Monoalfabetické substituční šifry

Ing. Jan Přichystal, Ph.D.

PEF MZLU v Brně

21. října 2010

Úvod

Jeden z prvních popisů substituční šifry se objevuje v Kámasútře z 4. stol, vychází však z rukopisů o 800 let starších.

Princip substitučních šifer spočívá v nahrazení písmen otevřené abecedy písmeny šifrovací abecedy.

Přiřadíme-li písmena naprosto náhodně je počet možných uspořádání 26! čili 4×10^{26}.

Pro efektivní dešifrování příjemcem je však potřeba držet se jednoduchého systému. Lze například použít klíčové slovo.

Pro svou jednoduchost a vysoký stupeň dominovala tato šifra tajné komunikaci po celé první tisíciletí n. l.

Caesarova šifra

Přibližně 2000 let starý a pravděpodobně nejznámější šifrovací systém. Autorství je připisováno Juliu Caesarovi, který ji používal jako jednu z mnoha šifer. Princip spočívá v posunutí šifrovaného písmene o tři místa dále v abecedě.

```
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
```

Příklad:

```
veni vidi vici
YHQL YLGL YLFL
```

Šifra ATBAŠ

Šifrovací systém, který prokazatelně vynalezli a používali Hebrejci. Zachovává princip vzájemné záměny písmen. Princip spočívá v tom, že se vezme písmeno, určí se jeho vzdálenost od začátku abecedy a nahradí se písmenem se stejnou vzdáleností od konce abecedy.

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

nebo

```
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
```

Příklad:

```
kain a abel byli bratri
XNVA N NORY OLYV OENGEV
```

Polybiův čtverec

Autorem šifry je řecký historik Polybios, který navrhoval tento systém používat pro signalizaci. Každé písmeno se kóduje pomocí dvojice čísel – číslem řady a číslem sloupce. Zpráva se pak mohla předávat pomocí různého počtu pochodní držených v levé a pravé ruce. Šifra se z kódu stane, pokud nejsou písmena ve čtverci vepsána abecedně.

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

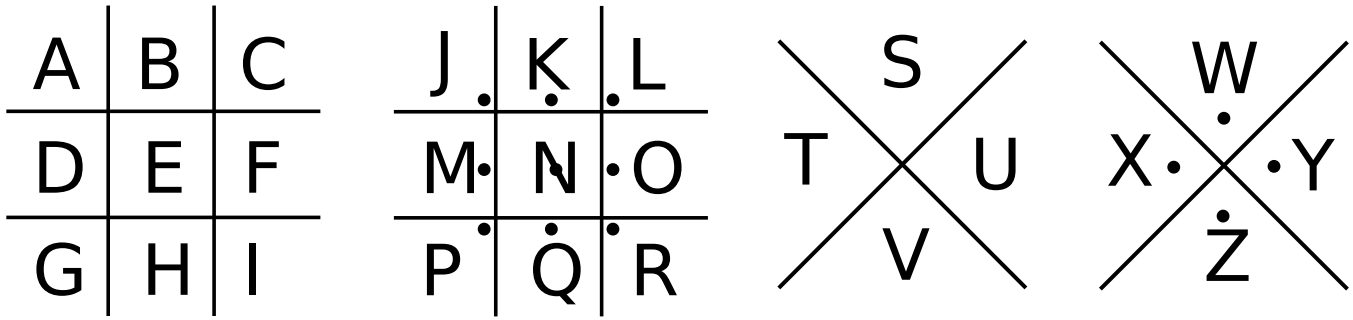
	1	2	3	4	5
1	P	O	L	Y	B
2	I	S	A	C	D
3	E	F	G	H	J
4	K	M	N	Q	R
5	T	U	V	WX	Z

Příklad:

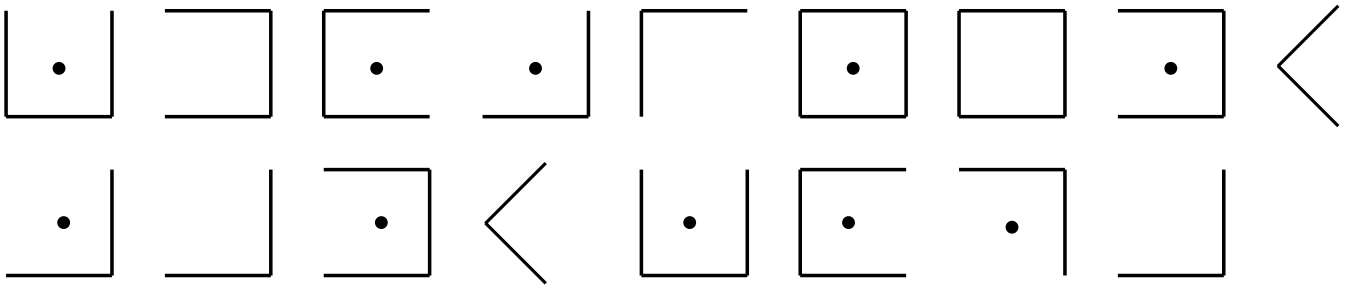
i	k	a	r	u	v	p	a	d
21	41	23	45	52	53	11	23	25

Šifrovací kříž

V historii se objevuje se v různých variantách. Jako jednu ze svých šifer ji používali svobodní zednáři.



Příklad:
kdo jinemu jamu kopa



Obecná substituce

Jde o velmi starý, dlouhou dobu a v různých modifikacích používaný šifrovací systém. Princip spočívá v nahrazení každého znaku otevřené abecedy jedním znakem šifrové abecedy. Možných variant je 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
C G F T K M Q I J A U Y O H P R W S L N V X D B Z E
```

Příklad:

```
mundus vult decipi
OVHTVL XVYN TKFJRJ
```

Substituce s klíčem

Zprávy zašifrované obecnou substitucí jsou pro velký počet možných uspořádání šifrových abeced velmi bezpečné. Problém je však předávání (pamatování) uspořádání. Z tohoto důvodu se často používá klíčové slovo zjednodušující rekonstrukci šifrové abecedy.

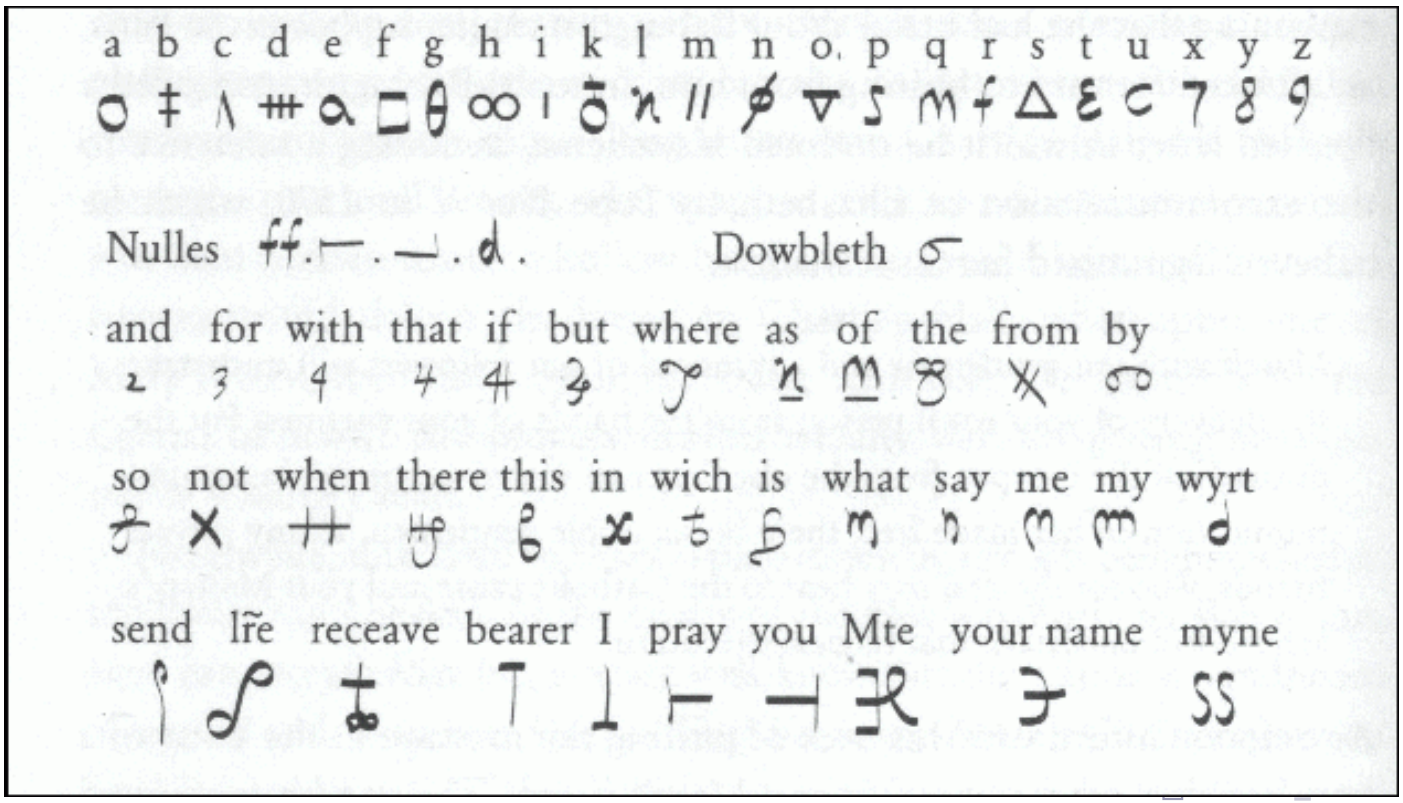
```
a b c d e f g h i j k l m n o p q r s t u v w x y z
P E T R K L I C A B D F G H J M N O Q S U V W X Y Z
```

Příklad:

```
zij vyrovnane
ZAB VYOJVHPHK
```

Vylepšení substitučních šifer

- použití speciálních symbolů pro nahrazení některých slov – nomenklátory
- znaky s klamajícím významem – nulový znak, smaž předchozí znak . . .
- zkomolení zprávy na základě zvukomalebnosti – „frikvence slof“



Homofonní šifra

Používala se jako vhodný kompromis mezi rychlostí šifrování a dešifrování a poskytovaným bezpečím. Odstraňuje hlavní problém monoalfabetických šifer – frekvenční charakter. Každé písmeno se zde nahrazuje řadou reprezentací, přičemž jejich počet je úměrný frekvenci písmene.

```
A 09 12 33 47 53 67 78 92
B 48 81
C 13 41 62
D 01 03 45 79
E 14 16 24 44 46 55 57 64 74 82 87 98
...
Y 21 52
Z 02
```

Příklad:

```
a  b  e  c  e  d  a
09 48 14 13 19 01 12
```

Šifra Playfair

Autorem je Charles Wheatstone, proslavil ji Lyon Playfair. Tato šifra nahrazuje každou dvojici dvojici písmen v otevřeném textu jinou dvojicí písmen. Šifrovací tabulka o rozměru 5×5 se vytvoří na základě dohodnutého klíče. Text se rozdělí na digramy, které se kódují podle speciálních pravidel.

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

Příklad:

```
me et me at ha mx me rs mi th br id ge to ni gh tx
GD DO GD RQ AR KY GD HD NK PR DA MS OG UP GK IC QY
```

Šifra Playfair – postup

- Zpráva se rozdělí na digramy, pokud je poslední písmeno liché, doplní se písmenem x.
- Ve dvojicích stejných písmen je použito nahrazení písmenem x.
- Písmena ve stejném řádku se nahrazují nejbližším písmenem vpravo od každého z nich.
- Pokud je jedno na konci řádku, nahradí se písmenem ze začátku řádku.
- Písmena ve stejném sloupci jsou nahrazena nejbližším písmenem pod každým z nich.
- Písmeno ve sloupci poslední je nahrazeno písmenem z vrcholu.
- Pokud nejsou ani na řádku ani ve sloupci, sestaví se z šifrovaných písmen pomyslný čtverec, jehož vrcholy reprezentují novou dvojici.
- Při dešifrování se postupuje obráceně.

Knižní šifra

Pro bezpečné šifrování lze využít i techniku sdílení rozsáhlé informace, kterou mají k dispozici obě komunikující strany. Typickým představitelem může být kniha. Zpráva může být zašifrována pomocí čísel udávajících písmena z knihy. Například:

- číslo x kóduje první písmeno x-tého slova,
- trojice čísel x, y, z kóduje z-té písmeno na y-tém řádku x-té stránky.

Výhodou je, že každé písmeno lze zašifrovat mnoha různými způsoby. Nelze tedy použít frekvenční analýzu. Nevýhodou je požadavek na stejnou knihu a zdlouhavost této metody.

Kryptoanalýza

Po dlouhá staletí byly monoalfabetické substituční šifry považovány za nerozluštitelné a bezpečné.

V době rozkvětu arabských zemí v 8. století došlo k rozvoji umění luštit šifry, které popsal v knize *Rukopis o dešifrování kryptografických zpráv* arabský učenec al-Kindí.

V Evropě se zásadní poznatky v tomto oboru projevily v době renesance. Myšlenky podněcovalo studium Starého zákona, který obsahuje ukázky kryptografie. Vedoucí úlohu zde hrály Itálie a Francie.

Kryptoanalýza sehrála významnou roli při odhalení spiknutí proti britské královně Alžbětě I.

Postup kryptoanalýzy

Luštění šifer má několik fází:

- Identifikace šifry – určení základního principu
- Odhalení klíče – odhalení klíče, který byl použit pro zašifrování
- Rozluštění – získání otevřeného textu zprávy

Útoky rozdělujeme podle informace, kterou máme k dispozici:

- Pouze zašifrovaný text (COA) – máme pouze zašifrovanou zprávu.
- Známý holý text (KPA) – máme zašifrovanou zprávu i otevřený text.
- Vybraný holý text (CPA) – máme zašifrovanou zprávu a vybraný otevřený text.

Metody kryptoanalýzy

Při luštění šifer se používají nejrůznější techniky z různých oborů lidského vědění (lingvistika, statistika, matematika, informatika . . .)
Velmi důležitá je znalost jazyka otevřeného textu a jeho charakteristik.

- Základní představu o monoalfabetických šifrách získáme pomocí frekvenční analýzy, neaplikujeme ji však slepě.
- Pokoušíme se odhalit samohlásky (40 % textu) – obvykle ve slově alespoň jedna, rozvrstveny relativně pravidelně, zřídka vedle sebe.
- Hledáme častá slova, typické digramy, trigramy.
- Zkoušíme delfskou metodu – hádat celá slova, odhadnout obsah zprávy.

Metody kryptoanalýzy

Další zajímavé sledovatelné charakeristiky:

- Vyskytuje se písmeno samostatně (jednopísmenné slovo)?
- Vyskytuje se písmeno často na začátku, uprostřed nebo na konci slova?
- Jaký je vztah písmene k ostatním písmenům?
- Jaká je pozice souhlásky ve skupině souhlásek (je na začátku nebo na konci)?

Velmi zajímavý popis kryptoanalýzy podává E. A. Poe v povídce Zlatý skarabeus.

La Disparition (Georges Perec) – povídka o 20 stranách, ve které se nevyskytuje ani jedno písmeno *e*.

Frekvence písmen

Pro každý jazyk je specifická frekvence jednotlivých písmen.
S touto znalostí jazyka lze rozpoznávat konkrétní znaky na základě jejich četnosti i po zašifrování.

FREKVENCE v %			
písmeno	relativní	písmeno	relativní
A	8.6	N	6.8
B	1.7	O	8.0
C	3.3	P	3.2
D	3.6	Q	0.0
E	10.5	R	4.9
F	0.2	S	6.3
G	0.2	T	5.1
H	2.2	U	4.0
I	7.5	V	4.3
J	2.2	W	0.0
K	3.6	X	0.1
L	4.2	Y	2.8
M	3.5	Z	3.2

FREKVENCE v %			
písmeno	relativní	písmeno	relativní
A	6.5	O	6.7
B	1.2	P	1.6
C	2.4	Q	0.1
D	3.1	R	5.2
E	10.7	S	5.0
F	2.3	T	8.6
G	1.3	U	2.1
H	4.3	V	0.8
I	5.6	W	1.6
J	0.1	X	0.1
K	0.3	Y	1.6
L	2.8	Z	0.1
M	2.0	-	18.2
N	5.8		

Další charakteristiky

Nejčastější digramy	st, ne, se, na, ni, po, pr, ov, ro, je, te, le, ko, od, ra
Nejčastější trigramy	pro, ost, sta, pri, pre, ter, eni, ova, kte, pra
Nejčastější slova	a, v, se, na, je, ze, s, o, z, i, do, to, pro, ve, k
Nejčastější zač. slov	p, s, v, z, n, o
Nejčastější kon. slov	e, i, a, o, u, y

Další charakteristiky

	Začátek slova	Konec slova	Samostatné slovo	Vazby na souhlásky	Vazby na samohlásky
e	×	●	×	el, em, en, ne, se, ta, je, le	×
a	○	●	●	ak, al, na, ra, ta, la	au
o	○	●	○	ov, po, to	ou
i	×	●	○	li, ni, ic, il	×
u	●	●	○	ku	ou, au
y	×	○	×	by, vy	×

Nepřímé útoky I.

Pro kryptoanalýzu jsou důležité nejen exaktní metody, ale i intuice, štěstí a rafinovanost.

Uhodnutí obsahu

- vojenská korespondence – typická oslovení, tituly, místa, pevná struktura zprávy . . .
- obchodní korespondence – odhadnutí předmětu dopisování (zboží, města, jména obchodních partnerů)
- osobní korespondence – forma oslovení, jméno odesílatele a příjemce
- historické texty – lze očekávat fráze „ve jménu Boha“ atp.

Nepřímé útoky II.

Chyby šifrantů

Slabým místem šifrovacích systémů je lidský faktor. Typickým příkladem je zaslání nezašifrovaného textu spolu se zašifrovaným nebo zaslání stejného obsahu zašifrovaného pomocí různých klíčů.

Intriky

- „Omylem“ vyzrazené komunikační klíče;
- Předstírání, že nemumím číst cizí šifru, ale přitom ji už dávno čtu;
- Použití šifry v šifře;
- Uloupení klíčů;

Děkuji za pozornost

Dotazy?