Prelomenie substitučnej šifry pomocou Metropolis-Hastings algoritmu

Tím 007

15. júna 2025

Cieľ zadania

Cieľom tejto úlohy bolo prelomiť substitučnú šifru, teda zistiť, ako boli znaky v texte poprehadzované, a pokúsiť sa text opäť správne poskladať. Na vstupe sme mali šifrovaný text, kde bola každému písmenu priradená iná náhrada podľa nejakej neznámej permutácie abecedy. Našou úlohou bolo túto permutáciu odhaliť a text rozlúštiť.

Na riešenie sme využili Metropolis-Hastings algoritmus, ide o metódu, ktorá pomocou pravdepodobností skúša rôzne možnosti a postupne sa približuje k tej, ktorá najlepšie zodpovedá reálnemu jazyku. V podstate teda prehľadáva priestor možných kľúčov a vyberá tie, ktoré dávajú zmysel.

Použitý prístup

1. **Príprava trénovacieho textu:** Najprv sme si vzali český román *Krakatit* a upravili ho do jednoduchej podoby, ostali len veľké písmená A–Z a miesto medzier sme použili podčiarkovník (_).

2. **Bigramová matica:** Z takto upraveného textu sme vytvorili prechodovú bigramovú maticu, ktorá zachytáva, aké dvojice písmen sa v texte vyskytujú najčastejšie. Táto matica potom slúžila ako referenčný model jazyka.

3. **Metropolis-Hastings algoritmus:** Algoritmus si náhodne skúšal rôzne poradia písmen a podľa toho, ako dobre sedeli na náš jazykový model, niektoré ponechal a iné zahodil. Postupne sa tak učil, čo dáva najväčší zmysel.

4. **Ukladanie výstupov:** Na konci sme pre každý šifrovaný súbor uložili rozlúštený text aj nájdený kľúč, aby sme si mohli skontrolovať, ako dobre to fungovalo.

Výsledky

- Algoritmus sme vyskúšali na viacerých šifrovaných textoch, ktoré sa líšili dĺžkou aj náročnosťou.

- Pri mnohých textoch sa nám podarilo dosiahnuť výsledky, ktoré sa veľmi podobali na pôvodný zašifrovaný text, najmä podľa hodnoty plausibility.

- Väčšina výstupov bola dobre čitateľná a niektoré dešifrované verzie boli takmer úplne správne, len pár písmen bolo občas prehodených.

Ladenie a úpravy počas vývoja

Na začiatku sme všetko stavali podľa pseudokódov zo zadania. Spočiatku sa nám však dešifrovať texty nedarilo, výstupy nedávali žiadny zmysel. Preto sme si museli trochu upraviť výpočet vierohodnosti bigramov. Po tejto zmene sa výsledky citeľne zlepšili, viaceré texty boli úspešne dešifrované, aj keď sa občas pomiešali dve alebo tri písmená. Najlepšie sa nám darilo pri dlhších textoch, kde bolo viac informácií na uhádnutie správneho kľúča.

Testovanie rôznych jazykových modelov

Bigramovú maticu sme vytvorili hlavne z románu *Krakatit*, no pre zaujímavosť sme vyskúšali aj iné texty, napríklad *Bibliu* a *Švejka*. Výsledky boli celkom podobné, čo ukazuje, že náš postup nie je až tak závislý od konkrétneho korpusu. Keďže ale niektoré šifrované súbory pochádzali práve z Krakatitu, nakoniec sme sa rozhodli použiť na finálne dešifrovanie práve jeho bigramovú maticu.

Porovnanie výsledkov počas iterácií

Počas behu algoritmu sme si priebežne zaznamenávali hodnoty plausibility, teda ako dobre momentálny kľúč sedí na náš jazykový model. Bolo vidieť, že po určitom počte krokov sa tieto hodnoty ustálili, čo naznačuje, že algoritmus sa ustálil na nejakej dobrej permutácii.

Export výstupov

Na základe zadania boli výsledky automaticky ukladané vo formáte:

- `text_{dlzka_textu}_sample_{id}_plaintext.txt`
- `text_{dlzka_textu}_sample_{id}_key.txt`

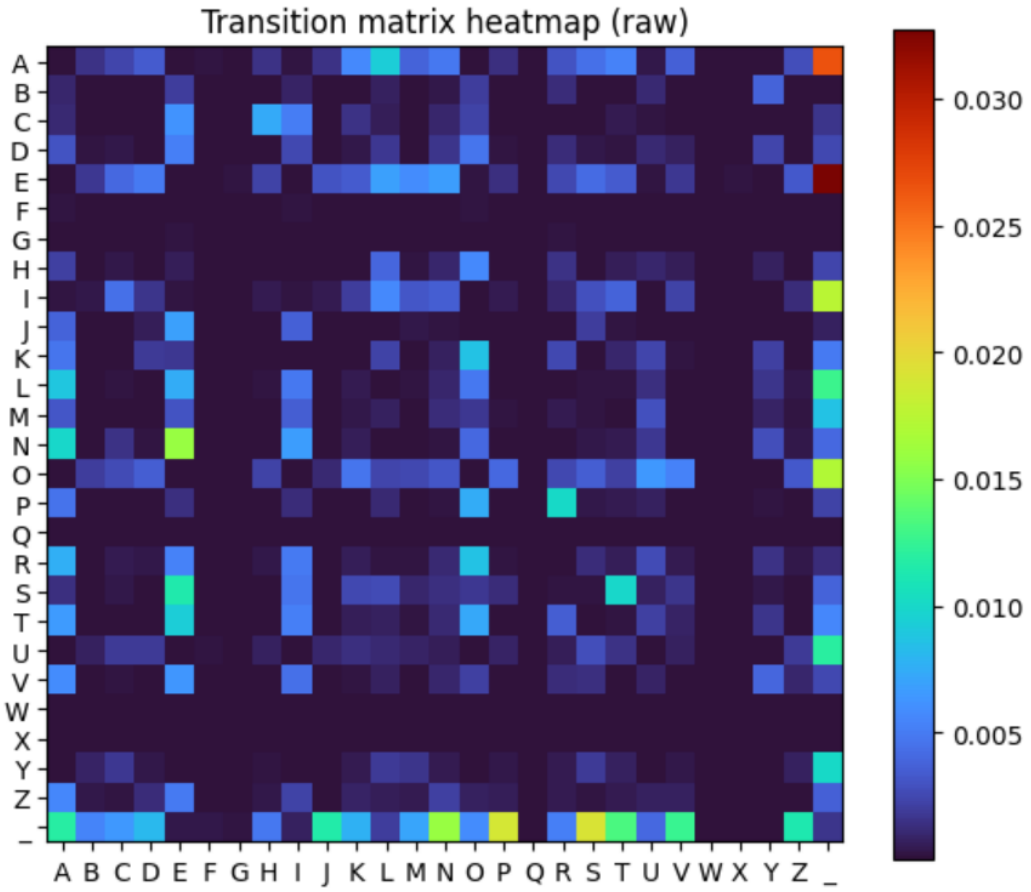
Skript pre každý zašifrovaný vstup automaticky spustil 20 000 iterácií a uložil si výsledný dešifrovaný text aj kľúč do samostatných súborov. Vďaka tomu sa dá celý postup jednoducho zopakovať a neskôr ďalej analyzovať, ak by bolo treba niečo upraviť alebo porovnať výsledky.

Ukážka dešifrovaného textu

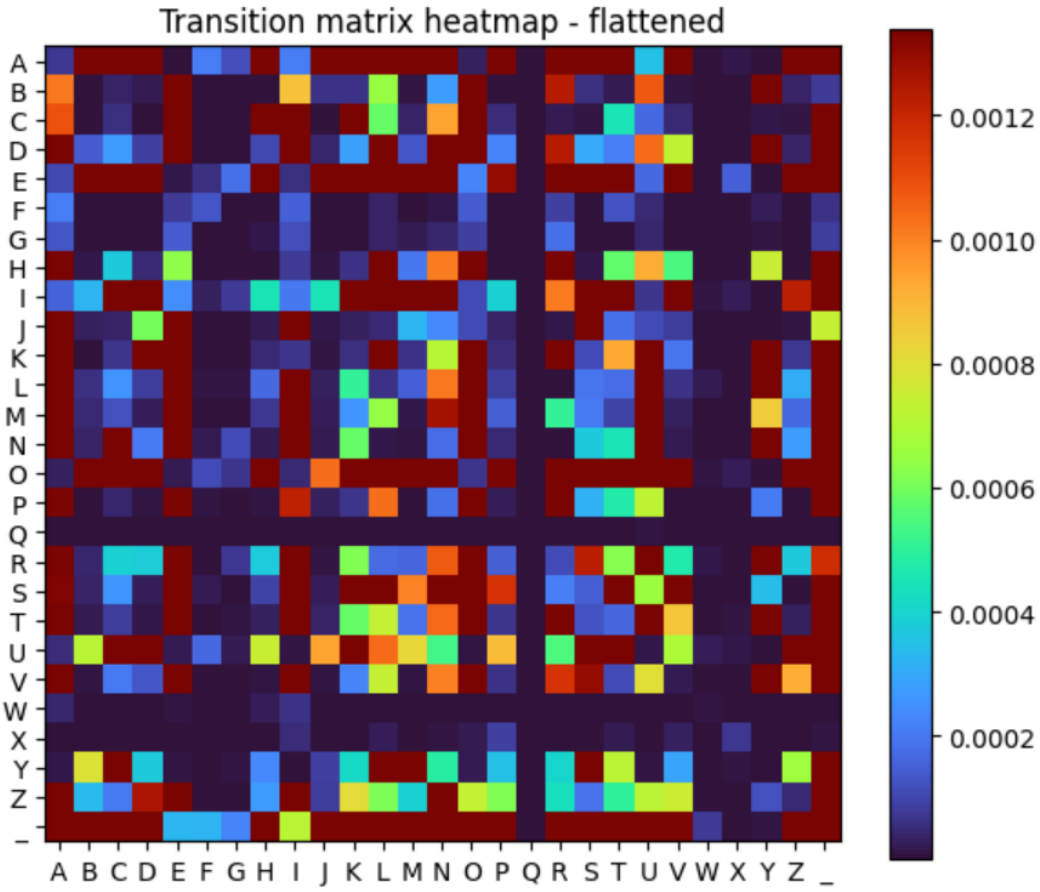
```
_VOZEM_DO_NEHO_A_ZAS_MNE_BEZI_DO_CESTY__ZACHVELA_SE_TAK_KUDY_VPRAVO_NEBO_VLEVO_
TEDY_JE_KONEC_OTEVREL_DVIRKA_VYSKOCIL_Z_VOZU_A_POSTAVIL_SE_PRED_KOLA_JED_REKL_
CHRAPTIVE_POJEDES_PRESE_MNE_UJELA_S_VOZEM...
```

Výstup má čitateľnú štruktúru a zodpovedá očakávanej štylistike češtiny podľa trénovacieho korpusu (Krakatit).

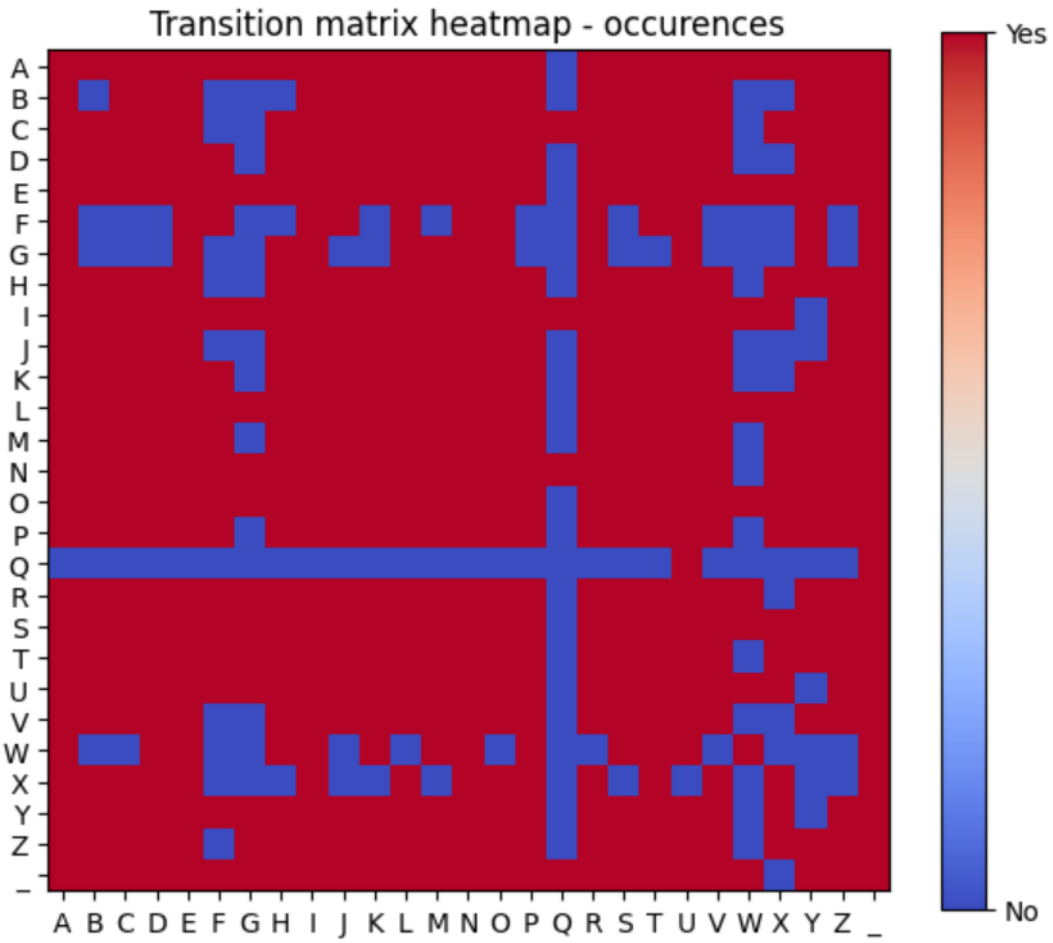
Vizualizácia prechodovej matice



Obr. 1: Základná heatmapa prechodovej (bigramovej) matice



Obr. 2: Zobrazenie matice po orezaní extrémnych hodnôt



Obr. 3: Binárne zobrazenie prítomnosti bigramov v texte

Záver

Zvolený prístup sa ukázal ako funkčný a pri viacerých testoch dokázal úspešne prelomiť substitučnú šifru. Aj bez znalosti pôvodného kľúča sa podarilo získať čitateľné a zmysluplné výstupy. Metropolis-Hastings algoritmus v kombinácii s bigramovým modelom fungoval spoľahlivo najmä pri dlhších textoch a ukázal, že sa dá využiť aj pri ďalšej automatizovanej analýze podobných šifrovaných údajov.