Vigenerova šifra

Martin Klíma

October 13, 2024

1 Výsledky

Repozitář:	`https://github.com/mugac/Vigenere`

CT:	`dejlogmnpnnravvncvtkrsegvbcjkrkejuakwejurvenzfvippvt`
	`adpskbtepskjnrcycpjrloebskvpejcbzskbttfsbpscpvvosbbp`
	`ifdkjmvyijuomblfkabpjvenraeuwolsegvbcjktfsbpscpvvosb`
	`ffveeibcvoambzlkekbkvoamjcvoequijjczmekfdvkiezcvtkvt`
	`trunfttzbklmtlsygpdcfsmfujuamzjvdejlaifplclzlagbrcbm`
	`votejdvnobsakjcbzpibvejskbtjmisfrrmnznskbtejmifzzned`
	`bpfmikjcbzmjzskfmvnzrmoqfnpnnrtvfcouoejpukfzzqocjtzd`
	`kpdhjurroayoukjhcbvfvskbtlkegseqjdvotifplclzlymscypl`
	`ezmrkeujnpnzrloepdrsnpnoihaefmafdmpubpmfsomzprslrnee`
	`ucvtkvsegvbcjkpoamscypllnotjvploeoejuoajvcbdrdejleif`
	`plclzlytfsbpjvaedfskszejmypsgpdrsskwidltvsagpdcfebpn`
	`fnitlytisfdirmnzdhrqocjtzdkpdhzodzlakprlkabpjvidgoaf`
	`cymbtvmezodvylzesbfhfsoqwoafieeeotvfcouztztkldizodvy`
	`smpbfeyzotvsnvuufecvozlsygbtijkmzsfdeipzmjnluydttruu`
	`dtvvuavloepmzdkpqaksiumejwekpvvcaelyupsbvpzoyksitftz`
	`keuoaefjsphruszdhjuakvsmftrtnvkvptszniwjnrocejmzqrzk`
	`mpwpfsomoaejsajnpnijuakzmrwecnidblpqoujlfcymbtvmzzki`
	`tjcyqouqrrieddhleoszvplaqvjvueqqodfrefnzakfvnvsomooj`
	`umvaiefjsphruszniroeadhlesznifcymbtvmirsecbtzwnvwymb`
	`zvoegseipzuflfwaejbfiakttmjnrqrzdpfqucbczniibnvaadfs`
	`koaepskjjvelfvhfeosfnzakrbpfepivmvsedwyjqeczcyaedjvz`
	`odvyuvlocpgzdkvttfqyafcvtkfppiptzoebueizmajnpnvptpvm`
	`ydaedjmdfnjjmvlocpgzdkpndcvzejkvncvtkfteumolioupbvsa`
	`ujmvaigbtebckoeaceqqetoeatitizvniebsmftv`

Délka klíče:	3
Klíč:	`abc`

OT:	Česko plným názvem Česká republika je stát ve střední Evropě. Samostatnost nabylo jako nástupnický stát Československa, předtím existovalo jako jedna ze dvou republik československé federace. Navazuje také na více než tisícileté dějiny české státnosti a kultury. Podle své ústavy je Česká republika parlamentní demokratický právní stát s liberálním státním režimem a politickým systémem založeným na svobodné soutěži politických stran a hnutí. Hlavou státu je prezident republiky, vrcholným a jediným zákonodárným orgánem je dvoukomorový parlament České republiky. Na vrcholu moci výkonné stojí vláda České republiky. Česko je země s tržním hospodářstvím, která podle ekonomických, sociálních a politických indikátorů, jako je HDP na obyvatele, index lidského rozvoje, index svobody tisku či index svobody internetu od cenzury, patří k vysoce rozvinutým státům světa. Ekonomicky patří dle Světové banky do skupiny třiceti jedna nejbohatších států světa s nejvyššími finančními příjmy v porovnání s jinými státy. Má velmi malý podíl obyvatel žijících pod prahem chudoby, vykazuje též poměrně nízkou nerovnost mezi nejbohatšími a nejchudšími obyvateli a relativně vyvážené přerozdělování bohatství napříč populací. Míra nezaměstnanosti je dlouhodobě nízká a pod průměrem vyspělých zemí. V indexu ekologické stopy je Česko oproti některým jiným vyspělým zemím menším ekologickým dlužníkem. Česko se dlouhodobě řadí mezi patnáct nejbezpečnějších zemí na světě.

2 Řešní

2.1 Délka klíče

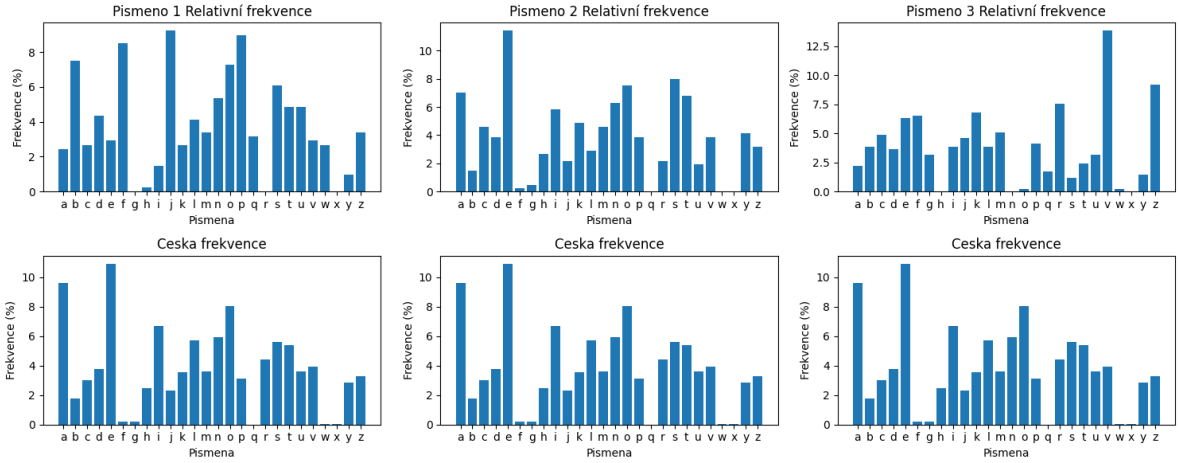
Pro zjištění délky klíče jsem využil index koincidence (pravděpodobnost, že dvě náhodně vybrané písmena z textu budou stejné). Aby se dala zjistit délka klíče, je zapotřebí text rozdělit do skupin podle odhadované délky klíče a pro každou skupinu spočítat index koincidence a udělat průměr všech skupin. Takto projdeme všechny možné hodnoty (většinou nějaké rozmezí + délky, které můžeme odhadnout například pomocí kasiskiho testu). Výsledky se následně porovnají s indexem koincidence pro jazyk, kterým je psán open text. Délka klíče je pravděpodobně ta hodnota, která se nejvíce shoduje.
Index koincidence v čj: 0.058
Index koincidence pro různé délky klíče:

2 :	0.045776882608718455
3 :	0.05811856786273908
4 :	0.0463213970495524
5 :	0.045428275949208494
6 :	0.058654984608098516
7 :	0.04508961982327601
8 :	0.04603564788574276
9 :	0.05815838800698883
10 :	0.04552157631571028
11 :	0.04557283982062743
12 :	0.059204264229963825
13 :	0.0458975794642088
14 :	0.045027054180483084
15 :	0.05708591668692233
16 :	0.04735231873389768
17 :	0.04692124454129199
18 :	0.058981166913372886
19 :	0.04549901852533432
20 :	0.04642693460250309

2.2 Prolomení klíče

Při znalosti délky klíče je možné text dešifrovat několika způsoby. Já jsem využil frekvenční analýzu, která 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. Text jsem rozdělil do skupin podle délky klíče((1,4,7)(2,5,8)(3,6,9)) a následně na každou skupinu provedl frekvenční analýzu, abych získal jednotlivé písmena klíče. Pro zjištění posunu jsem si frekvenci písmen znázornil do grafu a porovnal s grafem běžné distribuce písmen v čj. Spoléhat se na frekvenci písmen s vysokým výskytem může být v kratších textech nejisté, proto jsem se zaměřil na nízké výskyty.

Páry f-g a w-x, mezi kterými je ještě q mají skoro nulové výskyty, proto jsem je zvolil jako ideální patern pro zjištění posunu.
Grafy frekvence písmen:



Podle posunů jsem zjistil klíč: bar

2.3 Dešifrování

Když už mám klíč, tak stačí text dešifrovat. Pro dešifrování jsem použil stejný postup jako při šifrování, ale s opačným posunem. Pro každé písmeno jsem vzal korespondující písmeno klíče a provedl podle něho posun.