Networks and Markets, HW4

Disheng Zheng

Group, In collaboration with: David Hachuel, Akshay Jha

1. On State level, because most states are not swing states, the outcome of their preferred candidate can be fairly easy to predict. Therefore, it may induce some voters to not vote for their most preferred candidate to waste a vote. This is not DST and not system strategy proof. However, if there only exist 2 voters, it doesn’t give any incentive for voters to not vote truthfully. **There is always a DST and system strategy proof because voters vote truthfully.** Because when voters vote their preferred candidate, it maximizes their utility. And they minimize their utility if they choose the not preferred candidate.

The US Electoral college system **does not always elect a Condorcet winner**. We say that a candidate x ∈ X is a Condorcet winner if for every other x0 ∈ X at least n/2 voters prefer x to x0 (according to the preferences ~ µ). For the case of George W. Bush and Trump, both had less total votes than their opponents, meaning they won less than n/2 of the voters, but they won the electoral college by winning more swing states. They win electoral votes but not popular votes.

1. (a) the male-optimal

|  |  |  |
| --- | --- | --- |
| Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.10.14%20AM.png | Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.10.58%20AM.png | Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.11.27%20AM.png |
| Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.12.58%20AM.png | Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.13.58%20AM.png | Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.15.12%20AM.png |
| Network%20/hw4/MALE/Screen%20Shot%202017-11-27%20at%2012.16.45%20AM.png |  |  |

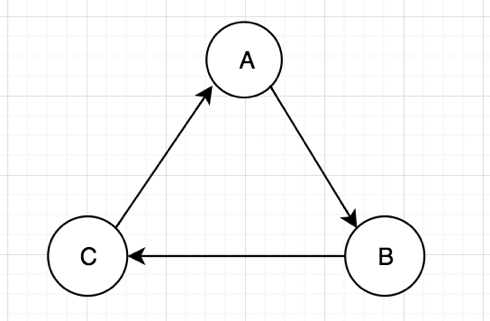
(b) the female-optimal

|  |  |  |
| --- | --- | --- |
| Network%20/hw4/FEMALE/Screen%20Shot%202017-11-27%20at%2012.03.50%20AM.png | Network%20/hw4/FEMALE/Screen%20Shot%202017-11-27%20at%2012.05.42%20AM.png | Network%20/hw4/FEMALE/Screen%20Shot%202017-11-27%20at%2012.06.15%20AM.png |
| Network%20/hw4/FEMALE/Screen%20Shot%202017-11-27%20at%2012.07.15%20AM.png | Network%20/hw4/FEMALE/Screen%20Shot%202017-11-27%20at%2012.08.36%20AM.png |  |

1. For non-bipartite matching setting,

Let’s consider the following case, an odd-length cycle:

|  |  |
| --- | --- |
| Individual | Preferences |
| A | B > C |
| B | C > A |
| C | A > B |

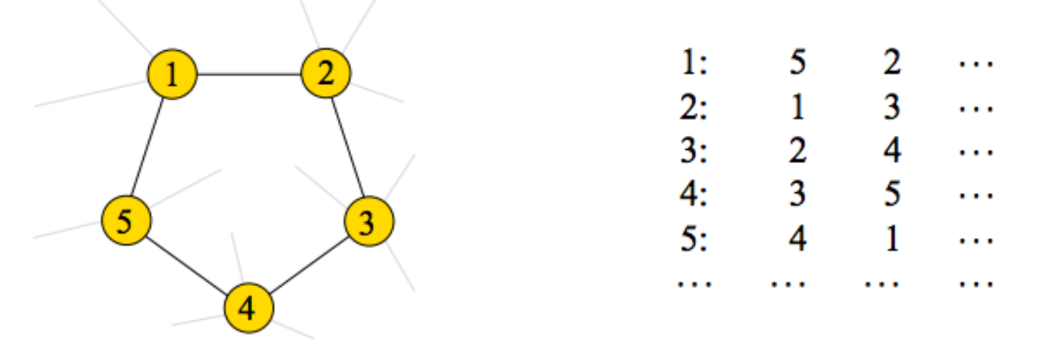


No stable matching exists.

A generalized Proof:

A non-bipartite graph must contain at least one cycle of odd length. Let v1, v2, ..., vk be such a cycle. We construct the preference table T such that vi ranks its predecessor vi−1 first and its successor vi+1 second. According to their mutual highest rankings, the members of this cycle prefer to stay among themselves. Therefore, we call such a cycle exclusive. In an exclusive cycle with an odd number of members at least one member must find a partner outside the cycle. Let us assume that, under a matching M, this poor chap is vi. Then vi prefers vi+1 to its current situation (whether vi is covered by M). Since vi is vi+1’s first choice, vi+1 prefers vi to its current situation and (vi, vi+1) form a blocking pair. Hence matching M is not stable. [1]

Here is another example of an odd exclusive cycle that prevents any matching from being stable:



1. P(+|N) = 15%

P(-|D) = 25%

P(+|D) = 75%

P(-|N) = 85%

P(D) = 1%

P(N) = 99%

Use Bayes’ Rule:

P(D|+) = =

= = 4.808%

The probability that someone who tests positive and actually has the disease is 4.808%

Let’s assume without generality, that the true state is w = 0:

P(w = 1) = P(w = 0) = ½

P(w = 0 | x = 1) = P(w = 1 | x = 0) = ½ - ε

P(w = 0 | x = 0) = P(w = 1 | x = 1) = ½ + ε

For the first person:

If he/she gets a wrong signal, x1 = 1. He will follow the wrong lead and guess w1 = 1.

For the second person:

P1 = P(w = 0 | x1 = 1, x2 = 0) =

P2 = P(w = 1 | x1 = 1, x2 = 0) =

Note that the denominators of the two are the same.

So we are only comparing and

= (½ – ε)( ½ + ε)

= (½ + ε)( ½ - ε)

For the Third person:

P1 = P(w = 0 | x1 = 1, x2 = 1, x3 = 0) =

P2 = P(w = 1 | x1 = 1, x2 = 1, x3 = 0) =

= (½ – ε)2( ½ + ε)

\* = (½ + ε)2( ½ - ε)

Basically, we want cascading incorrectly, meaning given that the player get’s a correct signal, he/she would ignore and still choose the wrong one based on previous players’ choice. so we need P2 > P1:

For each following case, we want

(½ - ε)k( ½ + ε) < (½ + ε)k( ½ - ε)

**Now consider the case where each player value their own signal c times more.**

(½ - ε)k( ½ + ε)c < (½ + ε)k(½ - ε)c

let :

W = ½ - ε

X = (½ + ε)c

Y = ½ + ε

Z = (½ - ε)c

Since W, X, Y, Z are in range (0, 1)

log(WkX) > log(YkZ)

= k\*logW + logX > k\*logY + logZ

= k\*log(W/Y) > k\*log(Z/X)

k > = = c

k > c makes sense because number of wrong signals must be bigger than “my own signal” for “me” to follow the lead.

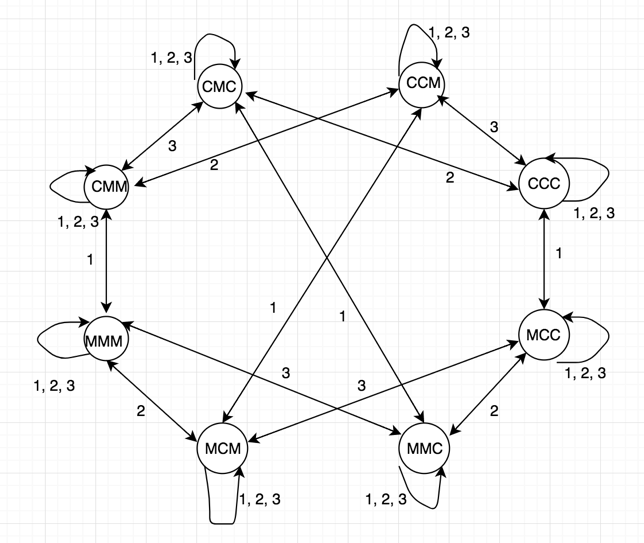
The above condition must be true for cascade to happen.

**The probability for the cascade to happen is:**

P(cascade) = (½ - ε)c+1

Because the cascade happens when at least k = c+1.

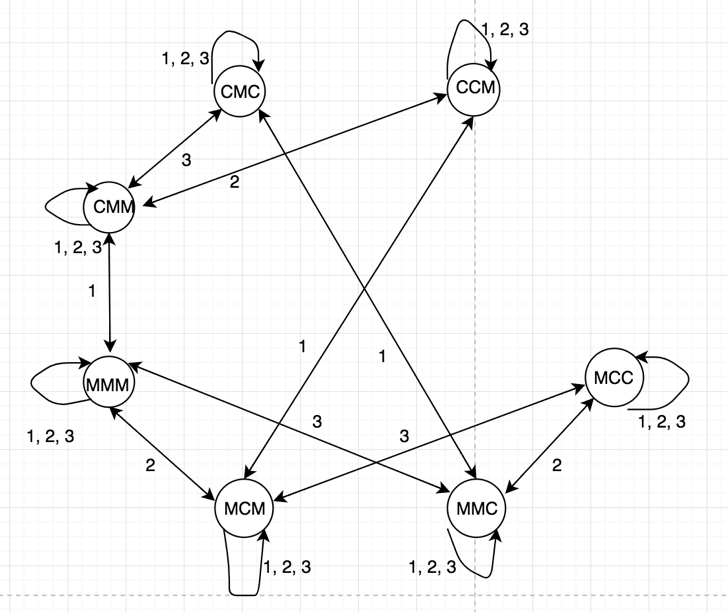
(a) The knowledge network for the muddy children example, 3 total children case.



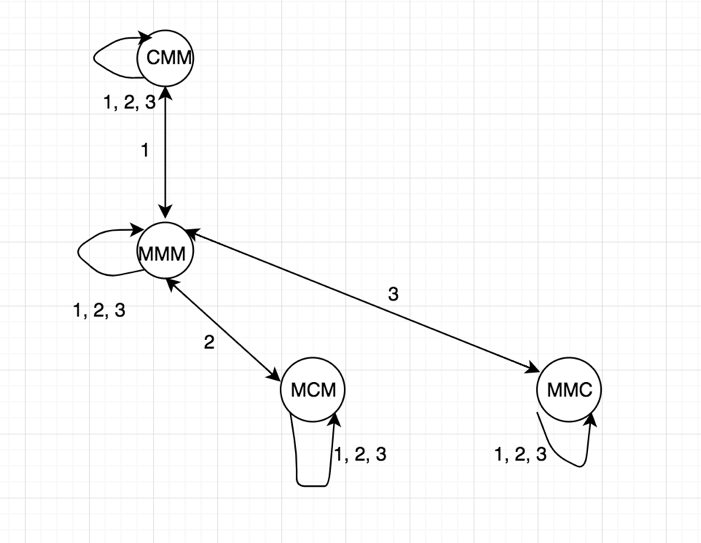
(b)

We begin by showing P(1). Because the father mentions that there are some muddy children, if there is only one muddy child, they will see nobody else in the room with mud on their forehead and know in the first round that they are muddy. Conversely, if there are two or more muddy children, they are unable to discern immediately whether they have mud on their own forehead; all they know for now is that some children (which may or may not include themselves) are muddy.

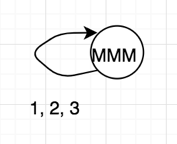
The first time father announces “there is muddy children” kills the CCC world. Because someone must has mud on their face. And for 1 muddy child case, they would see all other kids have clear face. They would know that they have muddy face and went to clean up. This leaves us possible worlds:



Then we show P(2). Suppose there are exactly 2 muddy children. Since there are more than 1 muddy children, nobody will say “yes” before round 2. In 2nd round, each muddy child sees 1 other muddy children, and knows thus that there are either 1 or 2 muddy children total. However, they are able to infer that, were there only 1 muddy children, someone would have said “yes” in the previous round; since nobody has spoken yet, each muddy child is able to deduce that there are in fact 2 muddy children, including themselves. Thus the only the world with at least 2 muddy children left:



Then we show P(3). Suppose there are exactly 3 muddy children. Since there are more than 2 muddy children, nobody will say “yes” before round 3. In 3nd round, each muddy child sees 2 other muddy children, and knows thus that there are either 2 or 3 muddy children total. However, they are able to infer that, were there only 2 muddy children, someone would have said “yes” in the 2nd round; since nobody has spoken yet, each muddy child is able to deduce that there are in fact 3 muddy children, including themselves. Thus the only the world with at least 3 muddy children left:



(from the notes):

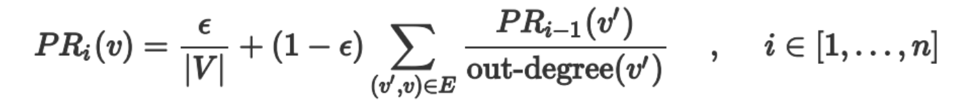
Now assume that P(k) is true for 0 ≤ k ≤ n; we will show P(n + 1). Suppose there are exactly n + 1 muddy children. Since there are more than n muddy children, by the induction hypothesis nobody will say “yes” before round n+1. In that round, each muddy child sees n other muddy children, and knows thus that there are either n or n+1 muddy children total. However, by the induction hypothesis, they are able to infer that, were there only n muddy children, someone would have said “yes” in the previous round; since nobody has spoken yet, each muddy child is able to deduce that there are in fact n + 1 muddy children, including themselves.

If there are strictly more than n + 1 muddy children, however, then all children can tell that there are at least n + 1 muddy children just by looking at the others; hence, by the induction hypothesis, they can infer from the start that nobody will say “yes” in round n. So they will have no more information than they did initially in round n + 1, and will be unable to tell whether they are muddy as a result.

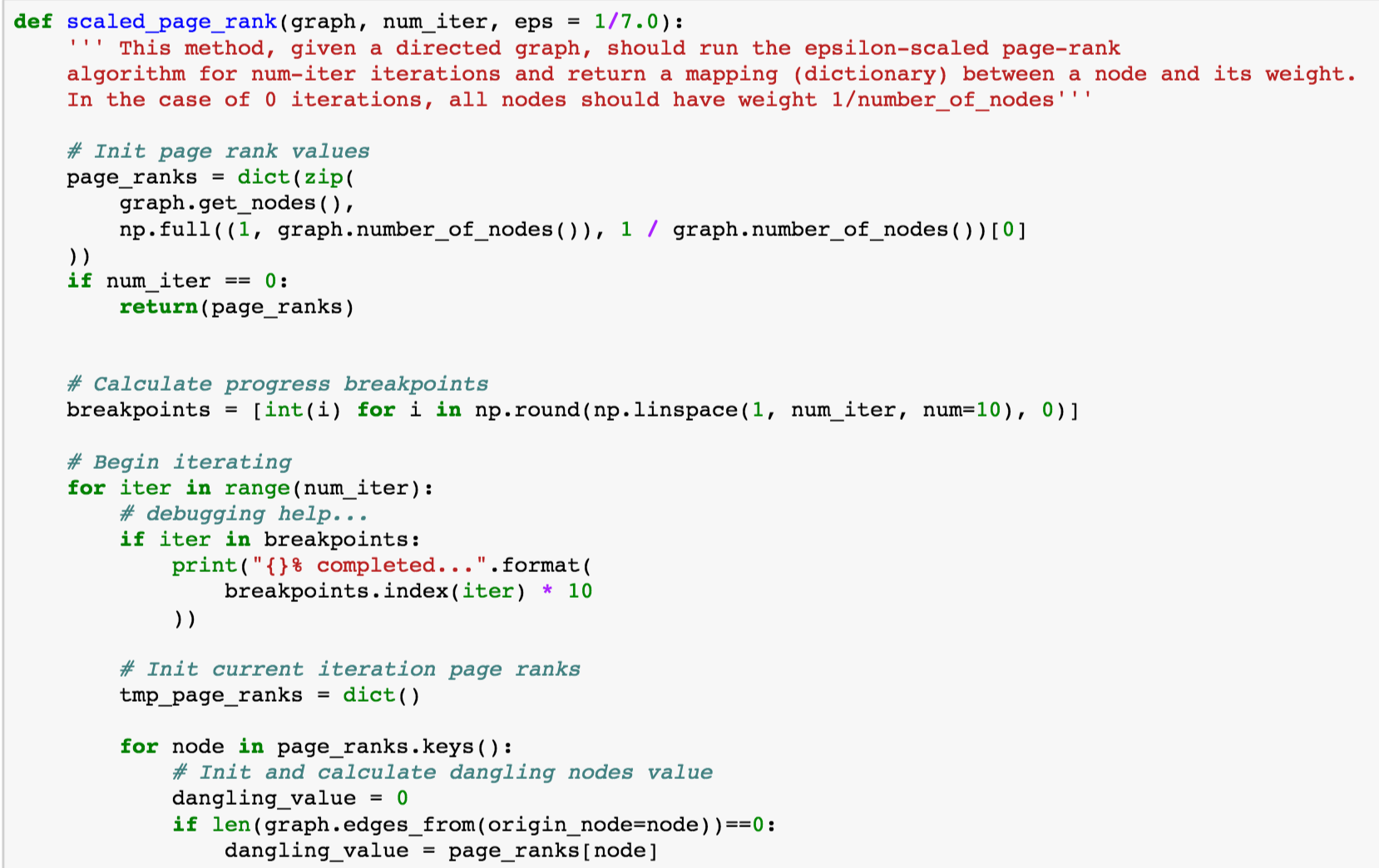
Given a graph G = (V, E), for each iteration i:

The ε-scaled PageRank algorithm has the following formula:

PR(v) =



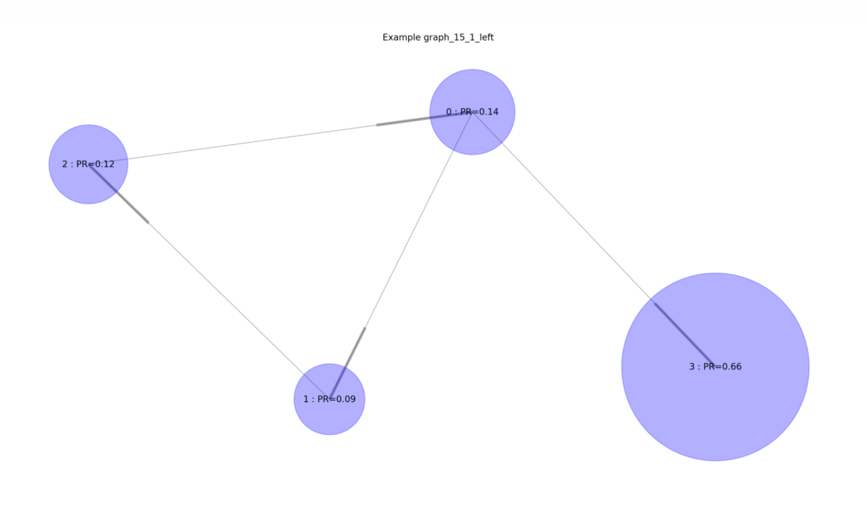
Here is the code for it:



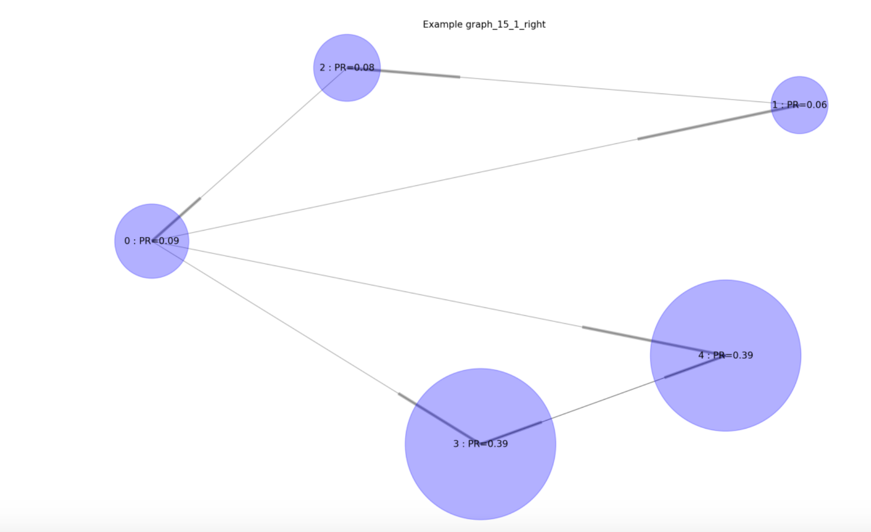




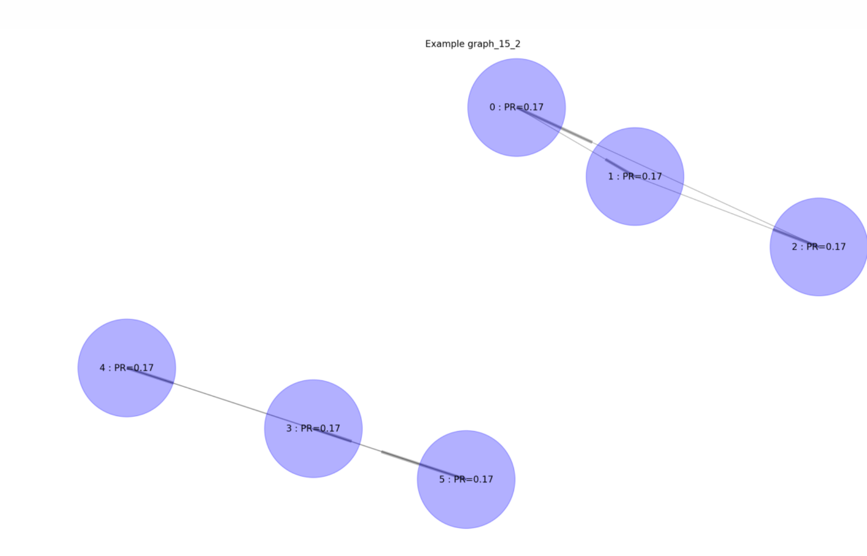
Then I test the scaled page rank on graph 15.1 Left:



Then I test the scaled page rank on graph 15.1 right:

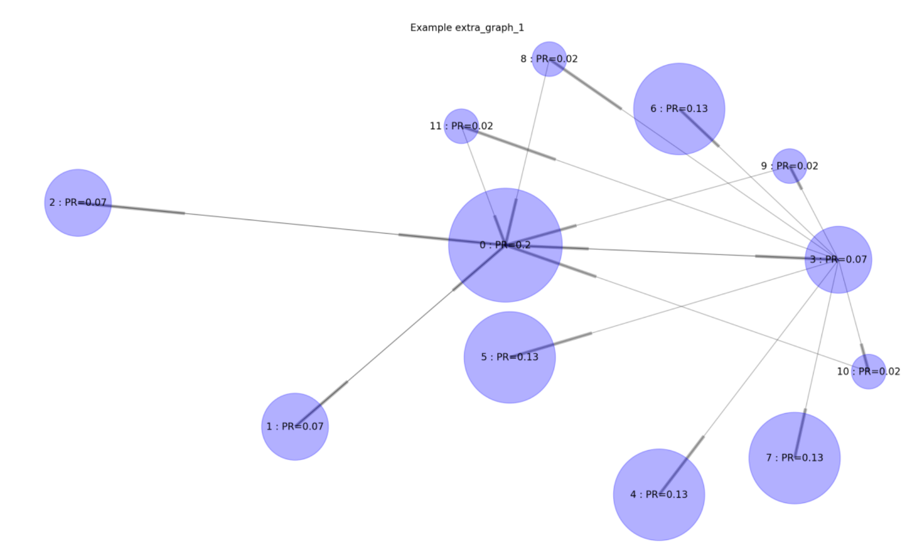


Then I test the scaled page rank on graph 15.2: The 2 sets n = 10 result in same result

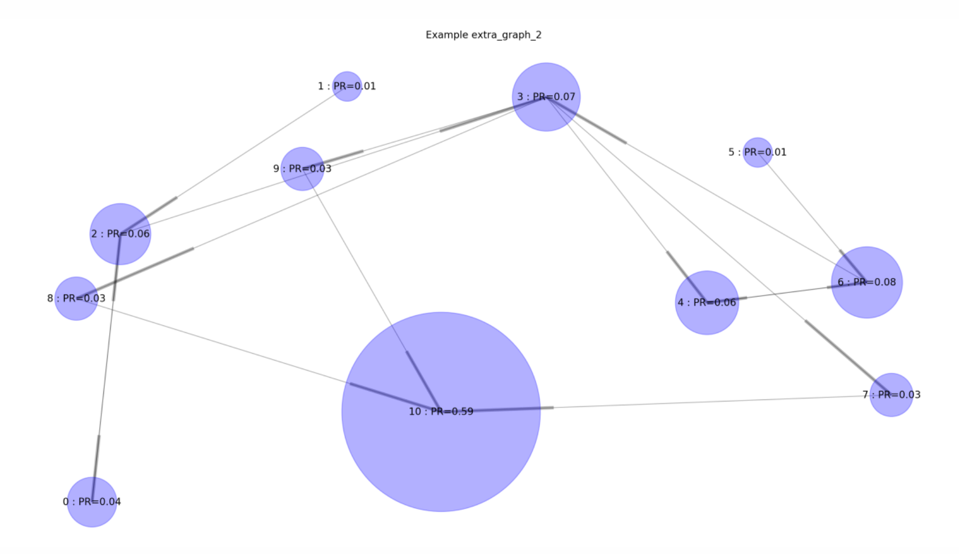


Then I tested it on 2 other self-generated examples:

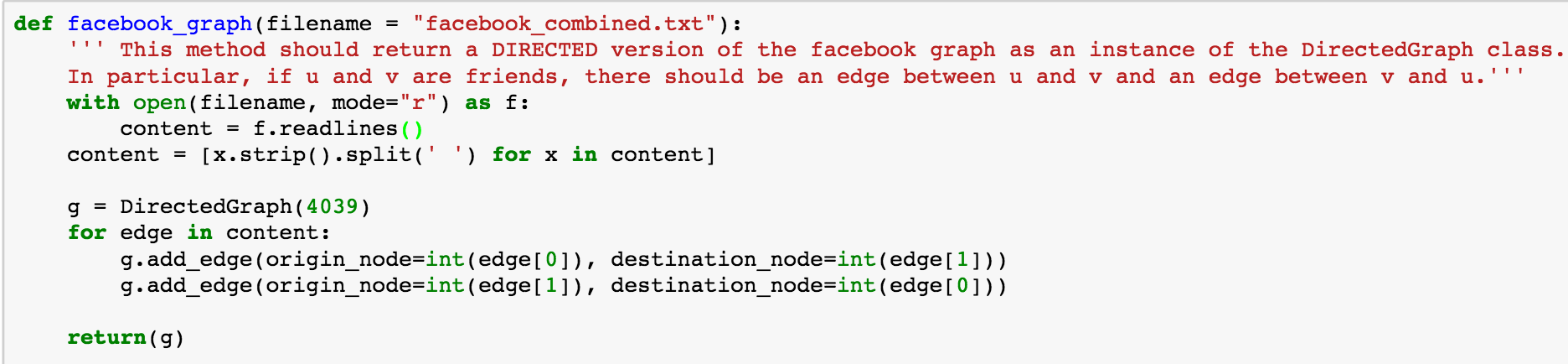
extra\_graph\_1:



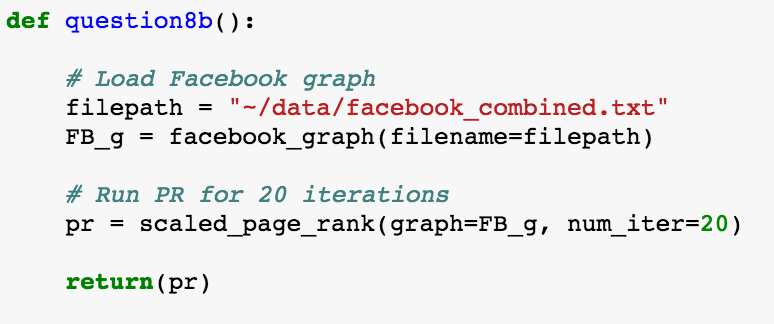
extra\_graph\_2:



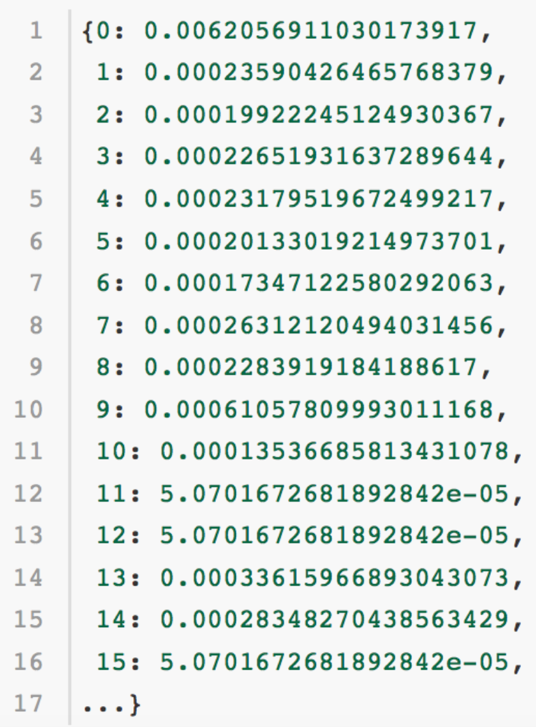
1. In this part, I need to turn facebook data into a directed graph:



1. To run page rank on Facebook data. It converges after n = 20.

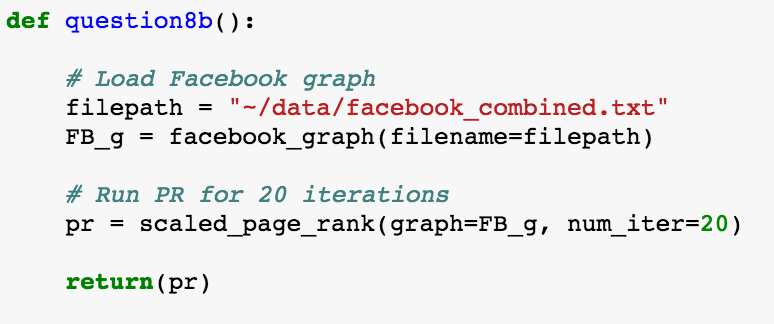


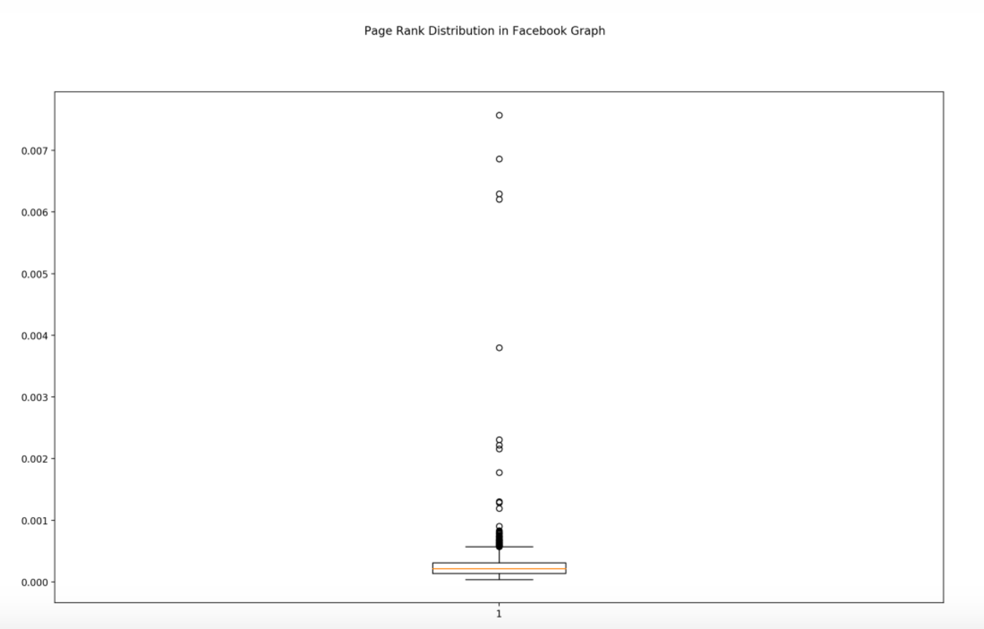
The output:



(c)

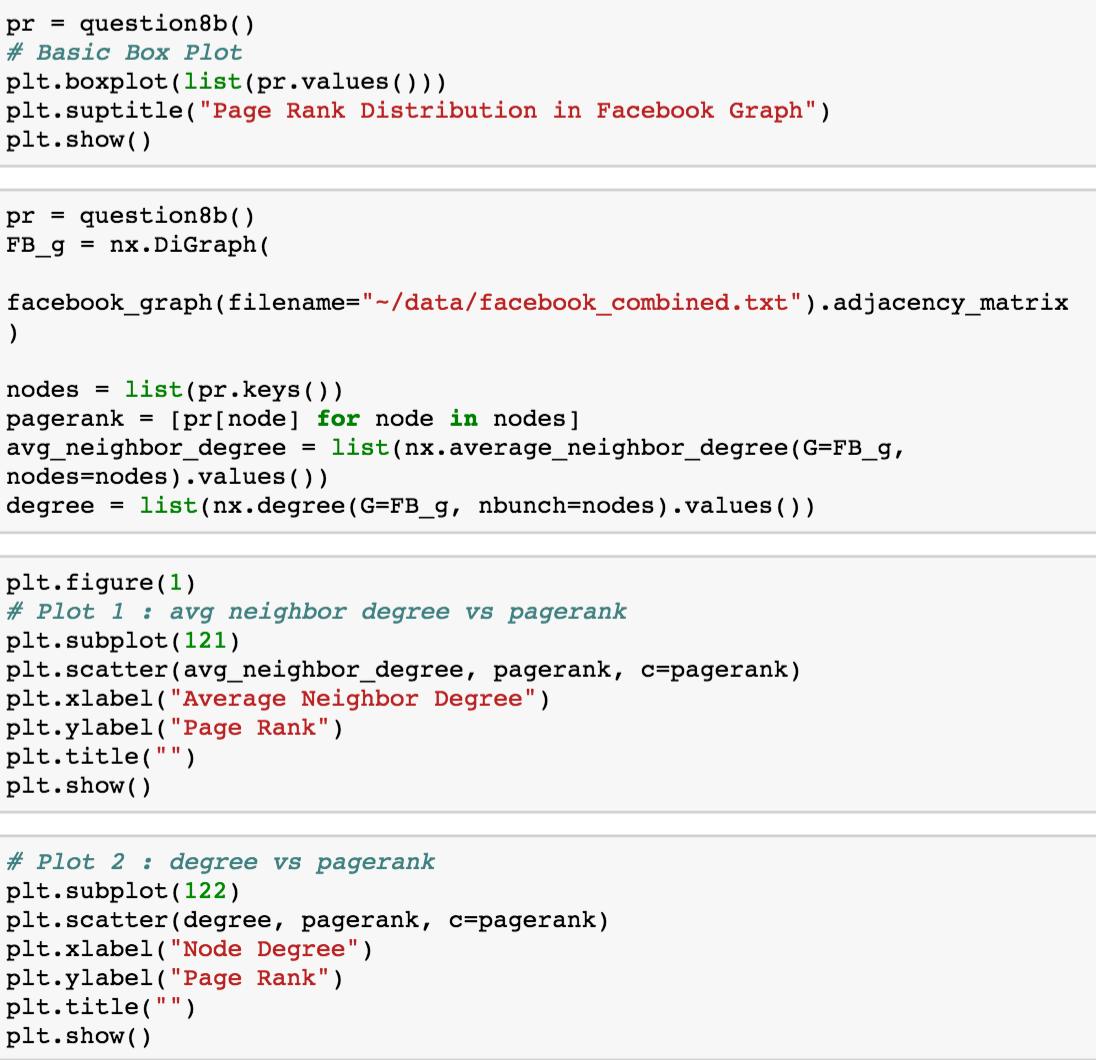
After running ε-scaled PageRank for a few times on facebook, we realized the following distribution.



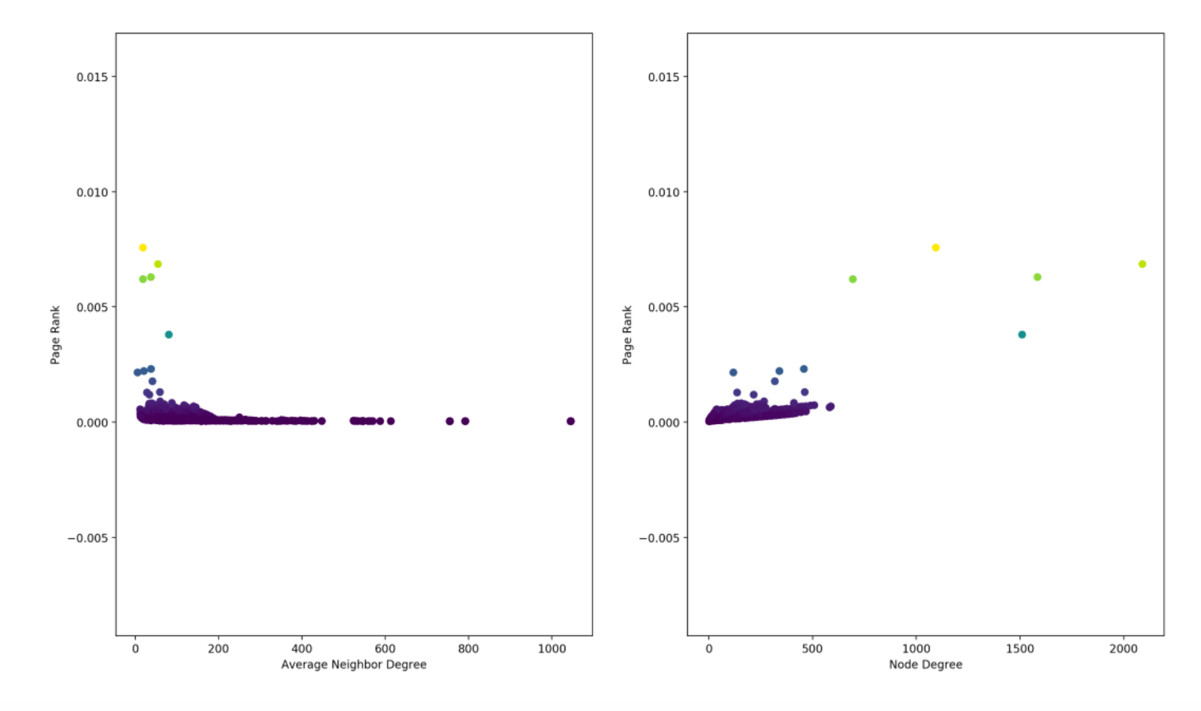


There are a few nodes that are outliers that are far skewed from the range.

The highest scores are concentrated in a small group of people. The distribution of PageRank is a right long-tailed and changes with the damping factor .



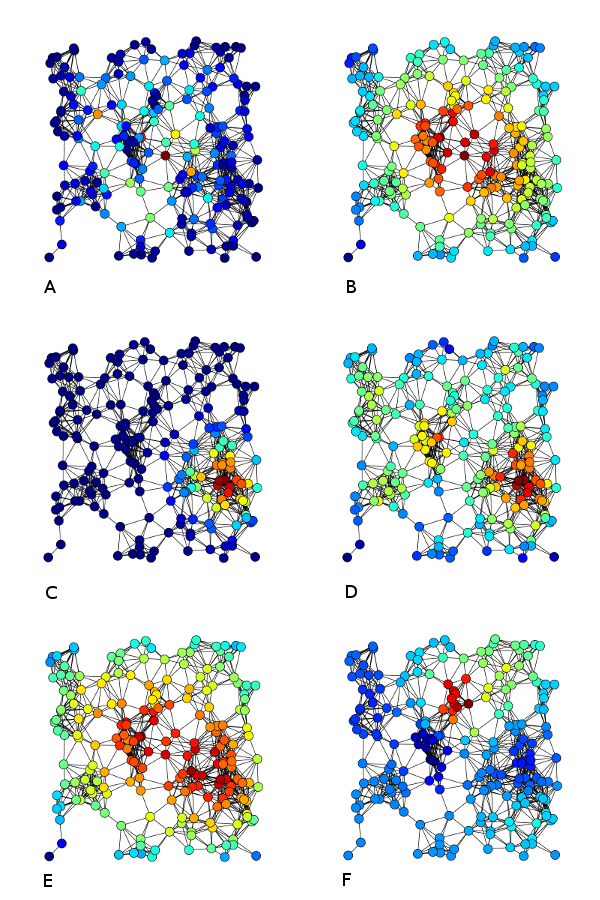
I then graph PageRank against highest and lowest page rank node’s degree and average neighbor node degree.



The node degree is positively correlated to page rank, meaning highly connected nodes have higher page rank. However, it seems that nodes with high page rank score have poorly connected nodes as neighbors. Also, as scores flow though the graph in a not well connected neighborhood, the score would most likey to end in a most well connected node. This might be that those highly connected nodes have more “fake accounts” as “fake friends”.

(d)

The influence in a social network can be determined by centrality. Page rank is one way to determine centrality which rank nodes who are connected to other influential nodes higher. Page rank is a special case of eigenvector centrality. A node could be well connected but with a low page rank score, because it is connected to neighbors of with low page rank score too.

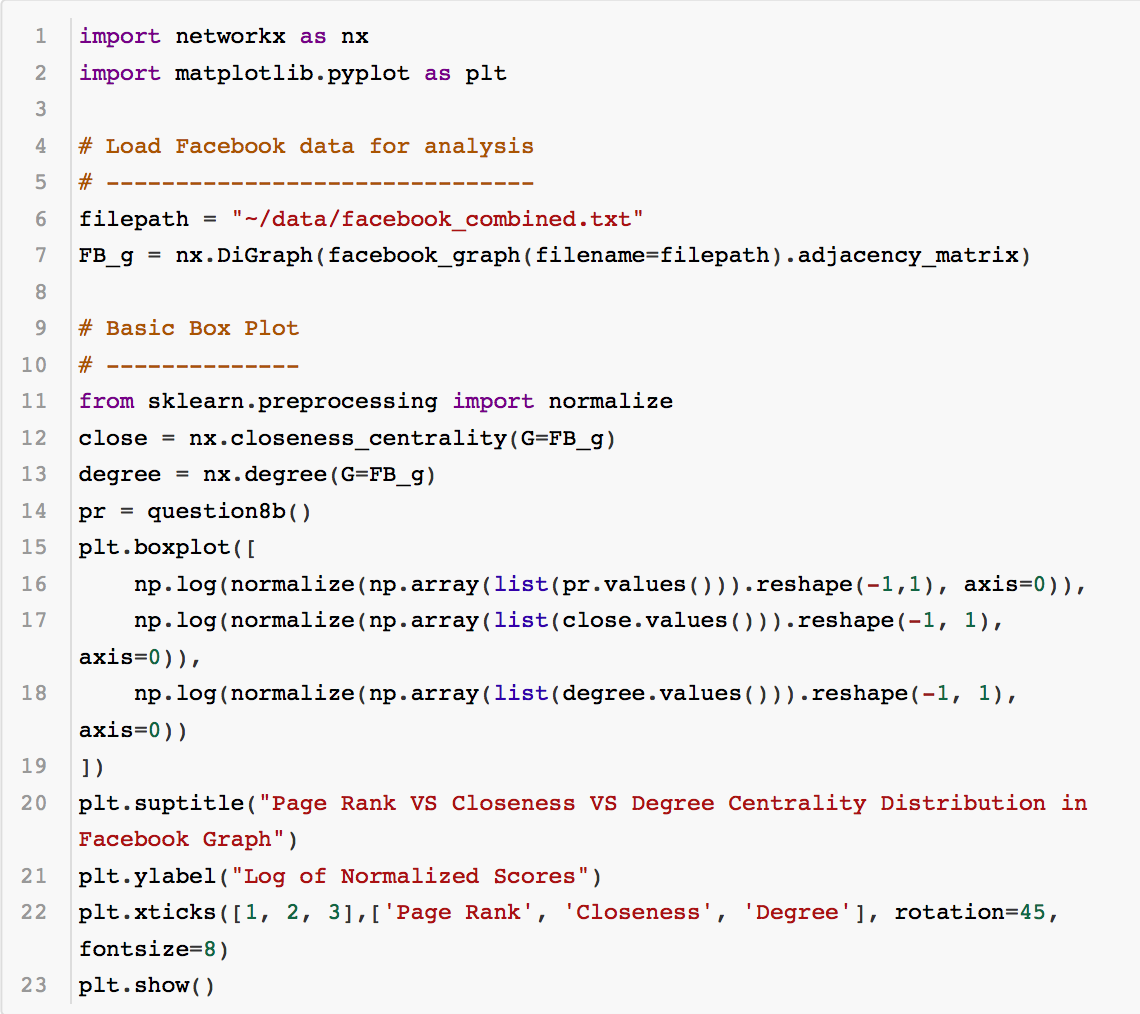


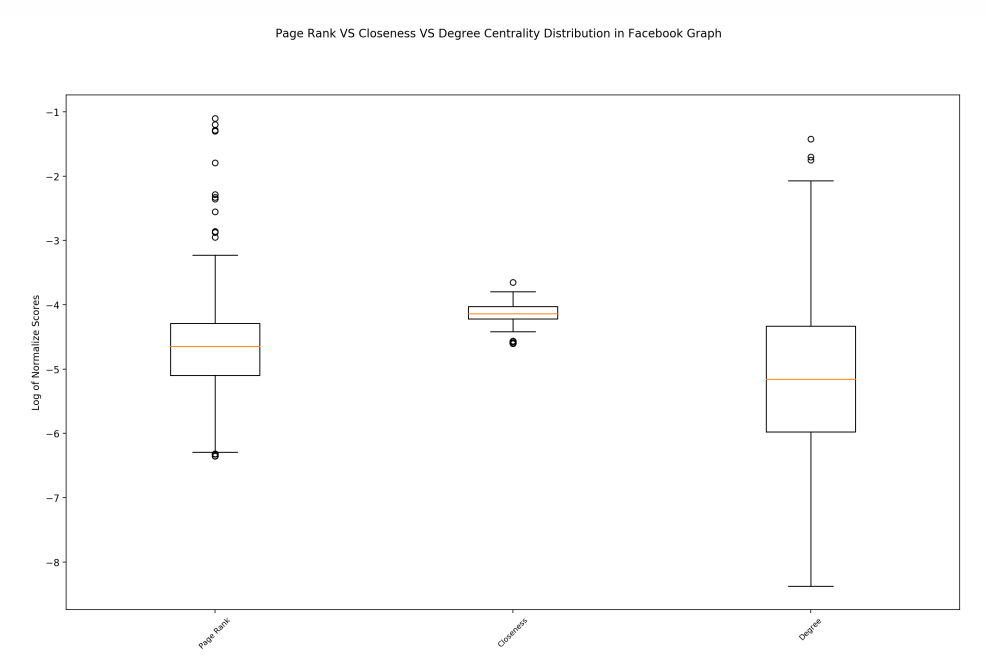
[2]

1. Betweenness Centrality: The number of shortest path go through a node
2. Closeness Centrality: How close (degree of distance) that a node is to the rest of the network.
3. Eigenvector Centrality
4. Degree Centrality
5. Harmonic Centrality
6. Katz Centrality

For applications in reality, when a company is trying to promote a new product, they may seek influential person (node) as initial adopters. If they found the right persons, cascading effect could be very effective. In this case, the company might prefer closeness centrality over degree centrality because nodes with high closeness centrality has higher influence and spreading strength.

Here are the comparisons of different centrality based on the Facebook network.



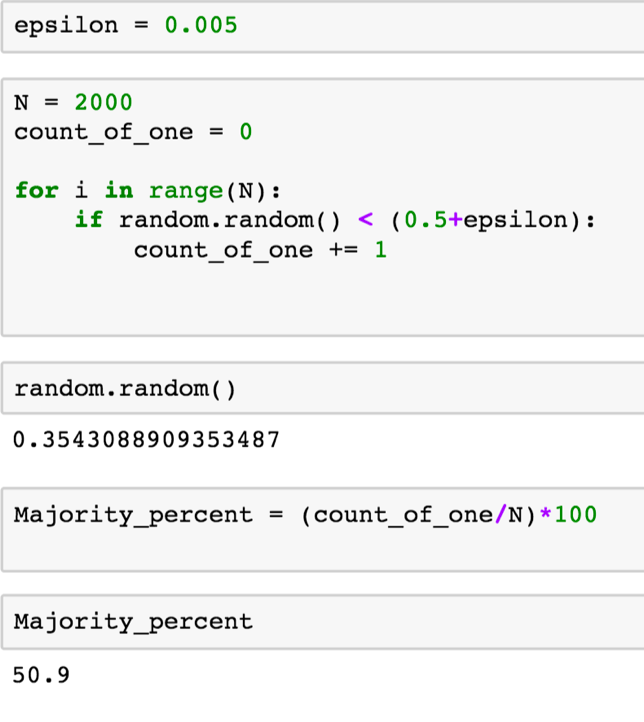


The above distribution is log normalized to create a contrast in distribution. It can be seen that closeness centrality has much smaller spread than page rank score and degree centrality. As a company trying to promote a product, they may want to avoiding using closeness centrality because it’s hard to focus resource on a target when there is lack of differentiability. Degree centrality on the other hand could be used to estimate the cost of allocating resources.

9.

(a)

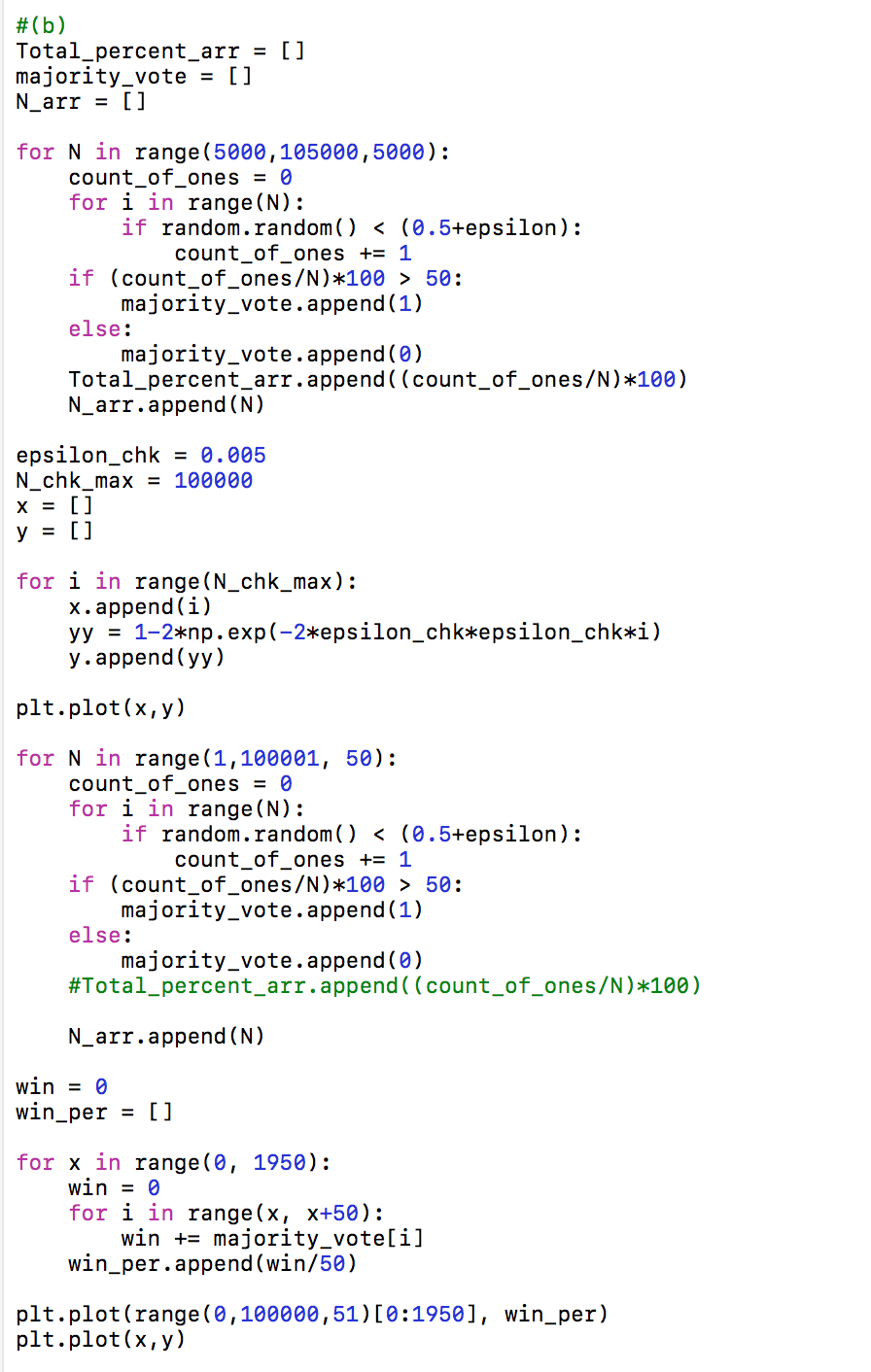
Here I wrote a simple program that generates N random pieces of “evidence” independently, each of which is 1 with some probability (1/2 + ε).

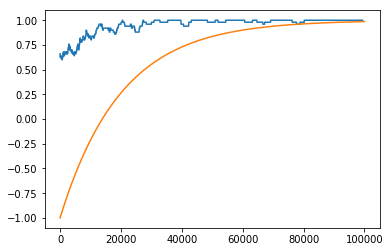


The result, percentage of people voting 1 is 50.9%. The majority is 1. (correct) This is about the same as the probability 1/2 + ε. It is the majority but approximately 50%. It’s highly likely that the majority is 0 (wrong) with a slightly lower probability.

(b)

By running 100 trials with N ranging from (5000 to 10000, increment 5000), and ε = 0.005.





Here the yellow line is Chernoff Bound. Blue line is the experimental result of the percentage times that the right (correct 1) majority wins.

This indeed agree with Theorem 16.2 because the right (correct 1) majority wins with the probability at least as great as the Chernoff Bound.

It can be seen that on small number of people scales, poll level, the actual probability that right (correct 1) would win has greater variability within itself and that it is in greater distance away from the Chernoff Bound. It is much harder to predict and capture the correct result. On the other hand, with large number of people, election level, the result varies less and it is close to the Chernoff bound.

The election result of Trump and Hilary in 2016 can be explained as: on poll level, the result consist of much greater volatility. Especially since the number of votes are quite close to each other for the two candidates, the result remains much uncertainty on poll level.

Another thing is that: voters cannot be modeled as independent variables. They affect each other as a group. Also the electoral college plays a role for dependencies.

References:

[1] <http://www.dcs.gla.ac.uk/~pat/jchoco/roommatesMorph/papers/0509221.pdf>

[2] <https://en.wikipedia.org/wiki/User:Rocchini>