Part 2

We implemented the Random Surfer algorithm with the Damping factor according to the formula given in class. After building the graph we iterated over our vertex list until the PageRank change from one iteration to another was smaller than epsilon (a constant we sat to be 0.01) for all vertices.

Results for different Damping factors:

For Damping=0:

1. Cnn
2. Yahoo
3. Aol

For Damping = 0.1:

1. Cnn
2. Bbc
3. Nytimes

For Damping = 0.3:

1. Cnn
2. Bbc
3. Nytimes

Execution time (milliseconds): 60-200

For a Damping factor of 0.85 I tried initial PageRanks of 1,5 and 10 and the results were the same for all:

1. Cnn
2. Bbc
3. Nytimes

Part 4

1. We implemented the aggregation function as follows:

For each word we have a sorted Vector of the documents with the word’s TF score in them.

* If the OR operator is given, the score is calculated by the Maximum between TF scores of all words in all documents, which is then multiplied by that document’s PageRank.
* If the AND operator is given, the score is calculated by the average of