Výpočetní technika a lékařská informatika

Šifrování

Mgr. Markéta Trnečková, Ph.D.

Komunikace

- Alice = odesílatel (např. lékař)
- Bob = příjemce (např. laboratoř)
- Ctirad = útočník (neoprávněná osoba)

Otázky:

- Jak zabránit Ctiradovi číst zprávu?
- Jak zajistit, že zpráva nebyla změněna?
- Jak ověřit, kdo zprávu odeslal?

Co znamená „bezpečná komunikace“?

- **Důvěrnost (confidentiality)**
  - nikdo nepovolaný nemůže číst zprávu
- **Integrita (integrity)**
  - zpráva nebyla během přenosu změněna
- **Autenticita (authenticity)**
  - víme, kdo zprávu odeslal
- **Nepopiratelnost (non-repudiation)**
  - odesílatel nemůže popřít odeslání zprávy

- přenos zdravotnických dat (lékař $\leftrightarrow$ laboratoř)
- ukládání dat pacientů v databázích
- zabezpečení komunikace (HTTPS)

Příklad:

- laboratorní výsledek je zašifrován před odesláním
- pouze oprávněný příjemce ho může přečíst

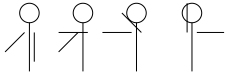
Steganografie

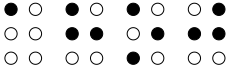
- **Steganografie** – ukrytí existence zprávy
- cílem je, aby si útočník nevšiml, že komunikace probíhá
- **Není totéž jako šifrování!**
  - šifrování: zpráva je vidět, ale není čitelná
  - steganografie: zpráva není vidět
- **Příklady:**
  - neviditelný inkoust
  - digitální steganografie – skrytí dat v obraze (např. v pixelech)
- **Poznámka:**
  - v praxi se často kombinuje se šifrováním

Co oči nevidí,
to srdce nebolí.

Kódování

- **Kódování** – převod informace do jiné reprezentace
- **Příklady:**
  - abeceda (text → znaky)
    AHOJ
  - Morseova abeceda
    $\cdot - \,|\, \cdot\cdot\cdot\cdot \,|\, - - - \,|\, \cdot - - -$
  - Semaforová abeceda

    
  - Braillovo písmo

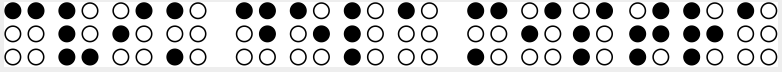
    
  - Kódování pro počítače – ASCII, UTF8

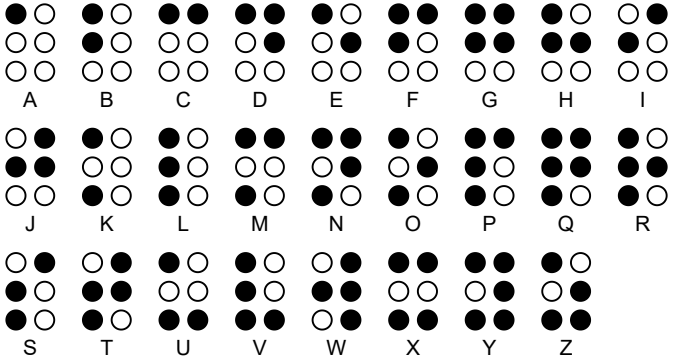
- **Vlastnosti**
  - bezpečný (bezpečnější než šifry)
  - náročnější na realizaci, zejména distribuce a utajení slovníku

Příklad

Příklad

Dekódujte zprávu psanou pomocí Braillova písma.





Šifrování

- **Holý text** – původní (čitelná) zpráva
- **Šifrovaný text** – nečitelná zpráva bez klíče
- **Klíč** – tajná informace řídící šifrování
- Šifra = algoritmus, který převádí holý text na šifrovaný

Cíl:

- zajistit **důvěrnost** informace

Historické principy

- **substituce** – záměna znaků
- **transpozice** – změna pořadí znaků

Poznámka:

- moderní šifry jsou matematicky mnohem složitější

Příklad

Jaký je rozdíl mezi kódováním a šifrováním?

Kódování vs. šifrování

- **Kódování**
  - cílem je reprezentace dat
  - příklad: UTF-8, Morseovka
- **Šifrování**
  - cílem je utajení informace
  - používá klíč
  - bez klíče není zpráva čitelná

Dělení šifer

- **Symetrické šifry:**
  - k rozšifrování se požívá symetrická operace k operaci zašifrování (používají stejný klíč)
  - Alice a Bob si soukromě domluví klíč
  - Alice tímto klíčem zprávu zašifruje, Bob jí pomocí stejného klíče dešifruje
  - nízká výpočetní náročnost
  - je potřeba sdílet klíč
- **Asymetrické šifry:**
  - používají se dva klíče – jeden pro šifrování, druhý pro dešifrování
  - Alice požádá Boba o šifrovací klíč, tím zašifruje zprávu a odešle jí
  - Bob dešifrovacím klíčem zprávu dešifruje
  - znalost šifrovacího klíče neumožňuje dešifrovat zprávu, je možné ho předat Alici veřejně (*veřejný klíč*)
  - dešifrovací klíč je tajný (*soukromý klíč*)
  - vyšší výpočetní náročnost
- Často se používají současně – asymetrické šifrování pro výměnu klíče pro symetrické (HTTPS, e-mailová korespondence)

Vědy o šifrách

- **Kryptografie:**
  - šifrovací postupy
  - to, co dělají Alice a Bob
  - přímá spojitost s matematikou
- **Kryptoanalýza:**
  - to, co dělá Ctirad
  - má k dispozici zašifrovaný text a snaží se ho rozšifrovat
  - na rozdíl od Boba nemá klíč, jak text rozšifrovat
  - neexaktní – frekvence písmen v textu, uhádnutí slova
- **Kryptologie:**
  - spojení kryptoanalýzy a kryptografie

Transpoziční šifry

- mění pouze pořadí písmen (znaky zůstávají stejné, jen se přeskupují)
- většina je založena na geometrickém postupu

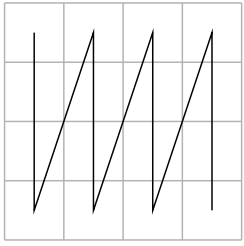
KDO SPI JAKO BY JEDL

- Kdos Pij Akoby Je dl
- KkDoObSyPjIeJdAl
- LDEJ YBOK AJIP SODK
- KOPJKBJD
  DS I AOYEL
- KPKJ DSIA OYEL OJBD

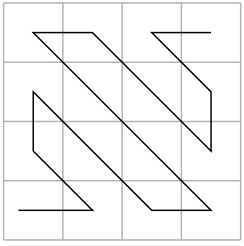
K P K J
D S I A O Y E L
O J B D

Transpoziční šifry

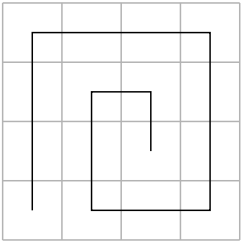
- stejná zpráva může mít více různých uspořádání
- klíčem je způsob přeskupení (např. mřížka)



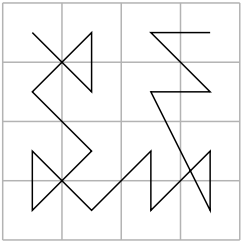
K	P	K	J
D	I	O	E
O	J	B	D
S	A	Y	L



O	B	D	L
S	K	Y	E
O	P	A	J
K	D	I	J



S	P	I	J
O	E	D	A
D	J	L	K
K	Y	B	O



K	O	D	L
S	D	J	E
J	P	K	B
I	A	O	Y

Kryptoanalýza

- **Hrubá síla** – vyzkoušíme všechny možnosti

Příklad

Kolik existuje různých možností pro zprávu délky m znaků?

Problém:

- počet možností rychle roste
- i pro krátké zprávy je výpočet náročný

Poznámka:

- u moderních šifer je hrubá síla prakticky nemožná

Transpoziční šifry dle klíče

- **Klíč:** určuje pořadí sloupců

K	L	I	C
K	D	O	S
P	I	J	A
K	O	B	Y
J	E	D	L

$\Longrightarrow$

C	I	K	L
S	O	K	D
A	J	P	I
Y	B	K	O
L	D	J	E

$\Longrightarrow$ SOKDAJPIYBKOLDJE

Kryptoanalýza

- SOKDAJPIYBKOLDJE
- neznáme délku klíče
- zpráva má délku 16 znaků $\Rightarrow$ klíč může mít délku 2, 4 nebo 8

- přepíšeme zprávu do n sloupců
- analyzujeme jednotlivé řádky

- **Pozorování:**
  - v českém jazyce tvoří samohlásky cca 40% znaků
  - nesprávné uspořádání tuto vlastnost poruší

- využijeme znalost jazyka (digramy, trigramy)

Proč tyto šifry dnes nepoužíváme?

- zachovávají statistické vlastnosti jazyka
- lze je prolomit analýzou textu
- malé množství možných klíčů

Důsledek:

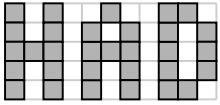
- nejsou bezpečné pro citlivá data

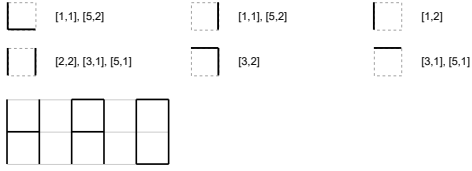
Příklad:

- zdravotnická dokumentace musí být chráněna moderní kryptografií

Grafické šifry

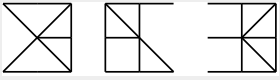
1, 1, 1, 2, 1, 2, 1, 2, 2, 1
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
3, 1, 3, 1, 1, 1, 1
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
1, 1, 1, 1, 1, 1, 1, 1, 2, 1





Příklad

Jak asi funguje následující grafická šifra?



Substituční šifry

- jednotlivá písmena nahrazuje jinými písmeny, či znaky
- pořadí znaků zůstává zachováno
- originální text, ŠIFROVANÝ TEXT
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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- **Obrácená abeceda** $25 - x$

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

  co nejde silou jde rozumem
  XL MVQWV HROLF QWV ILAFNVN
- **Lineární transformace** $(ax + b) mod 26$, a nesoudělné s 26
  $3x + 5$

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

Vlastnosti substitučních šifer

- zachovávají strukturu textu
- zachovávají frekvenci písmen

Důsledek:

- lze je prolomit pomocí statistické analýzy

Příklad:

- v češtině se často vyskytuje „E“, „A“, „O“
- nejčastější znak v šifře bude pravděpodobně „E“

Kryptoanalýza substitučních šifer

- **Frekvenční analýza**
  - analyzujeme četnost znaků
  - porovnáváme s přirozeným jazykem
- **Digramy a trigramy**
  - např. „ch“, „st“, „pr“
  - pomáhají rekonstruovat text
- **Závěr:**
  - substituční šifry nejsou bezpečné

Césarova šifra

- Substituční šifra
- Každý znak abecedy posune o určitý počet míst v abecedě (počet míst = klíč)
- Např. pro posun o 2: a $\rightarrow$ c, b $\rightarrow$ d, ...
  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 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
- PGJCBGLHNKPVWFQBKVC

Příklad

Jak můžeme tuto šifru zapsat matematicky?

Kryptoanalýza

Příklad

Kolik je možných zakódování zprávy?

Příklad

Pomocí metody hrubé síly dešifrujte zprávu:
ATCDRWHTCTYSGXKHTSGT
Naznačte, jak byste řešili tuto úlohu algoritmicky.

Vigenèrova šifra

- Substituční šifra, založená na Césarově
- Pro kódování se používá klíč – slovo
- každé písmeno se šifruje jiným posunem
- **Princip:**
  - klíč opakujeme pod zprávu
  - každý znak šifrujeme podle odpovídajícího písmene klíče
-

	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

...

- Klíč = "abc", zpráva = "ahoj"→ "AIQJ"

Příklad

Příklad

Je tato šifra bezpečnější než Césarova?

Proč je Vigenèrova šifra lepší?

- nepoužívá jeden posun (jako César)
- používá více různých posunů

Důsledek:

- frekvence písmen se „rozbije“
- jednoduchá frekvenční analýza nefunguje
- dlouho považována za „neprolomitelnou“

- **Ve skutečnosti:**
  - lze ji prolomit analýzou textu (složitější)
  - závisí na délce klíče

Shrnutí klasických šifer

- transpoziční šifry – mění pořadí znaků
- substituční šifry – mění znaky

- **Společný problém:**
  - zachovávají statistické vlastnosti jazyka
  - lze je prolomit analýzou textu

- **Důsledek:**
  - nejsou vhodné pro ochranu citlivých dat

Problém klasických šifer

- všechny klasické šifry jsme dokázali prolomit
- využívají vlastnosti přirozeného jazyka
- mají malý prostor klíčů

Otázka:

- Jak navrhnout šifru, kterou nelze prolomit?

Řešení:

- matematika + výpočetní složitost

Moderní kryptografie

- založená na matematice
- bezpečnost není založena na utajení algoritmu
- ale na **tajnosti klíče**

Princip:

- algoritmus je veřejný
- bezpečnost zajišťuje klíč

Kerckhoffsův princip:

- systém musí být bezpečný, i když útočník zná algoritmus

Symetrické šifry

- jeden klíč pro šifrování i dešifrování
- velmi rychlé

Příklad:

- AES (Advanced Encryption Standard)

Použití:

- šifrování disků
- databáze pacientů
- přenos velkých dat

Problém:

- jak bezpečně sdílet klíč?

AES – základní princip

- symetrická bloková šifra
- standard pro moderní šifrování

Vlastnosti:

- pracuje s bloky o velikosti 128 bitů
- používá klíče délky 128, 192 nebo 256 bitů
- provádí opakované transformace (tzv. kola)

Myšlenka:

- data se postupně „zamíchají“ a „přepočítají“

Jak funguje AES

- vstup: blok dat (128 bitů)
- výstup: zašifrovaný blok

Každé kolo obsahuje:

- **SubBytes** – nahrazení hodnot (substituce)
- **ShiftRows** – posun řádků
- **MixColumns** – promíchání dat
- **AddRoundKey** – přidání klíče

Počet kol:

- 10 (pro 128bitový klíč)

AES – malý ilustrační příklad

Zjednodušení:

- místo 128 bitů použijeme 4 čísla (2×2 matice)
- hodnoty 0–9 (místo 0–255)

Vstup:

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

1. SubBytes (nahrazení):

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

(použijeme jednoduchou tabulku nahrazení)

AES – malý ilustrační příklad

2. ShiftRows (posun řádků):

- každý řádek se posune doleva o jiný počet pozic
- v našem zjednodušeném příkladu:
  - 1. řádek: neposouváme
  - 2. řádek: posun o 1

$$\begin{pmatrix} 4 & 7 \\ 2 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & 7 \\ 5 & 2 \end{pmatrix}$$

3. MixColumns (promíchání):

- každý prvek je lineární kombinací všech prvků, např.

$$\begin{pmatrix} 4 & 7 \\ 5 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} (4+5) & (7+2) \\ (4+2) & (7+5) \end{pmatrix} = \begin{pmatrix} 9 & 9 \\ 6 & 2 \end{pmatrix} \pmod{10}$$

AES – malý ilustrační příklad

4. AddRoundKey:

Klíč:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 9 & 9 \\ 6 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 7 & 3 \end{pmatrix} \pmod{10}$$

Výsledek po 1 kole

Intuice AES

- kombinace:
  - substituce (jako klasické šifry)
  - permutace (přeskupení dat)

- opakování těchto kroků $\Rightarrow$ silné „zamíchání“

Důsledek:

- výstup nemá žádnou zjevnou strukturu
- nelze využít statistiku jazyka

Srovnání:

- klasické šifry $\rightarrow$ prolomitelné
- AES $\rightarrow$ prakticky neprolomitelné

Asymetrické šifry

- dva klíče:
  - veřejný (public key)
  - soukromý (private key)

Princip:

- veřejný klíč šifruje
- soukromý klíč dešifruje

Příklady:

- RSA
- ECC

Výhoda:

- není nutné tajně sdílet klíč

RSA – základní princip

- asymetrická šifra (dva klíče)
- veřejný klíč: šifrování
- soukromý klíč: dešifrování

Myšlenka:

- snadné: násobení velkých čísel
- obtížné: rozklad na prvočísla

Princip:

- vybereme dvě velká prvočísla p, q
- spočítáme $n = p \cdot q$
- vytvoříme dvojici klíčů

RSA – šifrování

- zprávu převedeme na číslo m
- zašifrování:

$$c = m^e \bmod n$$

- dešifrování:

$$m = c^d \bmod n$$

- (e, n) = veřejný klíč
- (d, n) = soukromý klíč
- e a d jsou navzájem inverzní modulo $\varphi(n) = (p-1)(q-1)$ (Eulerova funkce)

RSA – malý příklad

Zvolíme malá čísla (jen pro ilustraci):

- $p = 3$, $q = 11$
- $n = 33$

Klíče:

- veřejný: $(e = 3, n = 33)$
- soukromý: $(d = 7, n = 33)$

Zpráva: $m = 4$

$$c = 4^3 \mod 33 = 64 \mod 33 = 31$$

$$m = 31^7 \mod 33 = 4$$

Jak se to používá v praxi

- asymetrická kryptografie:
  - použije se pro výměnu klíče
- symetrická kryptografie:
  - použije se pro samotná data

Příklad: HTTPS

- výměna klíče pomocí RSA/ECC
- přenos dat pomocí AES

Hashovací funkce

- převod dat na krátký otisk (hash)
- např. SHA-256

Vlastnosti:

- jednosměrnost
- malá změna vstupu $\Rightarrow$ velká změna výstupu
- odolnost proti kolizím

Použití:

- kontrola integrity dat
- ukládání hesel

Digitální podpis

- zajišťuje:
  - autenticitu
  - integritu
  - nepopiratelnost

Princip:

- vytvoří se hash dokumentu
- hash se zašifruje soukromým klíčem

Ověření:

- pomocí veřejného klíče

Digitální podpis v praxi

- eRecept
- zdravotnická dokumentace
- komunikace s pojišťovnami

Příklad:

- lékař podepíše laboratorní výsledek
- příjemce ověří, že nebyl změněn

Certifikáty a důvěra

- jak víme, že veřejný klíč patří správné osobě?

Řešení:

- digitální certifikát
- certifikační autorita (CA)

Použití:

- HTTPS
- přístup do systémů

Co si odnést

- kódování $\neq$ šifrování
- klasické šifry nejsou bezpečné
- moderní kryptografie = práce s klíči
- kombinujeme více metod:
  - šifrování (AES)
  - asymetrie (RSA)
  - hash
  - digitální podpis

Cíl: chránit citlivá data (např. zdravotnická)