**How to run the program:**

**Input “python bonus.py” in command line. It will take about 10 minutes to return result.**

I used the 16 labels mentioned in <http://www.dmoz.org/> as topics.

Since the data set is about Mars, it is impossible to make a distinguishable label for each user according to tweet contents.

To capture the core idea of Topic-Sensitive PageRank, I adopt a simplified strategy---assign a label for each user randomly. My method is to store the 16 labels in ‘topicList’. For each userid, generate a random number in the range [0, 15], then select relevant label from ‘topicList’ for the userid.



The following work is similar to part 2. There are only two differences. One is the construction of teleporting matrix. Comparing Table 1 with Table 2, it is easy to find out a user can only teleport to another user with the same topic label now.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | userid1(N) | userid2(S) | userid3(G) | userid4(S) | userid5(N) |
| userid1(N) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| userid2(S) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| userid3(G) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| userid4(S) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| userid5(N) | 0.2 | 0.2 | 0.2 | 0.2 | 0.2 |
| N:News, S:Sports, G:Games | | | | | |

Table 1 original teleporting matrix

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | userid1(N) | userid2(S) | userid3(G) | userid4(S) | userid5(N) |
| userid1(N) | 0.5 | 0 | 0 | 0 | 0.5 |
| userid2(S) | 0 | 0.5 | 0 | 0.5 | 0 |
| userid3(G) | 0 | 0 | 1 | 0 | 0 |
| userid4(S) | 0 | 0.5 | 0 | 0.5 | 0 |
| userid5(N) | 0.5 | 0 | 0 | 0 | 0.5 |
| N:News, S:Sports, G:Games | | | | | |

Table 2 Topic-Sensitive teleporting matrix

The second difference is the construction of transition matrix. For example, if userid1(N) only mentions userid2(S) and userid4(N), the row for user1id in Topic-Sensitive transition matrix will be Table 3, not Table 4.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | userid1(N) | userid2(S) | userid3(G) | userid4(S) | userid5(N) |
| userid1(N) | 0 | 0 | 0 | 1 | 0 |

Table 3 Topic-Sensitive transition matrix for userid1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | userid1(N) | userid2(S) | userid3(G) | userid4(S) | userid5(N) |
| userid1(N) | 0 | 0.5 | 0 | 0.5 | 0 |

Table 4 original transition matrix for userid1

Program Result:

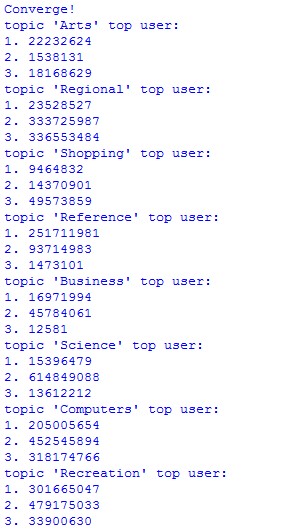


Figure 1 Topic-Sensitive PageRank Result

For convenience, I only show the top-3 userids in a few topics.