EXPERIMENT: Introducing peers with varying top-k strategies into an existing rational network  
  
DESCRIPTION: Given a network that has been established, see how various top-k strategies perform when added at different “ages” of the network. Ages refer to how old the network is (how many ticks has it been established for). The network will initially be populated with peers using a random top-k strategy. This experiment will be simplistic, documents and peers will have only one of two possible tags; this will limit the usefulness of some top-k strategies.

The top-k strategies being tested:

* document popularity
* peer popularity
* peer distance
* peer similarity
* follow similarity

The different of the ages of the network that will be tested are:

* Starting Network: the network is brand new (ticks = 0)
* Morning Network: the network has had few turns (ticks = 10)
* Noon Network: most peers have had a turn (ticks = 40)
* Evening Network: almost all peers are connected (ticks = 160)
* Complete Network: peers have made all or almost all connections (ticks = 1000)

Network Info:

* All peers:
  + Will have 1 of 2 tags.
  + The top-k will be of size 5, and will exclude files that have been in the top-k before.
  + A like strategy where, if the file shares at least one taste in common with the peer (and has not been liked before), it will be liked.
  + The top-m will be of size 3, and will rank peers based on the number of liked documents they share with the user. Peers that have no liked documents in common, are excluded. Peers that are already being followed are excluded.
  + A follow strategy that follows everyone who appeared in the top-m.
  + A payoff function which will increment the score of the user by one for every document in the top-k that shares at least one tag in common with the user, and has not been seen by the user before.
* The network will start with 20 peers who are using the random top-k strategy, and 50 documents. Exactly half of the starting peers and documents will have one tag, the other half will have the other tag.
* Only one peer will be added to each network for testing, it will have one of the 2 tag values

What will be measured and how?

* The new peer will be looped through 10 ticks in a row (the rest of the network will be frozen).
* After each tick, the new peers will have their utility analyzed.
* Results will be the average of 1000 simulations.

HYPOTHESIS:

The peer and document popularity strategies should be successful in the first turn, as they are not required to be connected to the existing network.

The peer distance, peer similarity, and follow similarity strategies should quickly take the lead after a turn or two, as they will isolate peers with the same tags as them.

RESULTS:

1. STARTING NETWORK:
   1. STRATEGY: Document Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles2** |
| 0 | 0 |
| 1 | 2.477 |
| 2 | 4.987 |
| 3 | 7.493 |
| 4 | 9.996 |
| 5 | 12.518 |
| 6 | 15.049 |
| 7 | 17.529 |
| 8 | 20.017 |
| 9 | 22.512 |
| 10 | 25 |
| **Grand Total** | **12.50709091** |

* 1. STRATEGY: Peer Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles3** |
| 0 | 0 |
| 1 | 2.497 |
| 2 | 5.027 |
| 3 | 7.483 |
| 4 | 9.953 |
| 5 | 12.452 |
| 6 | 15.04 |
| 7 | 17.532 |
| 8 | 20.035 |
| 9 | 22.471 |
| 10 | 25 |
| **Grand Total** | **12.49909091** |

* 1. STRATEGY: Peer Distance

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles4** |
| 0 | 0 |
| 1 | 2.522 |
| 2 | 5.063 |
| 3 | 7.512 |
| 4 | 10.086 |
| 5 | 12.57 |
| 6 | 15.004 |
| 7 | 17.48 |
| 8 | 20.009 |
| 9 | 22.448 |
| 10 | 25 |
| **Grand Total** | **12.51763636** |

* 1. STRATEGY: Peer Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles5** |
| 0 | 0 |
| 1 | 2.55 |
| 2 | 5.014 |
| 3 | 7.551 |
| 4 | 10.012 |
| 5 | 12.487 |
| 6 | 14.99 |
| 7 | 17.525 |
| 8 | 20.046 |
| 9 | 22.553 |
| 10 | 25 |
| **Grand Total** | **12.52072727** |

* 1. STRATEGY: Follow Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles6** |
| 0 | 0 |
| 1 | 2.508 |
| 2 | 4.986 |
| 3 | 7.504 |
| 4 | 9.997 |
| 5 | 12.541 |
| 6 | 15.032 |
| 7 | 17.585 |
| 8 | 20.045 |
| 9 | 22.525 |
| 10 | 25 |
| **Grand Total** | **12.52027273** |

1. EARLY NETWORK:
   1. STRATEGY: Document Popularity

|  |  |
| --- | --- |
| **Ticks** | **Average of Score** |
| 0 | 0 |
| 1 | 2.551 |
| 2 | 5.104 |
| 3 | 7.61 |
| 4 | 10.131 |
| 5 | 12.6 |
| 6 | 15.054 |
| 7 | 17.543 |
| 8 | 20.061 |
| 9 | 22.524 |
| 10 | 25 |
| **Grand Total** | **12.56163636** |

* 1. STRATEGY: Peer Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles3** |
| 0 | 0 |
| 1 | 2.44 |
| 2 | 5.36 |
| 3 | 8.333 |
| 4 | 10.956 |
| 5 | 13.485 |
| 6 | 15.924 |
| 7 | 18.286 |
| 8 | 20.582 |
| 9 | 22.825 |
| 10 | 25 |
| **Grand Total** | **13.01736364** |

* 1. STRATEGY: Peer Distance

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles4** |
| 0 | 0 |
| 1 | 2.523 |
| 2 | 6.1 |
| 3 | 9.676 |
| 4 | 12.605 |
| 5 | 15.01 |
| 6 | 17.19 |
| 7 | 19.258 |
| 8 | 21.207 |
| 9 | 23.137 |
| 10 | 25 |
| **Grand Total** | **13.79145455** |

* 1. STRATEGY: Peer Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles5** |
| 0 | 0 |
| 1 | 2.478 |
| 2 | 6.187 |
| 3 | 9.867 |
| 4 | 12.82 |
| 5 | 15.153 |
| 6 | 17.285 |
| 7 | 19.234 |
| 8 | 21.192 |
| 9 | 23.093 |
| 10 | 25 |
| **Grand Total** | **13.84627273** |

* 1. STRATEGY: Follow Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles6** |
| 0 | 0 |
| 1 | 2.498 |
| 2 | 5.736 |
| 3 | 8.875 |
| 4 | 11.618 |
| 5 | 13.978 |
| 6 | 16.245 |
| 7 | 18.45 |
| 8 | 20.622 |
| 9 | 22.854 |
| 10 | 25 |
| **Grand Total** | **13.26145455** |

1. NOON NETWORK:
   1. STRATEGY: Document Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles2** |
| 0 | 0 |
| 1 | 2.467 |
| 2 | 4.94 |
| 3 | 7.41 |
| 4 | 9.948 |
| 5 | 12.429 |
| 6 | 14.968 |
| 7 | 17.459 |
| 8 | 19.982 |
| 9 | 22.492 |
| 10 | 25 |
| **Grand Total** | **12.46318182** |

* 1. STRATEGY: Peer Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles3** |
| 0 | 0 |
| 1 | 2.51 |
| 2 | 5.457 |
| 3 | 8.511 |
| 4 | 11.593 |
| 5 | 14.316 |
| 6 | 16.679 |
| 7 | 18.747 |
| 8 | 20.696 |
| 9 | 22.679 |
| 10 | 25 |
| **Grand Total** | **13.28981818** |

* 1. STRATEGY: Peer Distance

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles4** |
| 0 | 0 |
| 1 | 2.505 |
| 2 | 7.355 |
| 3 | 12.344 |
| 4 | 17.303 |
| 5 | 21.69 |
| 6 | 23.221 |
| 7 | 23.697 |
| 8 | 24.121 |
| 9 | 24.57 |
| 10 | 25 |
| **Grand Total** | **16.52781818** |

* 1. STRATEGY: Peer Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles5** |
| 0 | 0 |
| 1 | 2.436 |
| 2 | 7.293 |
| 3 | 12.286 |
| 4 | 17.28 |
| 5 | 21.746 |
| 6 | 23.225 |
| 7 | 23.695 |
| 8 | 24.129 |
| 9 | 24.593 |
| 10 | 25 |
| **Grand Total** | **16.51663636** |

* 1. STRATEGY: Follow Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles6** |
| 0 | 0 |
| 1 | 2.461 |
| 2 | 7.349 |
| 3 | 12.342 |
| 4 | 17.323 |
| 5 | 21.634 |
| 6 | 22.996 |
| 7 | 23.514 |
| 8 | 24.002 |
| 9 | 24.498 |
| 10 | 25 |
| **Grand Total** | **16.46536364** |

1. EVENING NETWORK:
   1. STRATEGY: Document Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles2** |
| 0 | 0 |
| 1 | 2.416 |
| 2 | 4.92 |
| 3 | 7.327 |
| 4 | 9.793 |
| 5 | 12.27 |
| 6 | 14.773 |
| 7 | 17.354 |
| 8 | 19.881 |
| 9 | 22.421 |
| 10 | 25 |
| **Grand Total** | **12.37772727** |

* 1. STRATEGY: Peer Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles3** |
| 0 | 0 |
| 1 | 2.514 |
| 2 | 7.377 |
| 3 | 12.354 |
| 4 | 17.339 |
| 5 | 22.32 |
| 6 | 24.847 |
| 7 | 24.945 |
| 8 | 24.965 |
| 9 | 24.985 |
| 10 | 25 |
| **Grand Total** | **16.96781818** |

* 1. STRATEGY: Peer Distance

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles4** |
| 0 | 0 |
| 1 | 2.503 |
| 2 | 7.448 |
| 3 | 12.448 |
| 4 | 17.448 |
| 5 | 22.448 |
| 6 | 24.945 |
| 7 | 25 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.02181818** |

* 1. STRATEGY: Peer Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles5** |
| 0 | 0 |
| 1 | 2.499 |
| 2 | 7.46 |
| 3 | 12.46 |
| 4 | 17.46 |
| 5 | 22.46 |
| 6 | 24.961 |
| 7 | 25 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.02727273** |

* 1. STRATEGY: Follow Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles6** |
| 0 | 0 |
| 1 | 2.558 |
| 2 | 7.514 |
| 3 | 12.511 |
| 4 | 17.511 |
| 5 | 22.511 |
| 6 | 24.953 |
| 7 | 24.997 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.05045455** |

1. COMPLETE NETWORK:
   1. STRATEGY: Document Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles2** |
| 0 | 0 |
| 1 | 2.506 |
| 2 | 5.022 |
| 3 | 7.505 |
| 4 | 10.037 |
| 5 | 12.52 |
| 6 | 15.034 |
| 7 | 17.554 |
| 8 | 20.043 |
| 9 | 22.528 |
| 10 | 25 |
| **Grand Total** | **12.52263636** |

* 1. STRATEGY: Peer Popularity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles3** |
| 0 | 0 |
| 1 | 2.519 |
| 2 | 7.464 |
| 3 | 12.462 |
| 4 | 17.462 |
| 5 | 22.462 |
| 6 | 24.943 |
| 7 | 24.998 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.02818182** |

* 1. STRATEGY: Peer Distance

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles4** |
| 0 | 0 |
| 1 | 2.455 |
| 2 | 7.391 |
| 3 | 12.389 |
| 4 | 17.389 |
| 5 | 22.389 |
| 6 | 24.934 |
| 7 | 24.998 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **16.995** |

* 1. STRATEGY: Peer Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles5** |
| 0 | 0 |
| 1 | 2.474 |
| 2 | 7.423 |
| 3 | 12.419 |
| 4 | 17.419 |
| 5 | 22.419 |
| 6 | 24.945 |
| 7 | 24.996 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.00863636** |

* 1. STRATEGY: Follow Similarity

|  |  |
| --- | --- |
| **Row Labels** | **Average of sum [score] of profiles6** |
| 0 | 0 |
| 1 | 2.467 |
| 2 | 7.419 |
| 3 | 12.419 |
| 4 | 17.419 |
| 5 | 22.419 |
| 6 | 24.952 |
| 7 | 25 |
| 8 | 25 |
| 9 | 25 |
| 10 | 25 |
| **Grand Total** | **17.00863636** |

INITIAL COMMENTS:

* Document Popularity
  + Results are what is expected from random. A possible reason for this is that:
  + Because documents can only have 1 of 2 tags (with exactly even chances of having either), meaning that every document has the same chance of becoming “popular”.

**Expected to improve when some documents are more likable than others**

* Peer Popularity
  + Behaves slightly better than random, improving as the network becomes older. A possible reason for this is that:
  + Because the peers are evenly divided, the peer being tested will tip the balance of the network so that the group with similar tags will be favored. This produces better results as the network grows older since peer popularity becomes more uniform (the balance is tipped by the tested peer more reliably)  
      
    **Expected to worsen when peer groups are unevenly divided**
    - **May improve relatively when some peers have had more turns than others**
* Peer Distance
  + Is effective

**Expected to worsen when peers have differing sets of tastes**

* + - **Improves the more the sets overlap**
* Peer Similarity
  + Is effective

**Expected to worsen when peers have differing sets of tastes**

* + - **Improves the more the sets overlap**
* Follow Similarity
  + In the early network, during its first behaves a little worse than Peer Distance and Peer Similarity. This is probably because more connections between peers are required, and connections between peers are harder to form than connections between files (immediately because topm is smaller than topk, possibly because follows are dependent on likes already being formed)
  + In the early and noon networks, Follow Similarity drops off a little (in the early network, it falls back to basically being Random). This is probably because it exhausts the peers that have formed follow links, making following results be basically random. And since it has already had some early success, there will be more documents that do not match its taste than do… so a random strategy will be more likely to return document that are not liked from then on.  
      
    **Expected to improve when peers link more between each other than with files**

T-Test comparisons  
1-2, 1-3, 1-4, 1-5  
2-3, 2-4, 2-5…   
Total: 10  
  
There are 5 states being tested (the first can be ignored): 40 tests  
However, evaluating at each tick leads to 400 tests -> unpractical.

Proposed way to do it? Perhaps this article shouldn’t seek to establish statistical results, but instead be a probe for future experiments.