vyuka_l01/kos3.htm
5. KRÁLÍK, Jan. *Statistika českých grafémů s využitím moderní výpočetní techniky.* Slovo a slovesnost XLIV, 1983, s. 295-304.
6. PIPER, F. C. a MURPHY. Sean. *Kryptografie*. Praha: Dokořán, 2006.
7. SCHNEIER, Bruce. *Applied cryptography*. New York: Wiley, 1996.
8. SINGH, Simon. *Kniha kódů a šifer: tajná komunikace od starého Egypta po kvantovou kryptografii*. Praha: Dokořán, 2009.
9. *Svět hardware* [online]. [cit. 2016-04-16]. Dostupné z: http://www.svethardware.cz/sifrovani-a-biometrie-pod-drobnohledem/25723
10. TILBORG, Henk. *Fundamentals of cryptology*. Boston: Kluwer Academic Publishers, 2000.
11. TLUSTÝ, Pavel. *Obecná algebra pro učitele*. České Budějovice: Jihočeská univerzita, 2006.

8 Přílohy

Písmeno	Odborná č. %	Obecná č. %	Rozdíl	Písmeno	Odborná č. %	Obecná č. %	Rozdíl
A	8,2416	8,4548	-0,2132	**N**	6,5602	6,6167	-0,0565
B	1,5317	1,5582	-0,0265	**O**	8,8163	8,6977	0,1186
C	2,4515	2,5557	-0,1042	**P**	3,3577	3,4127	-0,055
D	3,4652	3,6241	-0,1589	**Q**	0,0021	0,0013	0,0008
E	10,8434	10,6751	0,1683	**R**	4,8526	4,9136	-0,061
F	0,3249	0,2732	0,0517	**S**	5,1708	5,3212	-0,1504
G	0,3086	0,2729	0,0357	**T**	6,0452	5,7694	0,2758
H	1,2715	1,2712	0,0003	**U**	3,9258	3,9422	-0,0164
Ch	1,1897	1,1709	0,0188	**V**	4,5321	4,6616	-0,1295
I	7,6126	7,6227	-0,0101	**W**	0,0091	0,0088	0,0003
J	2,2198	2,1194	0,1004	**X**	0,097	0,0755	0,0215
K	3,7962	3,7367	0,0595	**Y**	3,0924	2,9814	0,111
L	3,7384	3,8424	-0,104	**Z**	3,2077	3,1939	0,0138
M	3,3359	3,2267	0,1092	**∑**	100	100	

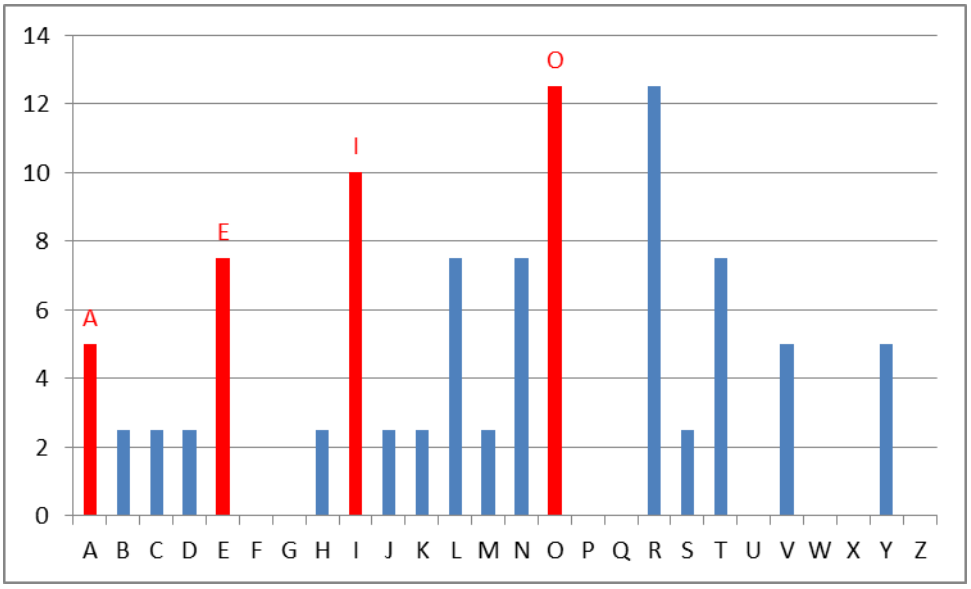
Tabulka 8. Frekvenční analýza odborné a obecné češtiny[36]



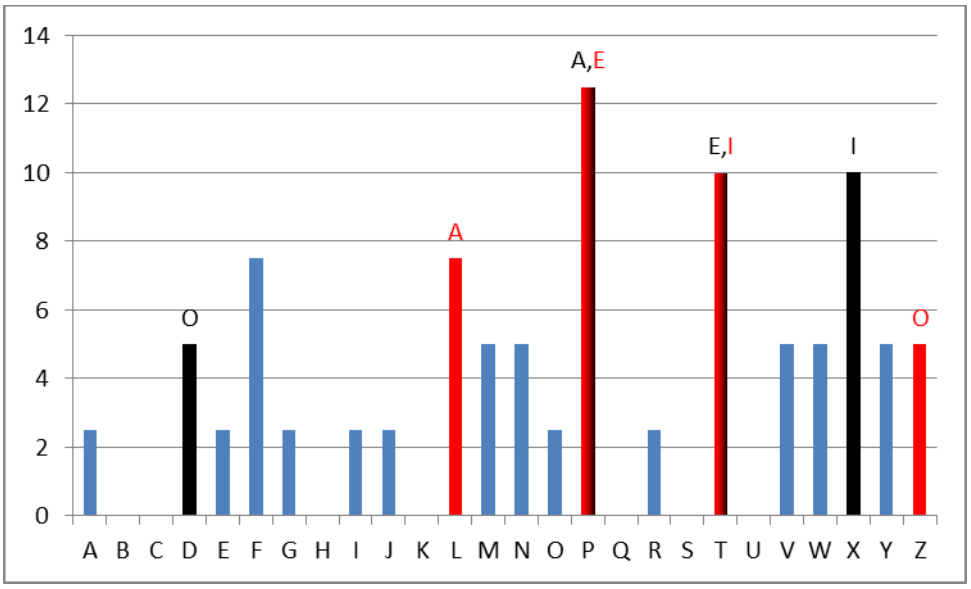
Obrázek 5. Šifra Marie Stuartovny[37]

[36] KRÁLÍK, Jan. Statistika českých grafémů s využitím moderní výpočetní techniky, Slovo a slovesnost XLIV, 1983, s. 295-304.

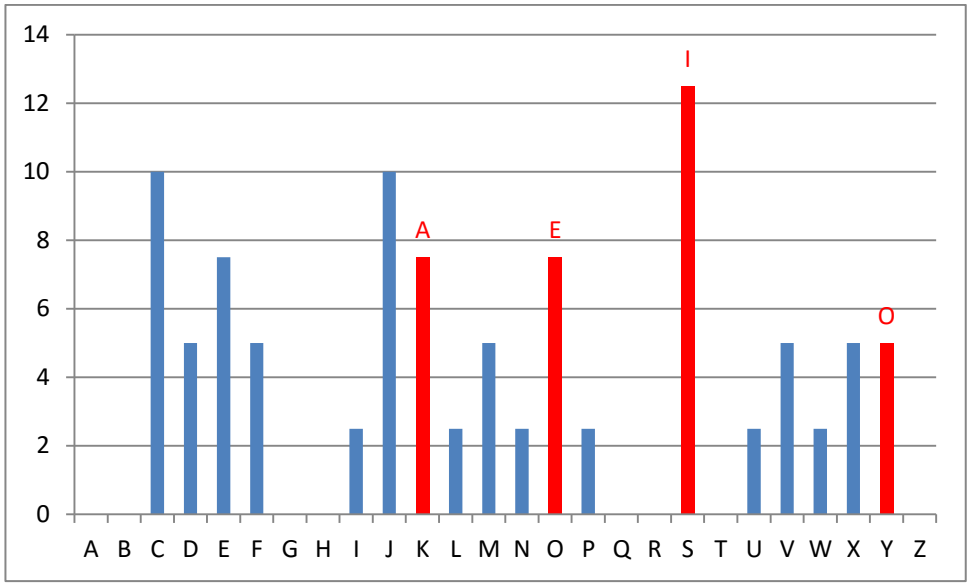
[37] http://www.svethardware.cz/sifrovani-a-biometrie-pod-drobnohledem/25723



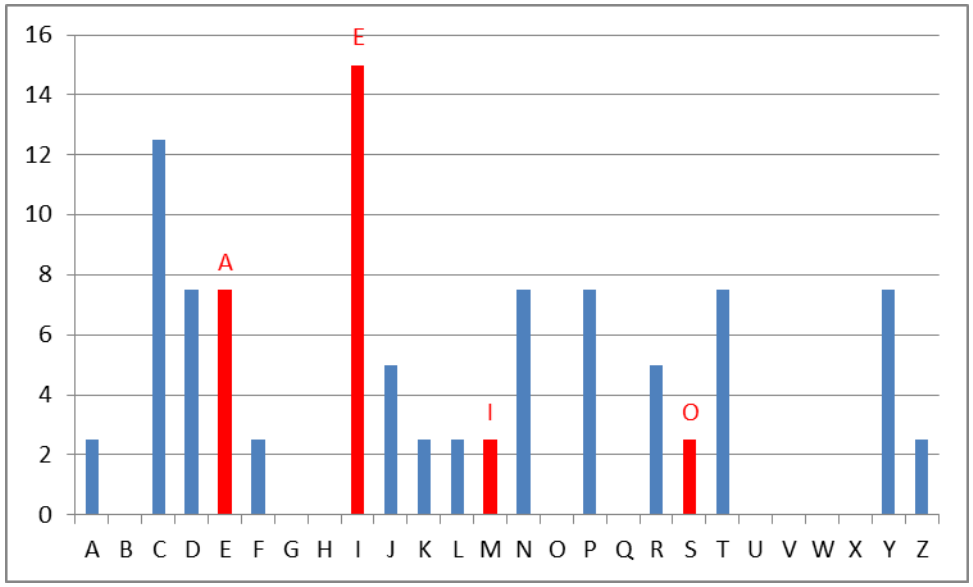
Obrázek 6. Četnost znaků u Vigenèra pro K_2



Obrázek 7. Četnost znaků u Vigenèra pro K_3



Obrázek 8. Četnost znaků u Vigenèra pro K_4



Obrázek 9. Četnost znaků u Vigenèra pro K_5

Samohlásky	Slovníkový tvar %	Obecný tvar %
A	10,237	7,7631
E	8,7681	14,9768
I	27,8703	21,0994
Y	20,5433	13,2375
O	2,362	5,2064
U	0,2718	11,006
∑	70,0525	73,2892

Tabulka 9. Frekvenční analýza koncovek[38]

[38] KRÁLÍK, Jan. Statistika českých grafémů s využitím moderní výpočetní techniky, Slovo a slovesnost XLIV, 1983, s. 295-304.