**Izveštaj – Druga lab vežba**

**Rabin-Karp algoritam:**

- Rabin-Karp algoritam koristi tehniku heširanja (rolling hash funkcija) kako bi brzo identifikovao potencijalna podudaranja izmedju stringa i zadatog podstringa. Rabin-Karp je u proseku efikasan, ali može biti osetljiv na određene scenarije (npr. ukoliko postoji dosta kolizija (spurious hits) – kada se hash kod reči poklapa sa hash kodom reči koja se traži u stringu međutim one nisu iste, to vidimo kada krenemo da proveravamo reč po reč).

U najgorem slučaju Rabin-Karp algoritam ima linearnu vremensku složenost, pa je efikasan i u tim scenarijima.

Best-Case Scenario: O(n-m+1) (kada uopšte nema kolizije)

Worst-Case Scenario: O(n\*m) (svaki put se poklapa hash kod ali se stringovi ne poklapaju)

**Ascii:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Broj reci u stringu | ASCII(Podstring 5) | ASCII(Podstring 10) | ASCII(Podstring 20) | ASCII(Podstring 50) |
| 100 | 4.6049 ms - 6.5075 ms | 7.7118 ms - 10.4226 ms | 7.6584 ms - 10.8723 ms | 8.8587 ms - 12.5417 ms |
| 1000 | 6.5107 ms - 15.7499 ms | 9.5642 ms - 16.2819 ms | 13.0506 ms - 21.646 ms | 11.9817 ms - 16.3766 ms |
| 10 000 | 9.1503 ms - 16.4507 ms | 10.5855 ms – 15.57ms | 11.6121 ms - 15.7594 ms | 9.9825 ms - 16.5605 ms |
| 100 000 | 9.6671 ms - 20.6666 ms | 10.086 ms - 31.2488 ms | 15.5922 ms - 23.6287 ms | 16.1371 ms - 31.6785 ms |

**HexaDecimal:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Broj reci u stringu | HEXA(Podstring 5) | HEXA(Podstring 10) | HEXA(Podstring 20) | HEXA(Podstring 50) |
| 100 | 3.4029 ms - 7.4253 ms | 5.0257 ms - 10.1266 ms | 6.9572 ms - 10.8715 ms | 7.5078 ms - 14.8142 ms |
| 1000 | 8.5214 ms - 10.221 ms | 7.3034 ms - 11.8499 ms | 10.1406 ms - 13.4892 ms | 11.4171 ms - 15.9347 ms |
| 10 000 | 8.0164 ms - 16.0724 ms | 15.8905 ms - 16.0807 ms | 15.98 ms - 16.4181 ms | 15.7069 ms -20.1235 ms |
| 100 000 | 15.007 ms - 20.6389 ms | 15.6923 ms - 23.8777 ms | 16.586 ms – 31.173 ms | 16.1098 ms – 32.13 ms |

**Levenshtein algoritam:**

- Levenshtein algoritam je algoritam koji se koristi za izračunavanje mere za razliku (poklapanje) dve string sekvence. Može biti zahtevan za duže stringove zbog veličine matrice koja se mora izračunati (zato treba i više vremena pri izračunavanju u odnosu na Rabin-Karp).

Best-Case Scenario: O(1) (kada su dva stringa već potpuno jednaka, svi elementi matrice su već 0)

Worst-Case Scenario: O(m\*n) (dva stringa su potpuno različita, što zahteva veliki broj insert, delete i replace funkcija, m i n su dužine ova dva stringa)

**Ascii:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Broj reci u stringu | ASCII(Podstring 5) | ASCII(Podstring 10) | ASCII(Podstring 20) | ASCII(Podstring 50) |
| 100 | 3.0072 ms - 7.0134 ms | 5.0325 ms - 10.0754 ms | 8.1236 ms - 14.5087 ms | 8.3437 ms - 14.9795 ms |
| 1000 | 7.0099 ms - 15.9103 ms | 7.6198 ms - 18.5001 ms | 15.7325 ms - 21.899 ms | 15.9507 ms - 25.0171 ms |
| 10 000 | 15.51 ms - 18.6303 ms | 16.1149 ms - 28.8804 ms | 21.7257 ms - 37.047 ms | 56.9707 ms -69.8772ms |
| 100 000 | 78.4275 ms - 94.0559 ms | 156.8651 ms - 173.1133 ms | 251.5852 ms - 315.9425 ms | 563.6698 ms - 611.2458 ms |

**HexaDecimal:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Broj reci u stringu | HEXA(Podstring 5) | HEXA(Podstring 10) | HEXA(Podstring 20) | HEXA(Podstring 50) |
| 100 | 6.689 ms - 9.9217 ms | 7.1228 ms - 10.1873 ms | 10.9303 ms - 13.1559 ms | 10.9677 ms - 15.8342 ms |
| 1000 | 15.6363 ms - 16.0309 ms | 15.8692 ms - 17.0542 ms | 15.8379 ms - 18.8688 ms | 16.0355 ms - 30.7318 ms |
| 10 000 | 15.0192 ms - 16.0207 ms | 15.7442 ms - 28.2119 ms | 22.1921 ms - 46.8919 ms | 52.4081 ms - 79.3881 ms |
| 100 000 | 78.6981 ms - 110.249 ms | 144.3312 ms - 161.6257 ms | 262.9132 ms - 280.3768 ms | 565.8824 ms - 630.8442ms |

Anđela Mia Dončov 18627