1. Diameter & Average Path Length:

① if distance unconsidered, D=6, implying that we have to switch at most 6 times to go there (node 9—Tuluksak, Alaska) from here (node 332--West Tinian, a western island near Philippines); APL=2.74.

② if distance considered, D=8661 miles (node 37, Eareckson Air Station at Alaska to node 330, Babelthuap/Koror at Palau in western Pacific; APL=2033 miles. Note that the east-west straight distance of the US is about 2500 miles.

Conclusion: Network is rather compact, but some nodes that belong to the **overseas territories** or **remoted states** (e.g. Alaska) may influence the density, because the degree of these nodes tend to be relatively small. Node 9 has degree 4, node 37 and 330 have degree 2 and node 332 has merely 1 degree.

2. Degree—Consider some big-degree nodes as well as small ones!!!!!!!

Biggest 20 are: [('118', 139), ('261', 118), ('255', 101), ('182', 94), ('152', 94), ('230', 87), ('166', 85), ('67', 78), ('112', 70), ('201', 68), ('147', 67), ('293', 62), ('162', 62), ('176', 61), ('258', 60), ('248', 59), ('144', 59), ('47', 57), ('299', 56), ('217', 56)]

(**don’t forget to draw a bar pic**)

118- Chicago O’hare Intl-伊利诺伊州-Illinois

261- Dallas/Fort Worth Intl-得克萨斯州-Texas

255- Hartsfield-Jackson Atlanta Intl-佐治亚州-Georgia

182- Lambert-St Louis Intl-密苏里州-Missouri

152- Pittsburgh Intl-宾夕法尼亚州-Pennsylvania

230-Charlotte/Douglas Intl-北卡罗来纳州-North Carolina

166- Stapleton Intl-科罗拉多州-Colorado

(Now **Abandoned** and moved to Denver because of natural disaster) Denver also in Colorado！

67- Minneapolis-St Paul Intl-明尼苏达州-Minnesota

112- Detroit Metropolitan Wayne County Airport-密歇根州-Michigan

201- San Francisco Intl-加利福尼亚州-California

147- Newark Intl-新泽西州-New Jersey

293- Houston Intercontinental-德克萨斯州-Texas

162- Philadelphia Intl-宾夕法尼亚州-Pennsylvania

176- Cincinnati/Northern Kentucky Intl-肯塔基州-Kentucky

258- Phoenix Sky Harbor Intl-亚利桑那州-Arizona

248- Los Angeles Intl-加利福尼亚州-California

144- Salt Lake City Intl-犹他州-Utah

47- Seattle-Tacoma Intl-华盛顿州-Washington

299- Orlando Intl-佛罗里达州-Florida

217- Nashville Intl-田纳西州-Tennessee

It includes Illinois, Texas, Georgia, Missouri, Pennsylvania, North Carolina, Colorado, Minnesota, Michigan, California, New Jersey, Kentucky, Arizona, Utah, Washington, Florida and Tennessee.

We can see that the number of the big and busy airports is rather uniformly distributed to SOME states. Texas, Pennsylvania and California have **two** such airports, indicating that they are better in development, which accords with the situation in the real world.

Then we calculate the number of big airports in the WEST/MID/EAST part of the US. It shows that:

WEST-6

MID-6

EAST-8

We can see that it is also uniformly distributed. So this may be a proof that the US had a balanced development from that time (1997).

Now let’s see those airports with degree merely 1.

**Alaska-6**

Small scale or for special use:

Washington-1

North Dakota-2

Minnesota-1

Michigan-1

New York State-1

Illinois-2

Pennsylvania-1

Nebraska-1

Ohio-2

New Jersey-1+

Colorado-3

**California-6**

Missouri-2

West Virginia-2

Virginia-1

Kentucky-1

North Carolina-2

Arizona-2

Kansas-1

**Texas-7**

Louisiana-2

Oregon-1

Florida-2

Oklahoma-1

Overseas Territories:

Puerto Rico-1

Pacific Ocean-2

**Total-3**

We can see that except Alaska, Texas and California, airports with extremely small degree (merely 1) are also uniformly distributed to every state. As for Alaska, it has 6 small airports, possibly due to its harsh weather and its undeveloped tourism at that time (1997). With the prosper of tourism, Alaska may have more airlines. As for Texas and California, they own two super-big airports and they also have many small airports, which locate mostly near their scenic spots. The above facts indicate that Texas and California are well-developed in air transportation.

3. our network VS random network/small-world/scale-free (if node number is the **same**)

|  |  |  |  |
| --- | --- | --- | --- |
|  | Average Degree | Clustering Coefficient | Average Path Length |
| Our Network | 12.81 | 0.625 | 2.738 |
| Random | P=0.3, then 99.60 | 0.300 | 1.263 |
| Small-World WS | K=3, then 6 | P=0.2, K=3, then 0.192 | 5.863 |
| Scale-Free BA | 11.78 | 0.070 | 2.530 |

4. assortativity is negative. Then we investigate whether those big nodes connect with big nodes. (suppose those whose degree >=20 are big nodes)

{'118': 56, '261': 54, '255': 53, '182': 52, '152': 47, '230': 45, '166': 45, '67': 52, '112': 50, '201': 41, '147': 45, '293': 43, '162': 43, '176': 45, '258': 40, '248': 43, '144': 30, '47': 35, '299': 45, '217': 41}

Compared with their total degree:

[('118', 139), ('261', 118), ('255', 101), ('182', 94), ('152', 94), ('230', 87), ('166', 85), ('67', 78), ('112', 70), ('201', 68), ('147', 67), ('293', 62), ('162', 62), ('176', 61), ('258', 60), ('248', 59), ('144', 59), ('47', 57), ('299', 56), ('217', 56)]

We can see that the probability that big airports connect with big airports Pbb is roughly the same as big airports connect with small airports Pbs. This may be the reason for the negative assortativity.

5. Top 20 node betweenness:

[(118, 0.20731993223098968), (8, 0.16948031947190978), (261, 0.15241740407914336), (201, 0.09384593454509225), (47, 0.09241355545314715), (182, 0.08104336166564649), (255, 0.07085473163772363), (152, 0.06916485349203222), (13, 0.06509811101956116), (67, 0.06479638197389133), (313, 0.06215761735790646), (230, 0.05421400126936584), (144, 0.048700235274616095), (166, 0.04537799052973221), (65, 0.04354882086075502), (248, 0.03496949563237522), (112, 0.0320952394739868), (258, 0.025578157331868207), (329, 0.01797125331868534), (293, 0.01685547257189301)]

These nodes’ degrees are:

[139, 29, 118, 68, 57, 94, 101, 94, 14, 78, 24, 87, 59, 85, 41, 59, 70, 60, **4**, 62]

We can see that those with big node-betweenness tend to have relatively big degree, except for node 329.

329 is **Guam Intl** (关岛国际机场), whose neighbors are 313 (Hawaii), 327 (Saipan 塞班岛), 328 (Rota 罗塔岛) and **330 (Palau, mentioned above)**. In conclusion, Guam Intl is crucial in connecting those airports in some (**scenic**) islands!!!

In addition, 313’s degree is 24, 327’s degree is 4, 328’s degree is 2, 330’s degree is 2. That means the Guam Intl is an important **BRIDGE**!

Also, node **13—Bethel Airport**, located in Alaska,has a merely degree of 14. Its neighbors are: ['6', '7', '8', '9', '10', '11', '12', '14', '15', '17', '18', '19', '20', '21']. **ALL** of them belong to **Alaska**, and at the same time, all of them have a low degree (1~4) except node 8 (Anchorage Intl, where Anchorage is the biggest city in Alaska)—That is to say, node 13 is rather **crucial** to Alaska’s airline transportation.

Those whose degree is top20 but not top20 node-betweenness are:

['147', '162', '176', '299', '217’] ---generally they are at the back of the top20 degree list!!

As for the edge-betweenness, [(248, 331), (313, 331), (327, 330), (329, 330), (313, 329), (327, 329), (328, 329), (327, 328), (313, 326), (147, 325), (150, 325), (221, 325), (273, 325), (299, 325), (311, 325), (321, 325), (322, 325), (147, 324), (150, 324), (311, 324)] are the edges that have the most edge-betweenness.

We discover that node 324, 325, 328, 329, 330, 331 have repeatedly occurred in the above result.

324 locates in Puerto Rico (overseas territory) with a degree 4;

325 locates in Virgin Islands (overseas territory) with a degree 8;

328 locates in Rota Islands (overseas territory) with a degree 2;

329 locates in Guam Islands (overseas) with a degree 4;

330 locates in Palau (overseas) with a degree 2;

331 locates in American Samoa (overseas) with a degree 2.

We notice that almost all of the possible airlines of the above nodes have a high edge-betweenness while these nodes have small degree. Also, those with high node-betweenness tend to **influence its related edge-betweenness**!!!

To sum up, **overseas territories have strong impact on edge-betweenness**!!!However, node-betweenness has a closer relationship to node-degree, the larger the degree, the larger the node-betweenness may be. Overseas territories as well as Alaska have **some** influence on node-betweenness.

6. coreness of the network is 26, with 35 nodes’ coreness equalizes it.

['67', '112', '118', '201', '248', '166', '94', '109', '131', '147', '150', '152', '172', '176', '177', '182', '219', '230', '232', '255', '258', '261', '293', '311', '146', '159', '162', '167', '174', '179', '217', '292', '299', '301', '310']

Among them, 18 are in the top-20-biggest-degree list. Clear evidence that a large coreness node is inclined to be a large degree node.

(Those who are not the top-20-biggest-degree nodes, generally have relatively big degree.

{'94': 34, '109': 50, '131': 45, '150': 46, '172': 35, '177': 48, '219': 49, '232': 44, '311': 47, '146': 50, '159': 28, '167': 41, '174': 53, '179': 42, '292': 33, '301': 42, '310': 34})

**144 and 47** are top20 biggest degree nodes, but not the biggest coreness nodes. Their coreness, respectively 22 and 24, which is not low.

7. Distance between big nodes: Is the distance between small nodes generally longer than that between big nodes? We consider those degree <=5 are small nodes, those degree >=30 are big nodes. If nodes are not directly connected, we use the Floyd algorithm to compute shortest path length.

We found that the average distance between big-big nodes is **1043** miles with 741 pairs of big-big nodes, **much lower** than the Average Path Length 2033 miles; Distance between small-small nodes is 2525 miles with 15753 pairs; Distance between big-small nodes is 1985 miles, very close to the APL with 6942 pairs; Distance between ‘ordinary’ nodes (5<degree<30) is 1821, also close to the APL with 31510 pairs.

We can infer that:

① Big-big nodes tend to be very close. The farthest big-big nodes are 47 and 311, locating in Seattle (Northwest of the US) and Miami (Southeast of the US).

② Small-small nodes tend to be far away. The farthest two have been illustrated before (see 1.); Small-small nodes have a relatively big proportion among the total pairs, so we can say that small-small nodes have a **significant influence on APL**!

8. Investigating flight frequency. (the frequency is **normalized**)

Firstly, we investigate big-big degree (top20) nodes.

('152', '144'), ('230', '144'), ('147', '144'), ('293', '176'), ('162', '47'), ('144', '217'), ('47', '299'), ('47', '217') have no direct edge so their frequency is 0.0. **144** lies in **Salt Lake City.**

From 152 to 144 needs **1** transfer (转乘), and **so do other cases**. So this phenomenon may be for the sake of conserving money.

If omitting those 0.0 cases, the average frequency between big-big (top 20) degree nodes is 0.209

Then we perform such on big-big nodes (degree>=30), small-small nodes (degree<=5) and big-small nodes. The average frequencies of the above (only consider if nodes are adjacent) are respectively: 0.195, 0.032, 0.077 (the average is 0.134). We can know that on average, the flight frequency between big-big nodes are much higher than that between small-small nodes.

Additionally, we investigate those pairs with high frequency, to find out some relationship between frequency and distance/area.

Top 20 highest frequency pairs:

[((248, 331), '1.000'), ((118, 313), '0.765'), ((67, 313), '0.749'), ((255, 313), '0.736'), ((8, 112), '0.722'), ((8, 313), '0.718'), ((182, 313), '0.717'), ((8, 118), '0.686'), ((313, 331), '0.651'),

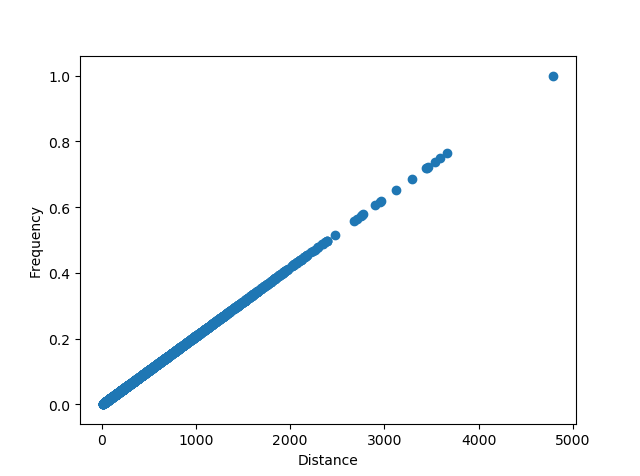
((261, 313), '0.617'), ((8, 67), '0.616'), ((166, 313), '0.605'), ((47, 313), '0.579'), ((47, 316), '0.576'), ((8, 248), '0.574'), ((313, 329), '0.565'), ((47, 311), '0.558'), ((8, 144), '0.515'), ((47, 109), '0.498'), ((109, 201), '0.498')]

Node 248: Los Angeles Intl (degree 59); Node 331: Pago Pago Intl (Overseas, degree is 2) **Strange!!!**

Node 118: Chicago O’Hare; Node 313: Honolulu in Hawaii (Overseas, degree is 24) Also Strange!

Note that **node 313 appears repeatedly!** 313 is the capital of Hawaii, which is a **tourism resort**. Thus, it may be inferred that a big airport may have one or some high-frequency airline linking to those resorts (especially some islands).

Then I investigate the relationship between distance and frequency (if two nodes are adjacent).



**STRANGE!** How to explain it???