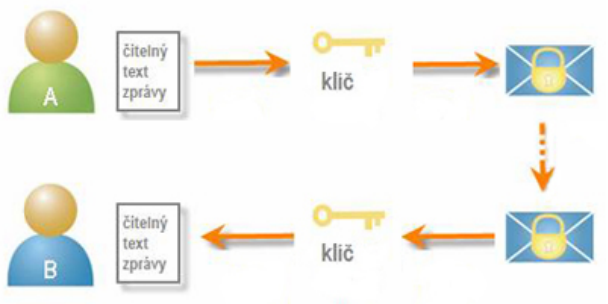
# Hillova šifra

Tomáš Růžička

20. 7. 2019

## 1 Úvod do šifrování

Cílem šifrování je převedení zprávy do nějaké vizuálně jiné podoby. Tedy změna prostého textu na zašifrovaný text. Převod probíhá pomocí předem domluveného klíče. Rozlišujeme symetrické a asymetrické šifrování. První možnost uvažuje použití stejného klíče pro zašifrování i dešifrování. Tedy dešifrování textu probíhá pouze opačně, než algoritmus použitý při zašifrování textu. Naopak asymetrické šifrování používá dva různé klíče - jeden k zašifrování a druhý k dešifrování. [1]




Obrázek 1: Symetrické šifrování

Mezi základní historické druhy symetrické kryptografie patří principy transpozice a substituce. Transpozici rozumíme jako permutaci znaků v textu. Takže zanechává počty výskytů jednotlivých znaků. Zatímco substituce pracuje na základě záměny znaků.

Následující odstavec je citován z [1]. *Šifrování na základě použití substituce rozlišuje monoalfabetické a polyalfabetické šifry. Obě tvoří abeceda s písmeny, symboly nebo jejich vzájemnými kombinacemi. V prvním případě se jedná o nejjednodušší formu, kde se podle jedné tabulky nahrazuje písmeno textu za symbol v zašifrovaném tvaru. "E" je vždy například "X" atp. Druhý případ dovoluje plynule měnit šifrovanou abecedu v průběhu kryptografie - jeden stejný znak otevřeného textu se šifruje několika různými symboly.*

Právě do symetrické kryptografie spadá i Hillova šifra pojmenovaná po americkém matematikovi jménem Lester S. Hill. [2]

# 2 Hillova šifra

Jedná se o substituční šifru z roku 1929, která je navíc polygrafická - nahrazuje m-tice znaků za jiné m-tice. Dle [2]: *Jedná se o první polygrafickou šifru, která umožňovala pracovat na více než třech symbolech zároveň.*

Šifra představuje aplikaci jednoduchého maticového počtu (násobení matic, výpočet inverzní matice a jejího determinantu), protože právě matice je šifrovacím klíčem.

K šifrování budeme používat písmena anglické abecedy. Každému písmenu přiřadíme celé číslo od 0 do 25 (protože anglická abeceda má 26 písmen). Každé číslo bude použito právě jednou. Samozřejmě lze použít libovolná abeceda nebo tabulka o libovolném počtu znaků. Takovou tabulku potřebujeme mít pro šifrování i následné dešifrování zprávy.

| A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| M | N | O | P | Q | R | S | T | U | V | W | X |
| 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| Y | Z | | | | | | | | | | |
| 24 | 25 | | | | | | | | | | |

Obrázek 2: Příklad přiřazení čísel

Matice reprezentující klíč (označme ji $\mathsf{K}$) musí splňovat následující 3 podmínky:

- Je čtvercová řádu $m$ (kde $m$ je přirozené číslo od 1 do délky šifrovaného textu).
- Determinant matice a počet znaků abecedy jsou nesoudělná čísla.
- Matice je regulární.

Tyto podmínky byly převzaty z [2]. Třetí bod je zřejmý, neboť tu budeme potřebovat pracovat s inverzní maticí k matici $\mathsf{K}$. Díky prvnímu bodu má pak násobení matic smysl a jelikož invertování matic zachovává čtvercový řád, tak má smysl i násobení inverzní maticí (více o rozměrech těchto matic v podkapitolách 2.1 a 2.2). Druhý bod je na první pohled nejasný. Vrátíme se k němu na konci druhé kapitoly potom, co ukážeme, jak probíhá samotné šifrování a dešifrování.

## 2.1 Šifrování

Následující část je parafrázována z [3]. Chceme-li zašifrovat nějaké slovo, postupujeme následovně. Vytvoříme sloupcový vektor složený z čísel, která jsou přiřazena k písmenům v tabulce. Označme ho $\mathbf{x}$. Způsobem, že $i$-tá souřadnice vektoru je $i$-té písmeno (znak) v šifrovaném slově (pro $i$ od 1 do délky slova). Rozdělíme vektor na několik $m$-tic. Získáme tak několik $m$-složkových sloupcových vektorů. Pokud není délka šifrovaného slova dělitelná číslem $m$, musíme

text nejdříve rozšířit několika speciálními znaky, dokud ona délka není násobkem čísla $m$ (například vložení mezer nebo interpunkce).

Nyní k samotnému zašifrování použijeme klíč - matici $\mathsf{K}$. Zvolme $j$-tý $m$-složkový sloupcový vektor (pro $j$ od 1 do $\frac{délka\ šifrovaného\ slova}{m}$). Označme tento vektor $\mathbf{x}_j$. Provedeme-li $\mathsf{K}\mathbf{x}_j$ **mod** 26[1], dostaneme nový $m$-složkový sloupcový vektor $\mathbf{y}_j$. Tedy pro každé $j$ z výše napsané množiny platí $\mathsf{K}\mathbf{x}_j$ **mod** $26 \equiv \mathbf{y}_j$.

Nyní stačí výsledné vektory $\mathbf{y}_j$ složit dohromady v jeden sloupcový vektor. Označme ho $\mathbf{y}$. Opačným postupem, než jsme tvořili vektor $\mathbf{x}$, vytvoříme i slovo z vektoru $\mathbf{y}$. Tedy $i$-tá složka vektoru $\mathbf{y}$ představuje $i$-té písmeno v zašifrovaném slově.

## 2.2 Dešifrování

Dešifrování probíhá opačným procesem než šifrování. Nyní dostáváme na vstup zašifrovaný text a našim cílem je z něho vytvořit otevřený text. Opět nejprve vytvoříme sloupcový vektor, ve kterém budou čísla z tabulky (abecedy), pod kterými jsou jednotlivá písmena. Tento vektor označme $\mathbf{y}$. Rozdělme tento vektor na několik $m$-složkových sloupcových vektorů $\mathbf{y}_j$ (pro $j$ od 1 do $\frac{délka\ šifrovaného\ slova}{m}$).

Nyní opět pomocí stejného klíče $\mathsf{K}$ můžeme text dešifrovat. Využijeme inverzního vztahu ke vztahu popsanému v předchozí podkapitole. Tedy pro výše zmíněná $j$ platí $\mathbf{x}_j \equiv \mathsf{K}^{-1}\mathbf{y}_j$ **mod** 26. Zde jsme využili toho, že matice $\mathsf{K}$ musí být regulární. Z výsledných vektorů $\mathbf{x}_j$ můžeme sestavit výsledný sloupcový vektor $\mathbf{x}$.

Teď už jen zbývá najít v tabulce znaky, kterým náleží čísla ve vektoru $\mathbf{x}$ a popořadě je zapsat na výstup. Výsledkem je dešifrovaný text.

Nyní se podíváme na již zmíněný bod 2 z požadavků na matici $\mathsf{K}$. Nechť $b$ je determinant matice $\mathsf{K}$ a $m$ je počet znaků tabulky. Kdyby $b$ a $m$ byly soudělná čísla, tak může nastat problém ve zpětném dešifrování. Poté co vytvoříme vektor $\mathbf{y}$, tak provádíme modulo $m$ a následně v dešifrování pak násobíme prvky tohoto vektoru maticí inverzního klíče. Pak nám ale vyjdou jiná čísla ve vektoru $\mathbf{x}$, než byla složena z původního otevřeného textu. Proto dokážeme následující větu, která říká, že u nesoudělných čísel se to stát nemůže. Zachováme původní značení $b$ a $m$ jako výše.

**Tvrzení.** Nechť $b$ a $m$ jsou přirozená čísla a NSD($b$,$m$)=1. Potom pro všechna celá $c$ platí

$$\frac{c\ (mod\ m)}{b}\ (mod\ m) = \frac{c}{b}\ (mod\ m).$$

[1]Používáme anglickou abecedu s 26 písmeny. Lze použít i jiná tabulka s jiným počtem znaků.

**Důkaz.** Výraz upravíme na $\frac{c\ (mod\ m)\ -\ c}{b}\ (mod\ m) = 0$. Z toho $[\frac{1}{b}(c\ mod\ m\ -\ c)](mod\ m) = 0$. Výraz $(c\ mod\ m\ -\ c)$ můžeme přepsat jako $km$ pro nějaké celočíselné $k$. Pak tedy $\frac{km}{b}(mod\ m) = 0$, což platí, protože NSD($b$,$m$)=1.

Jestliže čísla nesoudělná nejsou, tak poslední tvar obecně naplatí, a neplatí tedy ani věta.

Determinant $b$ jsme v tvrzení dali do jmenovatele, protože při výpočtu inverzní matice dostaneme vždy determinant ve jmenovatelích prvků inverzní matice. Číslo $c$ představuje prvek vektoru $\mathbf{y}$.

Uvedená kriteria pro matici $\mathsf{K}$ platí, jestliže matice má jako prvky celá čísla. Můžeme rozšířit prvky matice na racionální čísla, ale musíme přidat předpoklad, že i všechny jmenovatele jsou nesoudělné s počtem prvků v tabulce. Jinak by nastala podobná situace jako výše, nebo by čísla vycházející ve vektoru $\mathbf{y}$ nebyla celá.

# 3 Příklady

Ukážeme si dva příklady, na kterých demonstrujeme šifrování a dešifrování slova. K oběma příkladům použijeme následující 28-znakovou tabulku s přiřazenými čísly.

| A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|
| 2 | 4 | 6 | 8 | 10 | 12 | 14 |
| **H** | **I** | **J** | **K** | **L** | **M** | **N** |
| 16 | 18 | 20 | 22 | 24 | 26 | 27 |
| **O** | **P** | **Q** | **R** | **S** | **T** | **U** |
| 1 | 3 | 5 | 7 | 9 | 11 | 13 |
| **V** | **W** | **X** | **Y** | **Z** | | **.** |
| 15 | 17 | 19 | 21 | 23 | 25 | 0 |

Obrázek 3: Zvolená tabulka písmen s přiřazenými čísly od 0 do 27

## 3.1 Příklad 1

**Zadání**: Zašifrujte text **AUTOBUS A AUTO** použitím klíče

$$\mathsf{K} = \begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix}$$

**Řešení**: Ověříme, že matice splňuje všechny tři podmínky z předchozí kapitoly. Matice $\mathsf{K}$ je čtvercová, je zřejmě regulární, a její determinant je 5, což tvoří spolu s počtem znaků v tabulce (což je 28) čísla nesoudělná. Matice splňuje podmínky, aby se stala klíčem šifrování.

Nyní vytvoříme sloupcový vektor $\mathbf{x}$, který bude složený z pořadových čísel písmen v textu, který máme zašifrovat.

$$\mathbf{x} = (2\ 13\ 11\ 1\ 4\ 13\ 9\ 25\ 2\ 25\ 2\ 13\ 11\ 1)^T$$

Řád matice je 2, tedy vektor $\mathbf{x}$ rozdělíme na 7 dvousložkových sloupcových vektorů. Každý takový vektor pak vynásobíme maticí $\mathsf{K}$. Dvousložkové vektory si tak můžeme napsat do nějaké matice, např. matice $\mathsf{B}$, a násobit rovnou celé matice.

$$\mathsf{B} = \begin{pmatrix} 2 & 11 & 4 & 9 & 2 & 2 & 11 \\ 13 & 1 & 13 & 25 & 25 & 13 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 2 & 11 & 4 & 9 & 2 & 2 & 11 \\ 13 & 1 & 13 & 25 & 25 & 13 & 1 \end{pmatrix} = \begin{pmatrix} 15 & 12 & 17 & 34 & 27 & 15 & 12 \\ 95 & 29 & 99 & 193 & 179 & 95 & 29 \end{pmatrix}$$

Z výsledné matice můžeme inverzním způsobem vyndat výsledné dvousložkové vektory a slít je do jednoho výsledného vektoru $\mathbf{y}$.

$$\mathbf{y} = (15\ 95\ 12\ 29\ 17\ 99\ 34\ 193\ 27\ 179\ 15\ 95\ 12\ 29)^T$$

Ještě je potřeba provést s prvky vektoru modulo počtem znaků v tabulce. Dostáváme tedy upravený vektor $\mathbf{y}$.

$$\mathbf{y} = (15\ 11\ 12\ 1\ 17\ 15\ 6\ 25\ 27\ 11\ 15\ 11\ 12\ 1)^T$$

Nakonec z tohoto vektoru složíme zašifrované slovo. Pouze poskládáme popořadě znaky z tabulky, kterým přísluší čísla ve vektoru $\mathbf{y}$. Výsledný zašifrovaný text je tedy **VTFOWVC NTVTFO**.

## 3.2 Příklad 2

**Zadání**: Dešifrujte slovo **KRKVZPZHGFIF**, které bylo zašifrováno klíčem

$$\mathsf{K} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

**Řešení**: Nejdříve opět ověříme, že daný klíč je korektně zadaný. Determinant matice je -1, což spolu s 28 jsou nesoudělná čísla. Matice je čtvercová. Nyní ověříme, že je regulární.

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & -1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & -0.5 \end{pmatrix}$$

Provedením Gaussovy eliminace jsme zjistili, že matice $\mathsf{K}$ je regulární. To je důležitou podmínkou pro existenci inverzní matice. Podle pořadových čísel v tabulce složíme sloupcový vektor $\mathbf{y}$.

$$\mathbf{y} = (22\ 7\ 22\ 15\ 23\ 3\ 23\ 16\ 14\ 12\ 18\ 12)^T$$

Vektor $\mathbf{y}$ rozdělíme na tři tří-složkové sloupcové vektory, protože matice $\mathsf{K}$ je řádu tři. Každý takový vektor pak vynásobíme zleva maticí inverzní k matici $\mathsf{K}$.

$$\mathsf{K}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & -2 \end{pmatrix}$$

Všechny tří-složkové vektory si můžeme zase napsat do nějaké matice, např. matice $\mathsf{B}$, a pak můžeme násobit rovnou celé matice.

$$\mathsf{B} = \begin{pmatrix} 22 & 15 & 23 & 12 \\ 7 & 23 & 16 & 18 \\ 22 & 3 & 14 & 12 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -1 & -2 \end{pmatrix} \begin{pmatrix} 22 & 15 & 23 & 12 \\ 7 & 23 & 16 & 18 \\ 22 & 3 & 14 & 12 \end{pmatrix} = \begin{pmatrix} 22 & 3 & 14 & 12 \\ 7 & 11 & 7 & 18 \\ -7 & 1 & 2 & -18 \end{pmatrix}$$

Výsledek zapíšeme do sloupcového vektoru $\mathbf{x}$ jako slití výsledných tří-složkových vektorů.

$$\mathbf{x} = (22\ 7\ -7\ 3\ 11\ 1\ 14\ 7\ 2\ 12\ 18\ -18)^T$$

Ještě zbývá provést se všemi složkami vektoru $\mathbf{x}$ modulo počtem prvků tabulky. Dostáváme upravený vektor $\mathbf{x}$.

$$\mathbf{x} = (22\ 7\ 21\ 3\ 11\ 1\ 14\ 7\ 2\ 12\ 18\ 10)^T$$

Nakonec jen sestavíme výsledné slovo podle pořadových čísel v tabulce. Tedy dešifrované slovo je **KRYPTOGRAFIE**.

# 4 Osobní hodnocení

Výhodou Hillovy šifry je, že zaměňuje hned několikatice symbolů najednou. Tedy nebude fungovat zkoumání na základě nejčastějších výskytů jednotlivých písmen v dané abecedě. To jsme mohli vidět v příkladu 2, kde například písmeno $R$ se nejdříve nahradilo samo sebou a druhé $R$ se nahradilo písmenem $H$. Tím se hned mění frekventovanost písmen ve slově. Pokud bude matice větší, bude šifra zaměňovat větší části textu.

Naopak nevýhodou této šifry je fakt, že pokud získáme otevřený text a k němu odpovídající šifrovaný text, tak spolu s tabulkou písmen (resp. abecedou) můžeme jako soustavu lineárních rovnic spočítat klíč, pomocí kterého se zrovna šifruje. Ukážeme na příkladu.

Otevřený text: **CERVENEC**
Zašifrovaný text: **HMKRSUHC**
Tabulka abecedy jako v příkladech 2.1 a 2.2. Budeme předpokládat, že známe

řád matice - v našem příkladu to bude 2.
Potom tedy víme, že šifra převádí dvojice na další dvojice. Takže si můžeme vzít například CE $\rightarrow$ HM a EN $\rightarrow$ SU, což zapíšeme vektorově pomocí hledaného klíče $\mathsf{K}$ jako $\mathsf{K}(6\ 10)^T = (16\ 26)^T\ (mod\ 28)$ a $\mathsf{K}(10\ 27)^T = (9\ 13)^T\ (mod\ 28)$. To můžeme zapsat pomocí matic

$$\mathsf{K}\begin{pmatrix} 6 & 10 \\ 10 & 27 \end{pmatrix} = \begin{pmatrix} 16 & 9 \\ 26 & 13 \end{pmatrix}\ (mod\ 28)$$

Z toho pak už jednoduše (soustavou rovnic nebo pomocí inverzní matice) lze spočítat, že

$$\mathsf{K} = \begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix}.$$

Jestliže řád matice $\mathsf{K}$ neznáme, tak musíme vyzkoušet více možností. Tolik, kolik má délka šifrovaného textu dělitelů.

Podle [4] lze prolomit klíč na základě znalosti šifrovaného textu a toho, že v původním otevřeném textu se objevuje nějaké slovo, popř. část slova. Potom lze zkoušet těchto několik písmen přiřadit k písmenům šifrovaného textu od začátku do konce (tzn. posouvat pouze písmena jedním směrem). Ale opět se musí předpokládat, že známe řád matice - kvůli počtu písmen, které přiřazujeme.

Jedno z možných vylepšení může být kombinace Hillovy šifry s nějakou další substituční šifrou (například jednoduchá substituce nebo Caesarova šifra). Potom se rozluštění šifry stává mnohem složitější. K šifrování bude tedy potřeba mít dva klíče. Jeden v podobě matice, pomocí které provedeme klasickou Hillovu šifru, a druhý klíč, který nám pouze zamění znaky už jednou zašifrovaného textu za nějaké jiné. Tím je otevřený text zašifrován dvakrát.

# 5 Reference


[**1** ] KOLÁČEK, Michal. Šifrování a biometrie pod drobnohledem. Svět Hardware [online]. 20.2.2009 [cit. 2019-07-16]. Dostupné z: https://www.svethardware.cz/sifrovani-a-biometrie-pod-drobnohledem/25723

[**2** ] Hillova šifra. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 30.3.2017 [cit. 2019-07-16]. Dostupné z: https://cs.wikipedia.org/wiki/Hillova_%C5%A1ifra

[**3** ] Matematické základy šifrování akódování: Úvod do šifrování [online]. In: [cit. 2019-07-17]. Dostupné z: https://kap.fp.tul.cz/attachments/article/251/RIF3_%C5%A1ifrov%C3%

A1n%C3%AD%20k%C3%B3dov%C3%A1n%C3%AD.pdf

[**4** ] Cryptanalysis of the Hill Cipher. Practical Cryptography [online]. [cit. 2019-09-20]. Dostupné z: http://practicalcryptography.com/cryptanalysis/stochastic-searching/cryptanalysis-hill-cipher/

# 6 Zdroje obrázků

**Obrázek 1** KOLÁČEK, Michal. Šifrování a biometrie pod drobnohledem. Svět Hardware [online]. 20.2.2009 [cit. 2019-07-16]. Dostupné z: https://www.svethardware.cz/sifrovani-a-biometrie-pod-drobnohledem/25723

**Obrázek 2** Hillova šifra. In: Wikipedia: the free encyclopedia [online]. San Francisco (CA): Wikimedia Foundation, 30.3.2017 [cit. 2019-07-19]. Dostupné z: https://cs.wikipedia.org/wiki/Hillova_%C5%A1ifra

**Obrázek 3** Vlastní tvorba