Vysoké učení technické v Brně
Fakulta informačních technologií

KRY - Kryptografie
Projekt – Vigenerova šifra

Peter Horňák (xhorna14)
xhorna14@stud.fit.vutbr.cz
4. apríla 2021

1 Friedmanov test

Pre výpočet dĺžky kľúča pomocou Friedmanovho testu, bolo najprv potrebné vypočítať index koincidencie zašifrovaného textu. Následne je možné určiť dĺžku kľúča na základe indexov koincidencie náhodného textu (0.038) a anglického textu (0.065) a to pomocou nasledujúcej rovnice:

$$\frac{Kp-Kr}{Ko-Kr}$$

Pričom Kp značí index anglického textu, Kr index náhodného textu a Ko vyrátaný index zašifrovaného textu.[1]

2 Kasiského test

Pre výpočet dĺžky kľúča je potrebné nájsť v texte identické n-gramy. Pre začiatok program hľadá n-gramy o dĺžke 5, v prípade že neexistuje žiadny opakujúci sa n-gram tak extrahuje n-gramy o dĺžke n-1. V prípade, že neexistujú opakujúce sa n-gramy o dĺžke 3, tak Kasiskeho test vráti hodnotu kľúča 0.

Následne sa vyrátajú vzdialenosti medzi rovnakými n-gramami. Postupne od najčastejšej vzdialenosti sa rátajú všetky delitele danej vzdialenosti, ignorujúc čísla 1 a 2. Pre reprezentáciu delitelov sa používa štruktúra reprezentujúcu množinu. Následne sa vypočíta prienik množín delitelov dvoch vzdialeností a výsledok sa uloží do výslednej množiny, ktorá bude následne opäť porovnávaná. Takýmto spôsobom vznikne prienik medzi deliteľmi 10 najčastejších vzdialeností. Výsledok Kasiského testu je najväčšia hodnota z výslednej množiny.

[1] https://en.wikipedia.org/wiki/Vigenère_cipher#Friedman_test

3 Určenie dĺžky hesla

Pre určenie dĺžky hesla sa postupne rozdeľuje šifrovaný text na stĺpce, kde každý N-tý charakter textu je súčasťou výsledného stĺpca. Takýmto spôsobom vznikne K stĺpcov, pričom K reprezentuje skúmanú dĺžku hesla.
Pre každý stĺpec sa následne vypočíta index koincidencie a ich priemer. Ak sa priemer stĺpcov približuje k indexu koincidencie anglického textu, tak označíme K ako predpokladanú dĺžku hesla.

4 Určenie hesla

Pre určenie hesla využijeme stĺpce textu, ktoré sme vyrátali pri určovaní dĺžky. Pre každý jeden stĺpec je potrebné najprv spočítať frekvenciu každého znaku v danom stĺpci. Vieme, že ak je dĺžka kľúča správna, tak každý stĺpec bol vytvorený pomocou monoalfabetickej šifry. Na základe toho, pre každé možné posunutie abecedy vypočítame index podobný indexu koincidencie pomocou následnej rovnice:

$$\sum_{i=0}^{25} \frac{p_i * f_{i+g}}{n}$$

Kde f_i značí frekvenciu výskytu daného písmena v stĺpci, p_i predstavuje pravdepodobnosť výskytu daného písmena v anglickom texte a g značí aktuálne posunutie. Zo všetkých posunutí vyberieme ten, ktorého hodnota je najbližšia k indexu koincidencie anglického textu a určíme znak abecedy na danom indexe ako znak, ktorý je súčasťou kľúča. Po zopakovaní tohto výpočtu pre každý stĺpec dostaneme kľúč, ktorým bol text zašifrovaný