

# Prolomení Vigeněrový šifry

Martin Klíma

October 8, 2024

## 1 Rychlo teorie

K prolomení vigeněrový šifry je prvně zapotřebí získat délku klíče, který byl použit pro zašifrování. K tomu se dá využít Friedmanův test. Při znalosti délky klíče se zašifrovaný text rozdělí do skupin podle délky klíče (např. pro klíč délky 3 se rozdělí text do tří skupin). Poté je zapotřebí pro každou skupinu zjistit posun vůči původní abecedě (pomocí frekvenční analýzy) a tím určit posun písmena klíče. Při znalosti klíče stačí pouze text pouze dešifrovat.

## 2 Teorie

### 2.1 Friedmanův test / Index koincidence

Pro zjištění délky hesla lze využít Friedmanův test. Tento test spočívá v porovnávání indexu koincidence skupin vytvořených ze zašifrovaného textu s indexem koincidence jazyka, kterým je psaný otevřený text. Index koincidence je pravděpodobnost, že dvě náhodně vybraná písmena z textu budou stejná. Pro opravdu náhodný text je index koincidence  $1/26$ . Každý jazyk má svůj index koincidence, pro češtinu se odává hodnota 0.058.

Pro výpočet indexu koincidence skupiny písmen délky  $n$  se používá vzorec:

$$IK = \sum_{i=1}^{26} \frac{n_i(n_i-1)}{n(n-1)}$$

Abychom udělali průměr pro všechny skupiny, použijeme vzorec:  $IK_{prum} =$

$$\frac{1}{n} \sum_{i=1}^n IK_i$$

Index koincidence můžeme počítat od minimální odhadované délky až po maximální odhadovanou délku klíče. Také je možné přidat hodnoty možných délek získaných například pomocí Kasiského testu.

Hodnoty koincidence pro různé delky hesla následně porovnáme s indexem koincidence pro daný jazyk. Délka klíče je pravděpodobně ta hodnota, která se nejvíce shoduje.

## **2.2 Prolomení**

Při znalosti délky klíče máme k prolomení šifry více možných postupů

### **2.2.1 Útok hrubou silou**

Útok hrubou silou je praktický pouze u krátkých klíčů, jelikož jeho časová a výpočetní náročnost exponenciálně roste s délkou klíče.

### **2.2.2 Frekvenční analýza**

Způsob pomocí frekvenční analýzy využívá toho, že v každém jazyce se každé písmeno vyskytuje ve své vlastní frekvenci oproti čistě náhodnému textu, kde by frekvence písmen měla být velmi podobná.

Běžně jsou nejčastější samohlásky E a A.