CSC443 Assignment 1 Part 3 Research Results

# Query Result

Total number of true friends: 21776094

Total number of distinct users: 11316811

|  |  |  |  |
| --- | --- | --- | --- |
| **Top 10 Celebrities** | | | |
| **User Id** | **In-degree** | **Out-Degree** | **Difference** |
| 5994113 | 564512 | 292 | 564220 |
| 7496 | 350885 | 6035 | 344850 |
| 1349110 | 341963 | 1472 | 340491 |
| 1629776 | 172231 | 2120 | 170111 |
| 8121005 | 155967 | 34 | 155933 |
| 2041453 | 152689 | 620 | 152069 |
| 797152 | 118826 | 74 | 118752 |
| 6623784 | 116002 | 183 | 115819 |
| 645019 | 107914 | 275 | 107639 |
| 3403 | 102877 | 2946 | 97931 |

# Summary

True = Total number of true friends

Total = Total number of distinct users

True/(Total2 – Total) = 0.000017%

Therefore, there’re just 0.000017% pairs of users on Twitters over all pairs of users are true friends. About 5% of all users are following the most popular celebrity who has UID 5994113. The question that I would like to ask is what the distribution of the difference of in-degree and out-degree would be like.