Kinaan Patel and Terry Lyu
Repository name: homework-1-kinaanpatel_terrylu

CSE 316A – HW 1

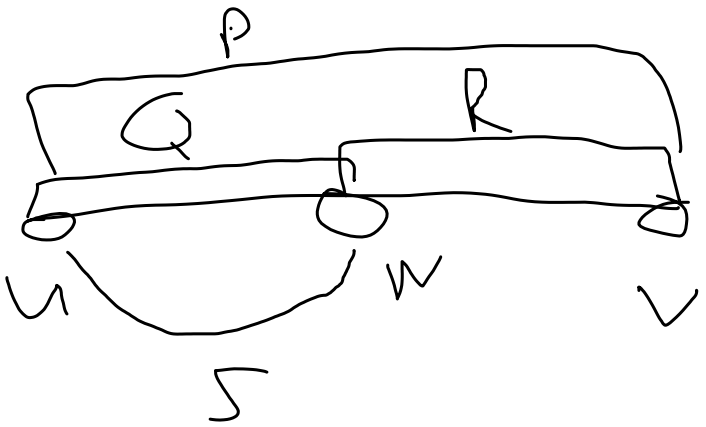
1.

a. $\frac{(originating\ edges)(destination\ edges)}{(2\ for\ double\ counting)} = \frac{(n)(n-1)}{2}$

b. Each node is connected to all $n-1$ other nodes, so the degree is $d_i = n-1$

c. Each node represents a person in the social network. Each edge represents interactions between people. The weight of the graph (non-negative integers) can represent the number of interactions between any two people.

d.



Shortest path P is the path Q between u and w plus R the path between w and v.

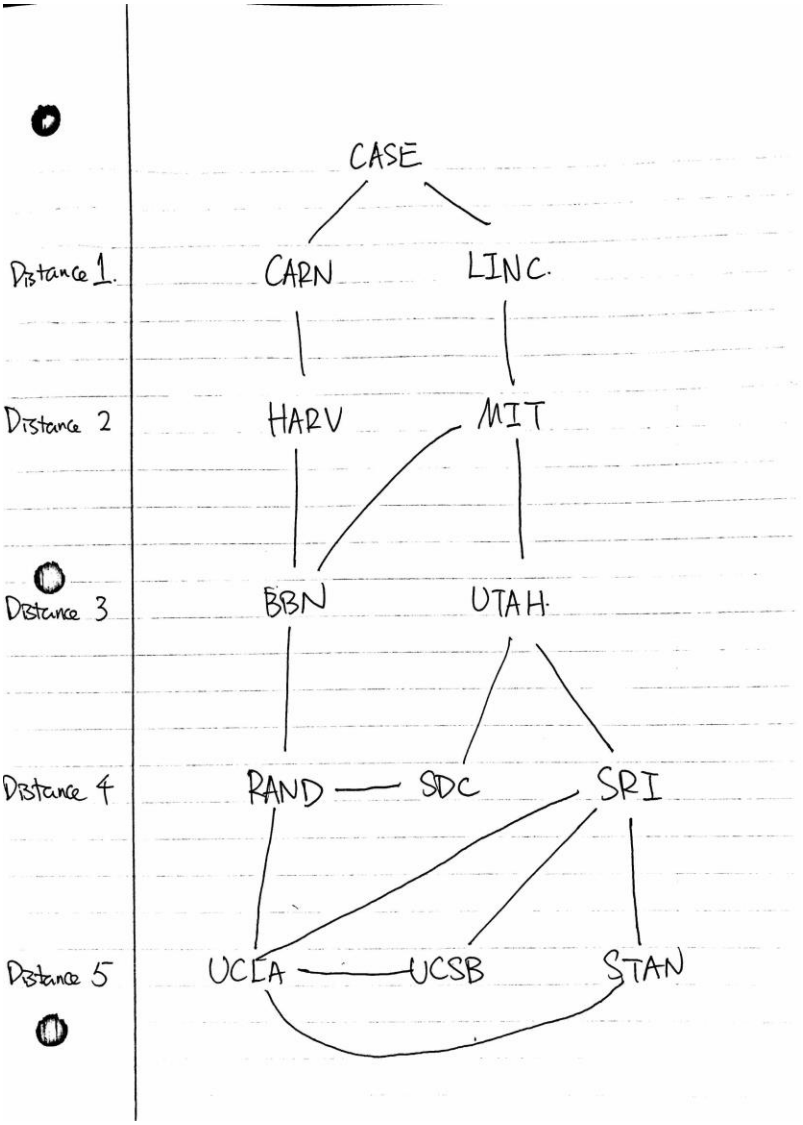
If S, another path between u and w, is shorter than Q then $S + R < Q + R \therefore P$ would not be the shortest path.

Because P is the shortest path there cannot exist a path between u and w shorter than Q

Kinaan Patel and Terry Lyu
Repository name: homework-1-kinaanpatel_terrylu

2.

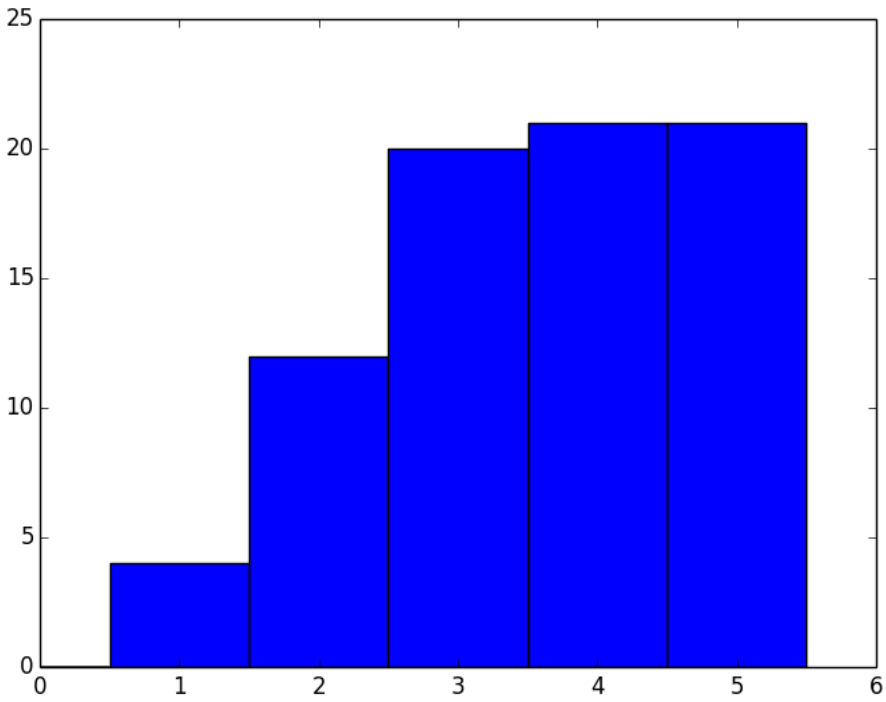
a.



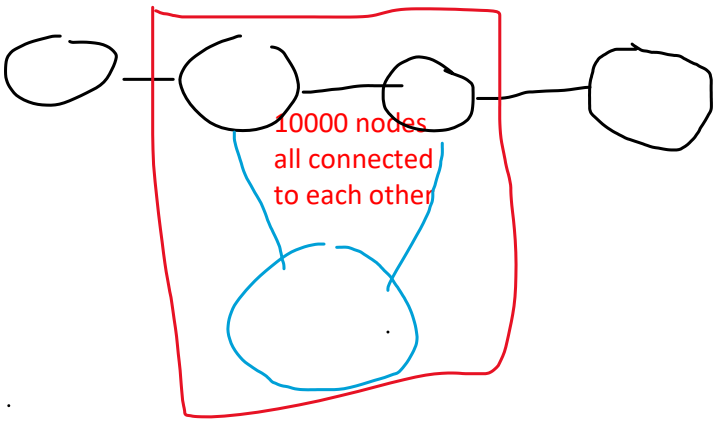
b. Repository name: **homework-1-kinaanpatel_terrylu**
Output from problem2b.py code:
[(('CASE', 'CASE'), 0), (('CASE', 'RAND'), 4), (('CASE', 'BBN'), 3), (('CASE', 'SRI'), 4), (('CASE', 'UTAH'), 3), (('CASE', 'UCLA'), 5), (('CASE', 'UCSB'), 5), (('CASE', 'CARN'), 1), (('CASE', 'SDC'), 4), (('CASE', 'STAN'), 5), (('CASE', 'HARV'), 2), (('CASE', 'LINC'), 1), (('CASE', 'MIT'), 2)]

Kinaan Patel and Terry Lyu
Repository name: homework-1-kinaanpatel_terrylu

c. Distance distribution for Arpnet network



d. One graph that will have an average distance of more than three times its diameter would be a graph that has 10000 nodes that are all connected to each other. One of these nodes will then be connected to a string of 2 other nodes. The diameter will then be 4, and the average distance will be less than 1.33. See below for an illustration.



e. This could be extended to produce a graph with a diameter more than c times as large as the average distance by adding more nodes on either side of the 2 nodes that are not connected to the other 10000. If the average distance becomes too large, the 10000 nodes that are all connected to each other can be increased to more nodes.

Kinaan Patel and Terry Lyu
Repository name: homework-1-kinaanpatel_terrylu

3.

a.

List of actors:

1. Davis, Mark (V)
2. Sanders, Alex (I)
3. North, Peter (I)
4. Marcus, Mr.
5. Tedeschi, Tony
6. Dough, Jon
7. Stone, Lee (II)
8. Voyeur, Vince
9. Lawrence, Joel (II)
10. Steele, Lexington
11. Ashley, Jay
12. Boy, T.T.
13. Cannon, Chris (III)
14. Jeremy, Ron
15. Bune, Tyce
16. Hanks, Tom
17. Michaels, Sean
18. Stone, Kyle
19. Hardman, Dave
20. Surewood, Brian

Most of these actors, except for Tom Hanks, are adult film actors. This allows them to make lots of movies because each movie takes less time to make. Because they make so many movies then can act with lots of other stars giving them a high degree centrality.

b.

1) Jeremy, Ron
2) Chan, Jackie (I)
3) Cruz, Pen�lope
4) Shahlavi, Darren
5) Del Rosario, Monsour
6) Depardieu, G�rard
7) Bachchan, Amitabh
8) Jackson, Samuel L.
9) Soualem, Zinedine
10) Del Rio, Olivia
11) Jaenicke, Hannes
12) Hayek, Salma
13) Pel�
14) Knaup, Herbert
15) Goldberg, Whoopi
16) Roth, Cecilia
17) Bellucci, Monica
18) Hanks, Tom
19) August, Pernilla
20) Kier, Udo

These actors all have acted in multiple genres, including the genres that are slightly removed from the largest part of the component such as the action or romance movies, so while they may not have a very large number of connections, their connections bridge groups that may otherwise be more removed from the graph.

c.

1. Jackson, Samuel L.
2. Goldberg, Whoopi
3. Berry, Halle
4. Diaz, Cameron
5. Hanks, Tom
6. Stiller, Ben
7. Myers, Mike (I)
8. Douglas, Michael (I)
9. Lopez, Jennifer (I)
10. De Niro, Robert
11. Willis, Bruce (I)
12. Cruise, Tom
13. Hopper, Dennis
14. Kidman, Nicole
15. Smith, Will (I)
16. Washington, Denzel
17. Travolta, John
18. Madonna (I)
19. Schwarzenegger, Arnold
20. Hoffman, Dustin

The actors with the top closeness are those that have the smallest average shortest path to all other nodes. Actors in the center of our network graph will have very short connections to all other actors in the center component, and will not be far to a connection with any of the other clusters because of their centrality. The actors with the highest degree centrality, on the other hand, will be in the red adult actor group because they are very well connected to each other, but these actors will be far from other areas. Actors who act as bridges between groups will have a high betweenness centrality, but because they are often primarily in or very near other genre pockets, they will not be near the center of the network in close proximity to all other nodes.