

1. Lý do chọn số lượng đỉnh của đồ thị

- Đỉnh của đồ thị : 5033

- Lý do chọn:

+ Bài tập không sử dụng sql nên bọn em chọn số đỉnh đồ thị nhỏ lại cho quá trình crawl nhanh hơn.

+ Chọn số đỉnh là 5033 sẽ tránh những link nhiễu do chọn quá nhiều đỉnh sẽ có nhiều đỉnh nằm ngoài nội dung ban đầu. Khi số lượng nhiễu quá nhiều thì kết quả in ra sẽ không đẹp. Có thể không liên quan đến kết quả ban đầu. kết quả có thể là những link hot nhất hiện tại trên wiki

+ Ước lượng số bài liên quan đến chủ đề trên wiki chỉ có thể lớn nhất là 5000 bài

2. Giải thích kết quả

- Kết quả đưa ra gần 100 kết quả khá đẹp. Nội dung những link có rank cao nhất liên quan đến chính trị, pháp luật và báo chí.

- Mọi người khá quan tâm đến tình hình chính trị Việt Nam. Trong gần 100 page có rank cao nhất chiếm khá nhiều tên tuổi những lãnh đạo của nhà nước Việt Nam.