**Họ và tên:** Trần Bá Trung **Mssv:** 20154000

**Báo cáo bài tập tuần môn: Big data**

**1) Đếm số thành phần liên thông**

Đếm số thành phần liên thông của đồ thị Gword bằng cách sử dụng

giải thuật duyệt cây BFS, duyệt BFS bắt đầu từ đỉnh đầu tiên: Mỗi lần duyệt BFS xác định được các đỉnh trong một thành phần liên thông, loại bỏ những đỉnh này khỏi danh sách các điểm cần duyệt BFS.

Tiếp tục thực hiện cho đến khi duyệt hết các đỉnh.

**Kết quả:**



Số thành phần liên thông là: **853**

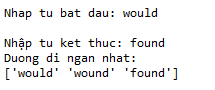
**2)** **Xác đinh đường đi ngắn nhất**

Xác đinh đường đi ngắn nhất của hai từ trên đồ thị Gword sử dụng giải thuật duyệt cây BFS:

Duyệt BFS bắt đầu với gốc tại từ bắt đầu , quá trình duyệt kết thúc khi gặp đỉnh kết thúc , trả về đường đi dựa vào thứ tự các đỉnh thăm trước đó của mỗi đỉnh

**Kết quả chương trình cài đặt:**

Chương trình yêu cầu nhập các từ bắt đầu và từ kết thúc để tiếp tục



**3) Cài đặt Girvan-Newman:**

**Kết quả chương trình cài đặt Girvan-Newman:**



* Cạnh có betweenness lớn nhất là: **chins-coins**
* Số betweenness lớn nhất là: **208275.41767595825**

**4)** **Hướng dẫn chạy chương trình:**

**B1:** Cho file WordNetwork.py và file data sgb-word cùng thư mục, vào terminal nhập: **py WordNetwork.py** (đối với windows) hoặc **python3** **WordNetwork.py** (đối với linux)

**B2:** Chương trình yêu cầu nhập các từ bắt đầu và từ kết thúc để tiếp tục, kết quả trả về đường đi ngắn nhất giữa hai từ.

**B3:** Chờ chương trình chạy hiển thị kết quả tính