KSIF - Semestrálne zadanie 19

Ema Richnáková

V semestrálnom zadaní riešime spôsob lúštenia zašifrovaného textu (ZT) šifrovaného Vigenerovou šifrou s neznámym kľúčom. Číslo zadania je 19. Boli nám ponúknuté 2 úvahy, ako sa máme zamýšľať nad riešením:

1. Lúštenie ako Vigenere. Rozdeliť lúštenie na 2, časovo menej náročné operácie. Najprv získať prvú polovicu kľúča (prvé 4 znaky hesla) a následne druhú polovicu kľúča (zvyšné 4 znaky hesla), nezávisle od seba. Takýmto spôsobom je možné za pár minút zvládnuť lúštenie aj pre heslo dĺžky 8 znakov. Pri použití správneho polovičného kľúča (či už prvá, alebo druhá polovica), presne 50% textu bude správne dešifrovaná.

2. Lúštenie Cézarovských posunov. Rozdeliť lúštenie na 8, časovo menej náročných operácií. V takomto prípade je potrebné riešiť jednotlivé Cézarovské posuny individuálne pomocou brute-force a je možné za zlomok sekundy zvládnuť jednu čiastkovú substitúciu. Vyriešením každej čiastkovej substitúcie získavame presne 1/8 výsledného textu.

Vybrali sme si druhú možnosť riešenia po neúspešnom riešení pomocou Genetického algoritmu.

Program vieme spustiť s argumentom, ktorý nás odkazuje na textový súbor so zašifrovaným textom. Ak mu argument neudáme, bude chcieť čitať súbor s názvom `"19.txt"`. Ak súbor nevie prečítať, hodí `IOException` chybu. Po úspešnom prečítaní si uložíme ZT a z neho si spravíme list slov, v ktorom budeme mať taký počet slov, aký má byť dlhý kľuč. V našom prípade je to 8. Náš list si predstavme teda ako tabuľku s 8 riadkami a počet stĺpcov nám údáva počet písmen v ZT delené 8. Do tejto tabuľky zapisujeme ZT po stĺpcoch a teda výsledný list nám výjde so slovami, ktorého všetky písmená v pôvodnom texte sa nachádzali na určitej pozícii s posunom 8. Slová z takto zadefinovaného listu môžeme Cezárovou šifrou pomocou brute-force metódy ohodnocovať, aký Cézarovský posun slova je najviac vyhovujúci. Zisťujeme to cez hodnoty bigramov anglického jazyka (tieto hodnoty sú uložené v súbore `"\_bigrams"`), kedy tieto hodnoty porovnávame s hodnotami bigramov posunutého slova a teda ako blízko sme s našimi hodnotami bigramov k tým anglickým. Cézarovský posun s najlepším ohodnotením si uložíme pre naše slovo a tento posun prevedieme na posun písmena a toto písmeno si uložíme do nášho kľuča. Výsledný kľúč použijeme vo Vigenerovej šifre na dešifrovanie ZT.

Pre naše zadanie číslo 19 nám touto metódou vyšiel takýto kľúč **zcbzbuie** s takýmito ohodnoteniami:

```
score: 1.2686288947313895 shift: z
score: 1.203053479282874 shift: c
score: 1.2337283720117527 shift: b
score: 1.134314993643638 shift: z
score: 1.2428607359267445 shift: b
score: 1.2280150641901915 shift: u
score: 1.3019849011986735 shift: i
score: 1.2773022556602127 shift: e
```

A teda po dešifrovaní nám vyšiel takýto text:

numbersdatanotshownbmpcrichelutriationfractionsobservedintimelapsecine matographyclustersofsmallroundcellsformedwithinhcellprocessesoccasionallyext endedfromthembuttheseretractedminuteslaterfigaindividualcellsweremotileand oftenleftthefieldbuttheclustersremainedintactafterhafewcellswithafibroblastlike morphologycouldbeseenbeneathandattheedgesoftheclustersthefibroblastlikecell sweremuchlargerthantheinitialcellsandquitemobileextendingandretractingusual lyaboutabroadfixedcuporpseudopodbydaysasignificantportionofthecellsretained theirelongatedformandlookedlikethecellsinfigblargeroundcellswerealsopresentth usitappearsthatbmpcsinthecirculationwerepresentassmallroundmononuclearcell sandtheirsubsequentmorphologyandfunctionweredictatedbycultureconditionsce llnumbersinthebmpcrichelutriationfractionfrommlofnormalhumanbloodelutriati onfractionswereselectedforquantificationofbmpcsbasedoncellsizeintermediatebet weenlymphocytesandmonocytesandgranularityfacsthispopulationcomprisedless thanlymphocytesandmorethanmonocytesnineteenconsecutivesampleshadanaver agetotalcellnumberofsemofwhichsemweremonocytesasubpopulationestimatedas ofthestartingelutriationfractionswasjudgedtoconsistofbmpcsonthebasisoftheirm orphologytheirstrongadherencetoplasticorglassandtheirabilitytoproliferateindm emfcswithoutaddedgrowthfactorsieofthestartingelutriatedcellsrepresentstobmp csthereforeitislikelythatmlofnormalbloodwillhaveseveralthousandbmpcscultures wereestablishedwithcellsfromtheelutriationfractionsandproliferationwasmeasure dondays

Nevyšiel nám týmto riešením dokonale odšifrovaný text, ale je na toľko zrozumiteľný, že nelogické časti by sme si vedeli doplniť manuálne.