

Klasická kryptologie: Historické šifry

L'ubomíra Balková

Úvod do kryptologie

13. února 2012

Klasická kryptografie

- končí 2. světovou válkou a nástupem počítačů
- praxe: rozluštění Enigmy polskými kryptoanalytiky počátkem roku 1933 a stavba prvního počítače v Bletchley Parku
- teorie: americký matematik Claude Shannon a jeho práce ze 40. let

Obsah

- 1 Základní pojmy
- 2 Substituce
- 3 Příklady monoalfabetických substitucí
- 4 Příklady polyalfabetických substitucí
- 5 Transpozice

kryptografie × steganografie
 nauka o utajování nauka o utajování
 obsahu komunikace komunikace

- Kahnova encyklopedická kniha *The Codebreakers* z roku 1967 ustálila dnešní použití

kryptografie kryptologie & kryptoanalýza
 navrhování odhalování slabin
 šifrovacích systémů šifrovacích systémů

- *kódy* = nahrazování skupin slov či vět vyhrazenými slovy či větami (*kódové knihy* = seznamy slov či vět otevřeného textu a odpovídajících šifrových textů)
- *kódování* = změna zápisu pomocí veřejně známých pravidel (morseovka, Baudotův kód, kódování písmen pomocí 0 – 25)

Definice pojmů

- \mathcal{M} = konečná množina otevřených textů (message space)
- \mathcal{C} = konečná množina šifrových textů (ciphertext space)
- \mathcal{K} = konečná množina klíčů (keyspace)
- *šifrovací transformace* = bijekce $E_e : \mathcal{M} \rightarrow \mathcal{C}$, kde $e \in \mathcal{K}$
- *dešifrovací transformace* = bijekce $D_d : \mathcal{C} \rightarrow \mathcal{M}$, kde $d \in \mathcal{K}$

Pozn.

existuje i obecnější definice, kdy šifrovací transformace nemusí být funkce, při šifrování možnost volby, ale dešifrovací transformace už funkce je (dešifrování musí být jednoznačné), takovým příkladem je homofonní substituce

konvence: otevřený text \rightarrow ŠIFROVÝ TEXT

Definice kryptosystému

Definice

Kryptosystém (šifra) je dvojice množin $\{E_e \mid e \in \mathcal{K}\}$ a $\{E_e^{-1} \mid e \in \mathcal{K}\} = \{D_d \mid d \in \mathcal{K}\}$ taková, že pro $(\forall e \in \mathcal{K}) (\exists_1 d \in \mathcal{K}) (\forall m \in \mathcal{M}) (D_d(E_e(m)) = m)$.

$(e, d) = \text{pár klíčů (keypair)}$

- na principu *confusion (směšování)*: zastírá závislost mezi m a c záměnou znaků (substituce)
- na principu *diffusion (rozptyl)*: rozprostře info obsažené v m po celé šifře pomocí permutace (transpozice)
- *šifry se symetrickým klíčem* (symmetric-key ciphers, one-key, single-key, conventional): pro každý pár klíčů (e, d) výpočetně snadné určit e ze znalosti d a naopak, často: $e = d$, odtud název
- *blokové*: otevřený text se rozdělí do bloků (strings) pevné délky a ty se šifrují stejným klíčem
- *proudové*: *keystream* (proudový klíč) se sčítá s m bit po bitu

Bloková šifra se symetrickým klíčem.

Definice

Nechť \mathcal{A} je abeceda, $|\mathcal{A}| = n$. \mathcal{M} je množina slov délky r nad \mathcal{A} .
Blokovou šifru s bloky délky r a prostorem klíčů

$$\mathcal{K} = \{e = (\sigma_1, \sigma_2, \dots, \sigma_r) \mid \sigma_i \text{ permutace na } \mathcal{A}\}$$

nazveme substitucí, pokud pro každé $m = (m_1, m_2, \dots, m_r) \in \mathcal{M}$ a každé $e \in \mathcal{K}$ platí

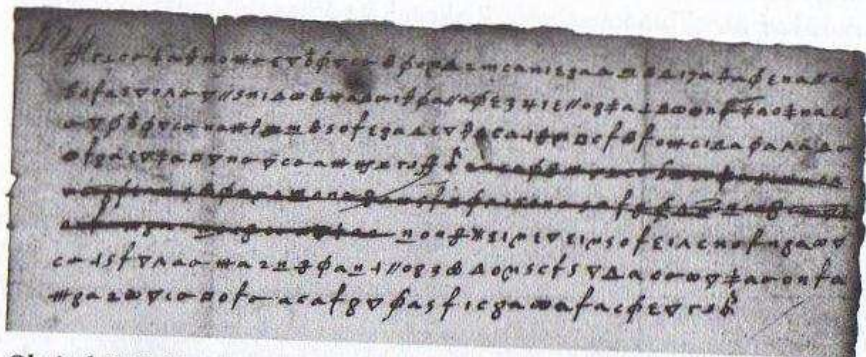
$$E_e(m) = (\sigma_1(m_1), \sigma_2(m_2), \dots, \sigma_r(m_r)) = c$$

$$D_d(c) = (\sigma_1^{-1}(c_1), \sigma_2^{-1}(c_2), \dots, \sigma_r^{-1}(c_r)) = m.$$

Počet možných klíčů $|\mathcal{K}| = (n!)^r$.

- 1 pro $\sigma_1 = \sigma_2 = \dots = \sigma_r$ jde o *monoalfabetickou* substituci (šifrovací tabulka má jen 2 řádky)
- 2 jinak *polyalfabetickou*

Šifra Marie Stuartovny



Obrázek 9: Padělané postskriptum, které doplnil Thomas Phelippes k Mariině zprávě, jež byla dešifrována pomocí nomenklátoru Marie Stuartovny (viz obrázek 8).

Caesarova šifra

Šifrová tabulka Caesarovy šifry

Plain	a	b	c	d	e	f	g	h	i	j	k	l	m
Cipher	D	E	F	G	H	I	J	K	L	M	N	O	P
Plain	n	o	p	q	r	s	t	u	v	w	x	y	z
Cipher	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Příklad

Pomocí Caesarovy šifry rozluštěte citát, který Caesar vyslovil při překročení řeky Rubikon určující hranici mezi Galií a Itálií. Vyhláškou tím obrazně válku Římu, protože jeho moc byla v té době omezena na Galii (prokonzul).

W, K, H G, L, H L, V F, D, V, W

Řešení: The die is cast. = Kostky jsou vrženy.

Posuvné šifry (shift ciphers)

- nejznámější je *Caesarova šifra*
- **Algoritmus:** $\mathcal{M} = \mathcal{C} = \{0, 1, \dots, 25\}^*$, $E_e(m_j) := m_j + b \pmod{26}$
pro $e = b$ a $D_d(c_j) := c_j - b \pmod{26}$
- šifra se symetrickým klíčem $e = -d = b$, bloková šifra s bloky délky 1
- **Kryptoanalýza:** posuvné šifry lehce rozluštitelné, stačí vyzkoušet jen $\#\mathcal{A}$ možností

Afinní šifry (affine ciphers)

- Algoritmus:** necht' $a, b, n \in \mathbb{N}$, $\mathcal{M} = \mathcal{C} = \{0, 1, \dots, n-1\}^*$,
 $E_e(m_j) := am_j + b \pmod n$ pro $e = (a, b)$, aby existovala
 $E_e^{-1}(c_j) = a^{-1}(c_j - b) \pmod n$, musí existovat $a^{-1} \pmod n$, to je
 splněno, právě když $\text{nsd}(a, n) = 1$
- takových a je $\phi(n)$, proto pro pevné n existuje $\phi(n)n$ afinních šifer
- posuvné jsou speciálním případem
- šifra se symetrickým klíčem $d = (a^{-1}, -a^{-1}b)$, bloková šifra s bloky
 délky 1

Příklad

$n = 26$, $\mathcal{M} = \mathcal{C} = \mathbb{Z}/_{26}\mathbb{Z}$, $e = (7, 5)$, pak $E_e(m_j) = 7m_j + 5 \pmod{26} = c_j$,
 a tedy $D_d(c_j) = 15(c_j - 5) \pmod{26} = m_j$. Dešifrujte text

$c = 19, 20, 17, 6, 8, 5, 18, 5, 4, 17, 1, 8$.

$m = 2, 17, 24, 15, 19, 0, 13, 0, 11, 24, 18, 19$

Řešení: cryptanalyst.

Substituce s klíčovým slovem

Plain	a	b	c	d	e	f	g	h	i	j	k	l	m
Cipher	J	U	L	I	S	C	A	E	R	B	D	F	G
Plain	n	o	p	q	r	s	t	u	v	w	x	y	z
Cipher	H	K	M	N	O	P	Q	T	V	W	X	Y	Z

Polybiův čtverec, šachovnice

- Algoritmus:** $E_e : \mathcal{A}^* \rightarrow (\hat{5} \times \hat{5})^*$, tedy písmenům anglické abecedy přiřazuje dvojice čísel od 1 do 5

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i/j	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

Příklad

Pomocí Polybiova čtverce dešifrujte citát přiřazený Augustovi v Životě 12 císařů od římského historika Suetonia (70-140 n.l.)

$c = 32112515 \ 2311434415 \ 433134523154$
m = Make haste slowly. = Spěchej pomalu.

Playfairova šifra

- **Algoritmus:** bigramová šifra, zašifruje pomocí tabulky

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

+

úprav otevřeného textu:

mm → mzm,

lichá délka textu → na konec se přidá z

- ▶ SF → BE (náhrada sousedními znaky vpravo)
- ▶ SI → IP (náhrada sousedními znaky dole)
- ▶ VU → ZO, KE → GB (diagonální řádková náhrada)

Příklad

Dešifrujte text pomocí klíčového slova Charles

$c = \text{ETGUUYYDPFUYP RDG}$

$m = \text{donotztrustzthem}$

Hillova šifra

- **Algoritmus:** šifruje r -gramy na r -gramy, dáno $n, r \in \mathbb{N}$,
 $\mathcal{K} = \{e \in (\mathbb{Z}/n\mathbb{Z})^{r \times r} \mid \exists e^{-1} \pmod{n}\}$, $\mathcal{M} = \mathcal{C} = (\mathbb{Z}/n\mathbb{Z})^r$, pro $\forall m \in \mathcal{M}$
 definujeme $E_e(m) = m \cdot e$ a $D_d(c) = c \cdot e^{-1}$
- platí, že e^{-1} existuje $\Leftrightarrow \text{nsd}(\det(e), n) = 1$

Příklad

$r = 2$, $n = 26$, $\mathcal{M} = \mathcal{C} = (\mathbb{Z}/26\mathbb{Z})^2$, $\mathcal{K} =$ všechny invertibilní matice 2×2 ,
 tj. $e \in \mathcal{K} \Leftrightarrow \text{nsd}(\det(e), 26) = 1$. Vezměme $e = \begin{pmatrix} 3 & 6 \\ 1 & 5 \end{pmatrix}$, jelikož $\det(e) = 9$,
 může hrát roli klíče. Zašifrujte text *movie*.

$$m = 12, 14, 21, 8, 4, 25$$

$$E_e(12, 14) = (12, 14) \begin{pmatrix} 3 & 6 \\ 1 & 5 \end{pmatrix} \pmod{26} = (24, 12) \text{ atd.}$$

$$c = 24, 12, 19, 10, 11, 19$$

Šifrový text je tedy *YMTKLT*. K dešifrování je třeba znát $e^{-1} = \begin{pmatrix} 15 & 8 \\ 23 & 9 \end{pmatrix}$.

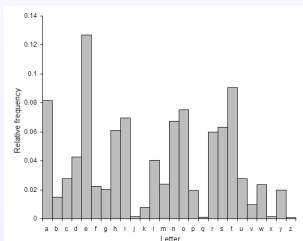
Relativní četnosti písmen v angličtině

a	b	c	d	e	f	g	h	i
8.167	1.492	2.782	4.253	12.702	2.228	2.015	6.094	6.966
j	k	l	m	n	o	p	q	r
0.153	0.772	4.025	2.406	6.749	7.507	1.929	0.095	5.987
s	t	u	v	w	x	y	z	
6.327	9.056	2.758	0.978	2.360	0.150	1.974	0.074	

Pořadí od nejčastějšího k nejméně častému písmenu

etaonirshdlucmpfywgbvjkqxz D.Kahn, 1967

etaonrshdlfcmugpywbvkvxjqz A.G. Konheim, 1981



Frekvenční analýza

9. stol. - al-Kindí: *Rukopis o dešifrování kryptografických zpráv*

Příklad

Pomocí frekvenční analýzy a nápovědy, že šifrová tabulka obsahuje klíčové slovo, prolomte:

*P GPQOEGPQSUSPH SN P LEVSUE RIM QTMHSHD UIRREE SHQI
QOEIMEGN. SR HTGAEMN PME HIQ AEPTQSRTF, HIQOSHD SN.*

Frekvenční analýza není všemocná, obecně funguje až pro texty s aspoň stovkou slov. "From Zanzibar to Zambia and Zaire, ozone zones make zebras run zany zigzags."

Vylepšení monoalfabetické substituce

- *klamače (nuly)* = šifrování pomocí širší abecedy (např. 0 až 99), mnoho z písmen nemá význam
- komolení pravopisu
- kombinace s kódovými slovy = *nomenklátory*

Homofonní substituce

Kompromis mezi monoalfabetickou a polyalfabetickou substitucí

- písmenu odpovídá takové procento znaků šifrové abecedy, kolik je frekvence písmene \Rightarrow znemožnění frekvenční analýzy
- **Kryptoanalýza:** lze vyjít ze vztahů mezi písmeny v anglických textech, např. za q následuje vždy u , q je reprezentováno jediným symbolem a u třemi symboly, najdeme tedy písmeno, za nímž se vyskytují jen tři další a máme pravděpodobně q a u

	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
09	48	13	01	14	10	06	23	32	15	04	26	22	18	00	38	94	29	11	17	08	34	60	28	21	02	
12	81	41	03	16	31	25	39	70			37	27	58	05	95		35	19	20	61		89			52	
33		62	45	24			50	73			51	59	07				40	36	30	63						
47			79	44			56	83			84	66	54				42	76	43							
53				46			65	88				71	72				77	86	49							
67				55			68	93				91	90				80	96	69							
78				57									99							75						
92				64																85						
				74																97						
				82																						
				87																						

Albertiho šifrovací disk

Alberti jako první použil kódovou knihu, číslům 11 až 4444 přiřadil věty, např. číslu 21 odpovídala věta: Poslat v plucích, příjemci pak poslal zprávu: Poslat v plucích & P + šifrový text

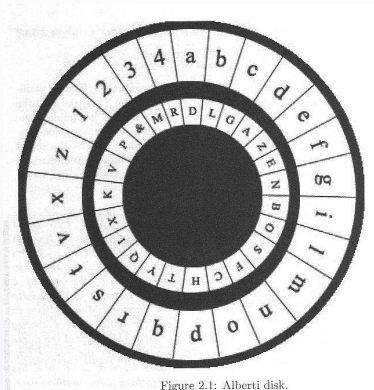


Figure 2.1: Alberti disk.

Vigenèrova šifra - Chiffre indéchiffrable

Nejčastější varianty:

- ① periodický klíč - opakující se slovo
- ② smysluplný klíč - text, který nesouvisí s otevřeným textem nebo kombinace *priming key* a otevřeného textu (autokláv otevř. text)
- ③ klíč = kombinace *priming key* a šifrového textu (autokláv šifrtext)
- ④ náhodný klíč \Rightarrow Vernamova šifra (Shannon)

Příklad

Dešifrujte Vigenèrovu šifru s periodickým klíčem period.

ELZTCVDTYG WV P KRUS ZXXY WPMTGKQJHH EEL BR GYCMG

Dešifrujte Vigenèrovu šifru s priming key plain.

BLTPRYAMPGE IL I ISUW WOTT VQTXZ RHO RG OOMRQHJEIU

Beaufortova šifra

- varianta Vigenèrovy šifry s periodickým klíčem
- **Algoritmus:** stejná šifrovací i dešifrovací funkce, tj. inverzní sama k sobě $\mathcal{M} = \mathcal{C} = \mathbb{Z}/n\mathbb{Z}$, $e = (e_1, e_2, \dots, e_r) \in \mathcal{K}$,

$$E_e(m_j) = e_j \bmod r - m_j \bmod n$$

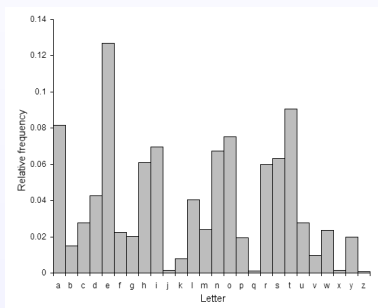
Příklad

Dešifrujte Beaufortovu šifru s periodickým klíčem period.

1,19,14 22,14,10,8,0,5,8,21,21,13,22,17,21,22 16,11,9,13,17 11,21,11.
 11,10,4,6,12 23,6,25,15,10,17,23 21,23,16,21 2,11,2,3,22,10
 9,2,12,3,23,19,3,14,19,13.

Vigenèrova šifra s periodickým klíčem

- *Kasiského test* = hledání opakujících se bloků v šifrovém textu, které zřejmě vznikly zašifrováním opakujících se úseků otevřeného textu stejným úsekem klíče, $l = \text{nsd}(\text{vzdálenosti opakujících se úseků})$ a to je pravděpodobná délka klíče
- sepsání šifrového textu do tabulky o l sloupcích a v každém z nich provedení frekvenční analýzy



Vigenèrova šifra se smysluplným klíčem

- v otevřeném textu předpokládáme výskyt častých slov (např. the), správně umístěným častým slovům bude odpovídat smysluplný úsek klíče
- nalezený úsek klíče hádáním rozšíříme a najdeme odpovídající nové úseky otevřeného textu

Příklad

Prolomte VHRMHEUZNFQDEZRWXFIDK.

Index koincidence (shody)

Definice

IC je pravděpodobnost, že dvě náhodně vybraná písmena v textu jsou stejná.

Indexy koincidence

angličtina	0,0676
němčina	0,0824
francouzština	0,0801
ruština	0,0470 (32 znaků v abecedě)
čeština	0,0577
náhodný text	$1/26 = 0,0385$

Bloková šifra se symetrickým klíčem.

Definice

Bloková šifra s blokem délky r a prostorem klíčů $\mathcal{K} = \{\text{permutace na } \hat{r}\}$ se nazývá jednoduchá transpozice (permutace), pokud platí pro každé $m = (m_1, m_2, \dots, m_r) \in \mathcal{M}$ a pro každé $e \in \mathcal{K}$

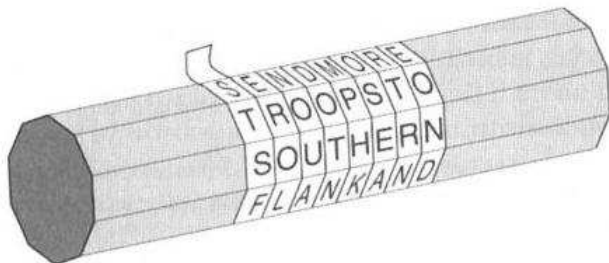
$$E_e(m) = (m_{e(1)}, m_{e(2)}, \dots, m_{e(r)}) = c$$

$$D_d(c) = D_{e^{-1}}(c) = (c_{d(1)}, c_{d(2)}, \dots, c_{d(r)}) = m.$$

Počet možných klíčů $|\mathcal{K}| = r!$

Spartánská transpozice - skytale, scytale

- **Algoritmus:** kožený pásek namotaný na tyči určitého průměru, na nějž se napsal text, pak se pásek sejmul, aby šel přečíst, musel mít příjemce tyč stejného průměru



Jednoduchá transpozice s heslem

- **Algoritmus:**

R	I	F	L	E
5	3	2	4	1
M	I	X	U	J
E	M	I	S	T
A	Z	A	C	H
O	V	A	P	I
S	M	E	N	A

→

JTHIA
XIAAE
IMZVM
USCPN
MEAOS

- **Kryptoanalýza:** snadná při znalosti delšího slova otevřeného textu

Příklad

Prolomte, víte-li, že otevř. text obsahuje slovo transpozice.

R R J C
M N I K
N I I H
T P D M
P T E I
I C C Y
I S V H
V Z N C
I A E H
I O E A.

Německá ADFGVX

- kombinace substituce a transpozice
- **Algoritmus:**

	A	D	F	G	V	X
A	b	3	m	r	l	i
D	a	6	f	ϕ	8	2
F	c	7	s	e	u	h
G	z	9	d	x	k	v
V	1	q	y	w	5	p
X	n	j	t	4	g	o

m = field cipher,

$c = DFAXFGAVGFFAAXVXFXFGAG$

Německá ADFGVX

- poté šifrový text zapsán do obdélníku s použitím číselného klíče odpovídajícího slovu RIFLE

R	I	F	L	E
5	3	2	4	1
D	F	A	X	F
G	A	V	G	F
F	A	A	X	V
X	F	X	F	G
A	G			

výsledný šifrový text *FFVGAVAXFAAFGXGXFDGFXA*

- ADFGVX vybrána, protože jejich znaky v Morseově abecedě odlišné ⇒ méně překlepů při telegrafování
- polygramová monoalfabetická substitučně-transpoziční šifra

- monoalfabetická substituce (Caesar, Marie Stuartovna, Hill, Polybius, Playfair)
- kryptoanalýza frekvenční analýzou (9. stol. - Arabové)
- polyalfabetická šifra (Alberti, Vigenère, Beaufort)
- kryptoanalýza testem periodicity + frekvenční analýzou (1854 - Babbage, 1863 - Kasiski)
- konec 1. svět. války - Vernamova šifra a šifrovací stroje