**ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH**

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN**





**Báo Cáo Bài Tập Lớn**

|  |  |  |
| --- | --- | --- |
| Sinh viên thực hiện: | | |
| STT | Họ tên | MSSV |
| 1 | Huỳnh Ngọc Công Danh | 19521322 |
| 2 | Trần Đình Khôi | 19521711 |

TP. HỒ CHÍ MINH – 7/2020

**Kết quả tổ hợp {1-grams,2-grams,3-grams}**

1 jrf1109.txt vs sra31.txt 4941

2 abf0704.txt vs edo26.txt 1822

3 abf0704.txt vs abf70402.txt 1639

4 catchmeifyoucan.txt vs tyc12.txt 1190

5 bef1121.txt vs edo14.txt 1182

6 catchmeifyoucan.txt vs hal10.txt 991

7 catchmeifyoucan.txt vs ecu201.txt 945

8 edo20.txt vs esv254.txt 621

9 sra126.txt vs sra42.txt 608

10 abf70402.txt vs edo26.txt 569

11 esv254.txt vs sra42.txt 536

12 esv254.txt vs sra126.txt 533

13 edo20.txt vs sra126.txt 533

14 edo20.txt vs sra42.txt 532

15 esv254.txt vs prz100.txt 514

16 abf0704.txt vs edo20.txt 510

17 edo20.txt vs prz100.txt 470

18 bef1121.txt vs sra126.txt 454

19 prz100.txt vs sra126.txt 437

20 bef1121.txt vs esv254.txt 432

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tổ hợp đơn vị so khớp** | **vector** | **binary search**  **tree** | **hashtable** | **map** |
| **{1-grams}** | 14.8167 | 2.68088 | 1.03773 | 1.35808 |
| **{2-grams}** | 16.7389 | 2.28962 | 1.39958 | 1.77182 |
| **{3-grams}** | 17.0196 | 2.78058 | 1.60327 | 2.01301 |
| **{1-grams,2-grams}** | 31.5111 | 4.69625 | 2.28738 | 1.77182 |
| **{1-grams, 2-grams, 3-grams}** | 49.4709 | 7.33164 | 3.69556 | 5.14021 |

**Nhận Xét:**

Qua bảng số liệu t nhận thấy được cấu trúc dữ liệu bảng băm có tốc độ chạy vượt trội so với các cấu trúc khác.Do chúng ta chỉ cần luu key của đối tượng mà chúng ta muốn sử dụng và từ đó chỉ cần 1 vòng lặp để tra ra nên có tốc độ vượt trội hơn hẳn tuy vậy tùy trường hợp mà chúng ta sẽ sử dụng cấu trúc dữ liệu sao cho hợp lý.

**Cấu hình máy:**

+ Ram 8gb

+ SSD 512gb

+ Window 10

+Chip intel i5-1.6ghz

**Các tài liệu tham khảo :**

**+** <https://nguyenvanhieu.vn/bang-bam-hash-tables/>

+ <https://khmt.uit.edu.vn/laptrinh/it003.k24/view_problem/7>

+<https://en.cppreference.com/w/cpp/chrono?fbclid=IwAR2gYIHhdAYZ6wuX1ulaw4x6AblVoDwZ8kZrAWny_J2BRJtZvrYgFig3FSE>

+<http://www.cplusplus.com/reference/cstring/strtok/>

+<http://www.cplusplus.com/reference/map/map/>

+<http://www.cplusplus.com/reference/vector/>

**TÀI LIỆU THAM KHẢO**

1. Họ và tên tác giả: Le Thi Ngoc Thanh
2. Tên trang web: <https://viblo.asia/p/tim-hieu-kien-truc-gitlab-BAQ3vVbZvbOr> (12/12/2019)

**PHỤ LỤC PHÂN CÔNG NHIỆM VỤ**

|  |  |  |
| --- | --- | --- |
| **STT** | **Thành viên** | **Nhiệm vụ** |
| 1 | Huỳnh Ngọc Công Danh | Tìm kiếm thông tin và soạn bài. |