**Degree Dist**

**Indegree dist**

**OutDegree dist**

Most of the graphs follow the power law. But for some of the graphs, we see strange patterns. We discuss these below.

1. Gnutella P2P: The Out-degree distribution is not a power-law pattern and it has sudden jump in the middle of the picture. It means that about 10 machines connects to all of the machines in the p2p cluster. It might be some master file service machines that stores all of the files in the cluster.

2. JDK dependency: The Out-degree distribution is not a power law pattern, and In-degree distribution shows lots of waves at the tail. We inferred that in Java the relationship between classes is not just simple independent as Java has inherit, interface, implementation. For all the classes, the basement class is Object.class.

3. Flickr Social: From the Out-degree we found about more than 40000 users who has zero out degree, but from In-degree we found that only 1000 users has zero in degree. This shows that in Flickr most users just upload their photos but they don't care others so they don't follow others. And there are some users who follows large amount of users. Most of the users are followed by some maybe administrative account as most of the users have at least one follower.

4. Pokec Social: The same observation with Flickr also takes place on Pokec.

5. Email EuAll: From the in degree we found that there are almost 200000 accounts that receives zero email, which shows a large amount of mistyped, no-existing or spam email addresses.

**PageRank**

We analyzed the pageRank score and the number of nodes with that score.Half of the graphs show the pattern of power-law, but others are not which described as below.

1. Digg Social: There are several small jump in the descending trend.

2. Flickr Social: the points are more scattered and distributed sparse compared with power-law. As we have mentioned above in degree distribution, due to the social relationship that most of the users are not following anyone else. Or those accounts are zombies accounts.

3. Hamsterster Social: The trending is more or less like power-law but it is not a standard one.

4. Pokec Social: Like Flicker, the points are more scattered. due to the social relationship that most of the users are not following anyone else. Or those accounts are zombies accounts.

5. Youtube Social: After a few ascending, it follows the power-law. We inferred that the reason is most of the Youtebe users using Google account to log in and do the social work. As not all of the google account are active users, there are small amount of very inactive user account which combines the start of the line.

6. Gnutella P2P: As in a P2P network, every node is totally equal to others, there is no way to say which one is more important or not. But there are pagerank difference is just because maybe some of the users has more files than others, but if they are normal users, no one has extraordinary numbers of files unless it has some purpose.

7. JDK dependency: As described in degree distribution, JDK classes has its internal dependency following Obejct Oriented Language Design.

8. Hepth Citation: There is a small ascending trend and then goes with power-law.

**Connected Component:**

We analyze the component size and the number of components. We can observe a power relation for components if there are hundreds of components. In all the graphs, we see that either the graph is fully connected, or it has too little number of components or it follows the power-law.

1. Fully connected graph:

JDK dependency is fully connected because it needs to follow JAVA rules.

Spanish Text is fully connected because it is a graph from words of a book. And words in the book are continuous, which means that all the words are of course in the place one by one and must be fully connected.

Caida AS is fully connected because all the ISP must be connected to the internet network, and make sure all their customers can reach out to where ever they want in the universe, so of course they are fully connected.

2. Few Components:

If a graph shows there are few components, it shows that there are strong connection relationship between nodes. Like in Digg, Hamsterster, Pokec, their users are strongly connected to each other. And for Gnutalla P2P, the number of components is just 12. In protein biology graph, it shows that the protein structure is complicated as they highly connect to each other. And in Hepph and HepTh citation, as they are focusing on papers on High Energy Physics which is really high technical, there are not so many papers like some other fields, so the paper citation is limited.