NETWORK ANALYSIS

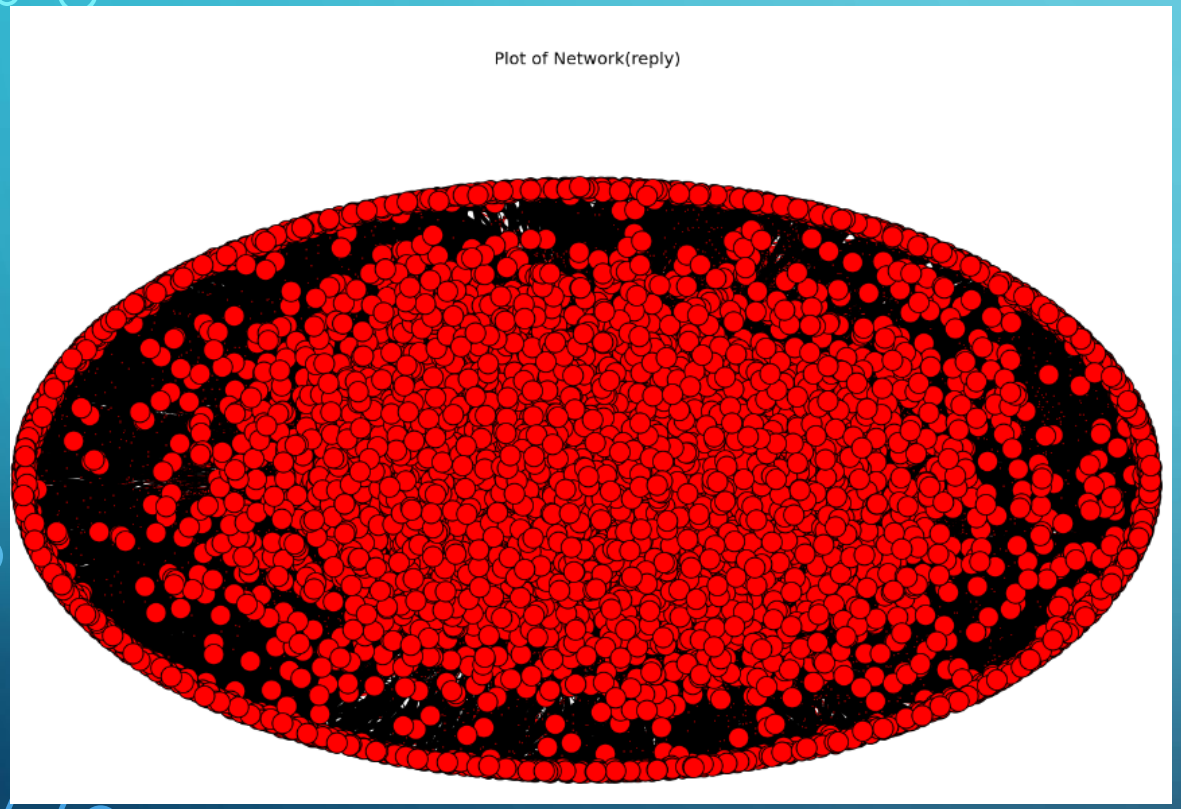
DATASETS AND TOOLS

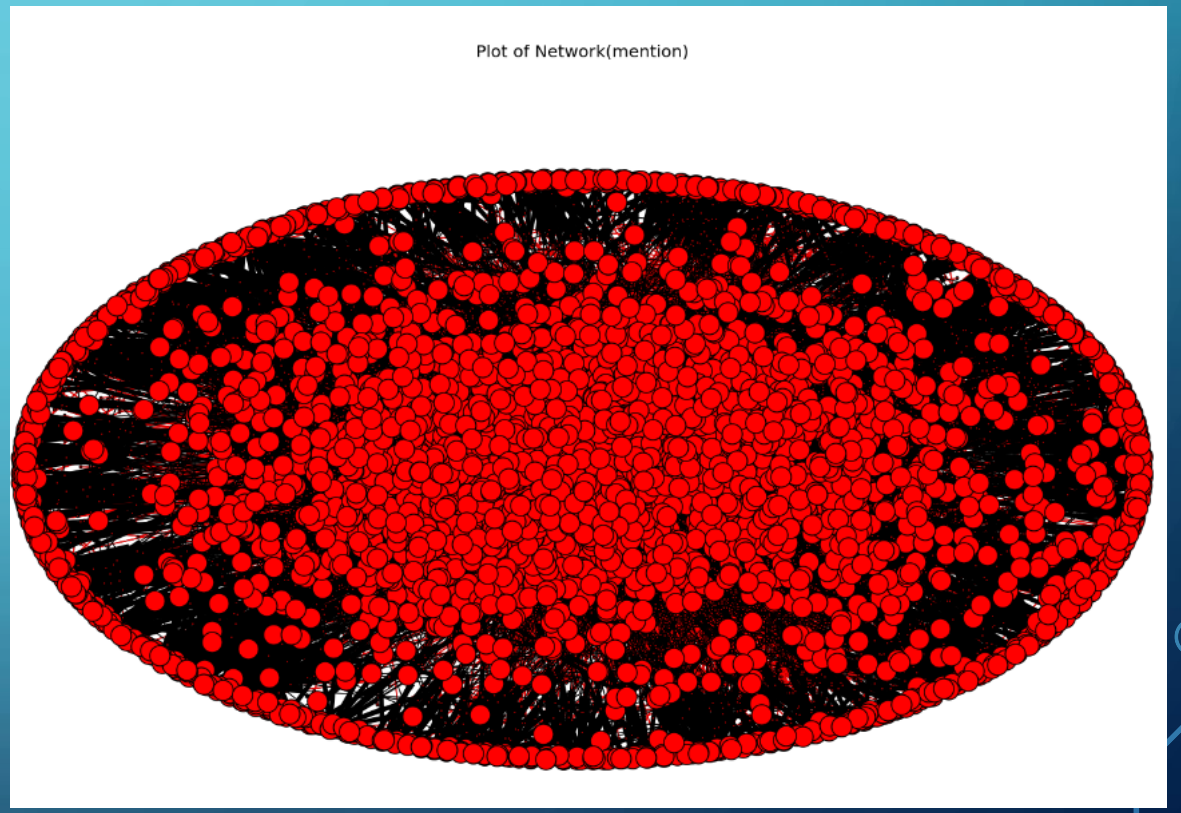
- Data: mrredges-no-tweet-no-retweet-poi-counted

Graph	Nodes	Edges	Self-loop	Density	Is_connected
DiGraph of Reply	4038	28992	0	0.00177849689775	NA
DiGraph of Mention	3513	17227	0	0.00139629440146	NA
UnDiGraph of Reply	4038	22062	0	0.00270676038619	False
UnDiGraph of Mention	3513	15429	0	0.00250112339005	False

- Tools: Python, Networkx, Matplotlib

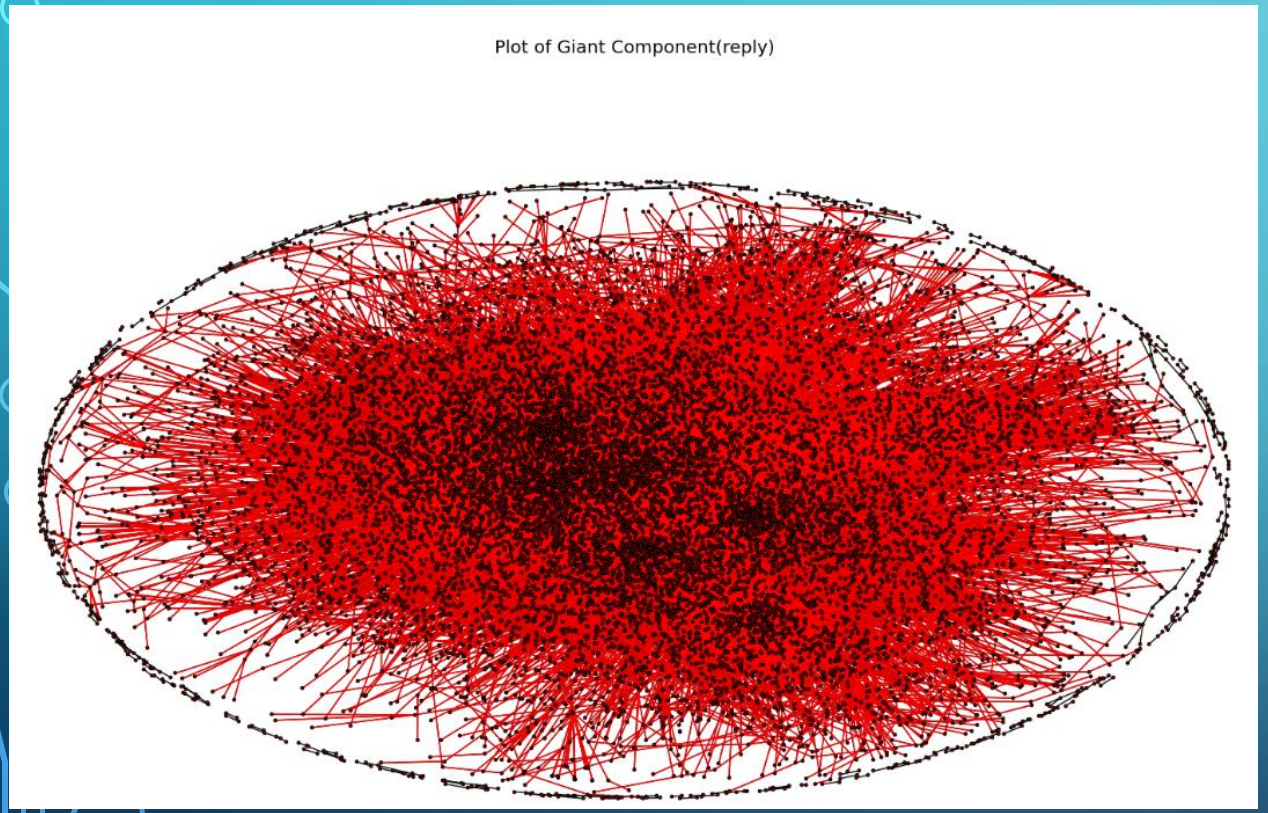
DIGRAPH TOPOLOGY OF NETWORK



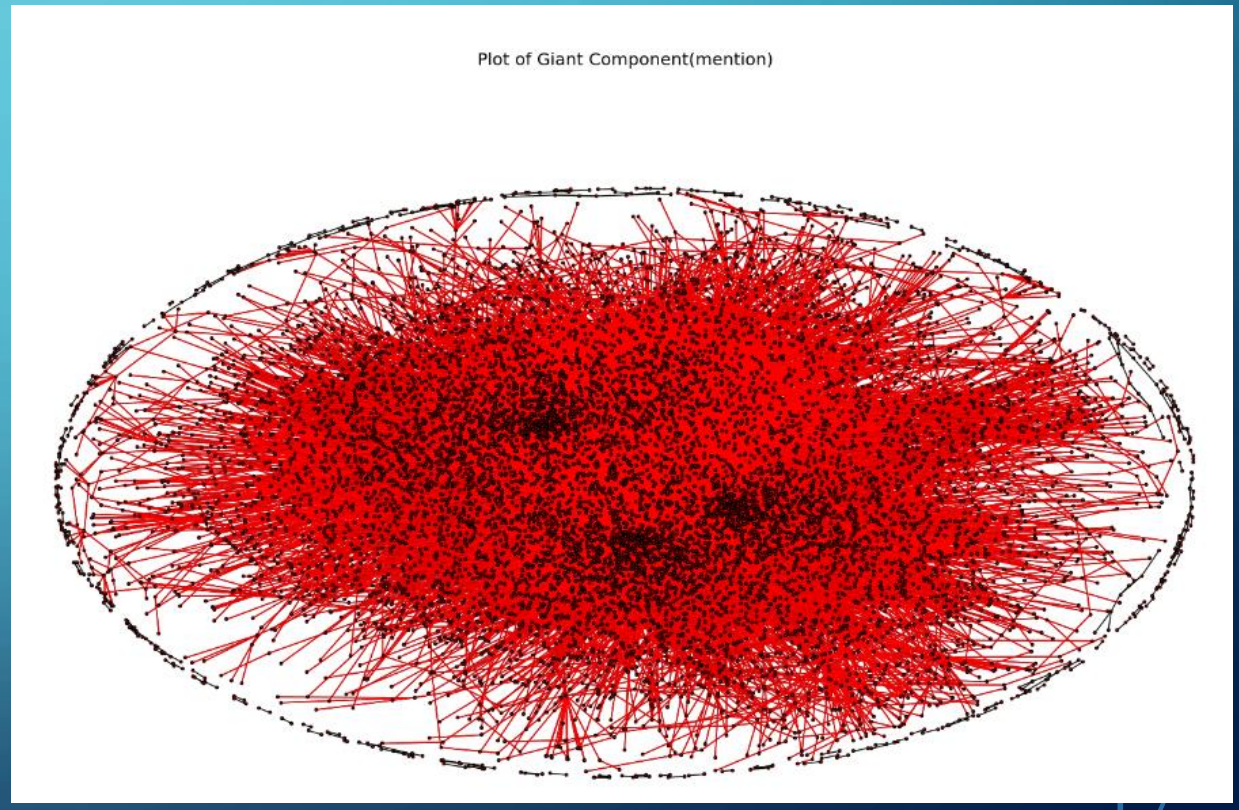


The nodes in mention network seems more partitionable than those in reply network.

COMPONENTS (UNDIGRAPH)

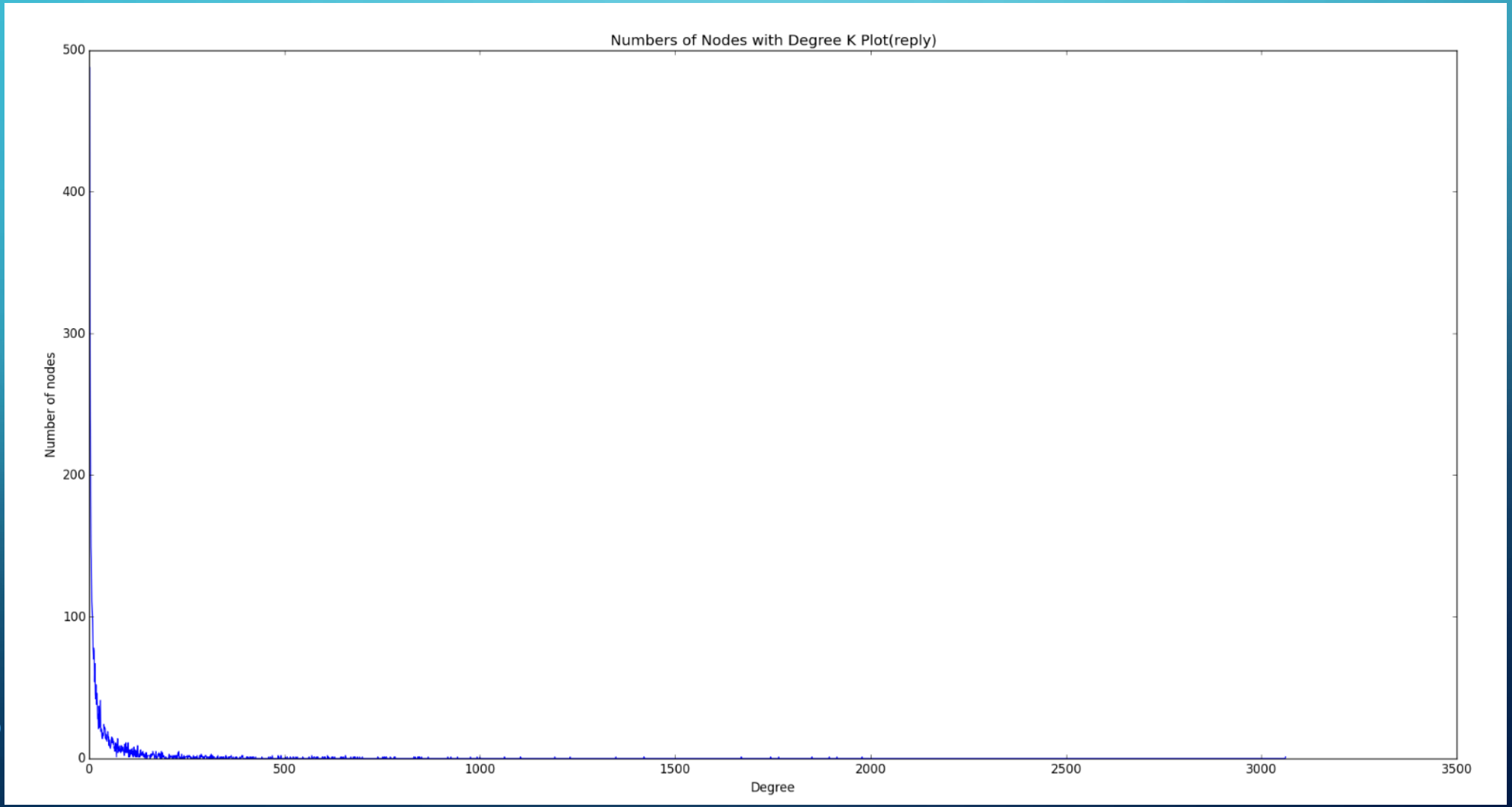


number of components in network: 66
the size of giant components: 3892/4038 ratio: 96.4%

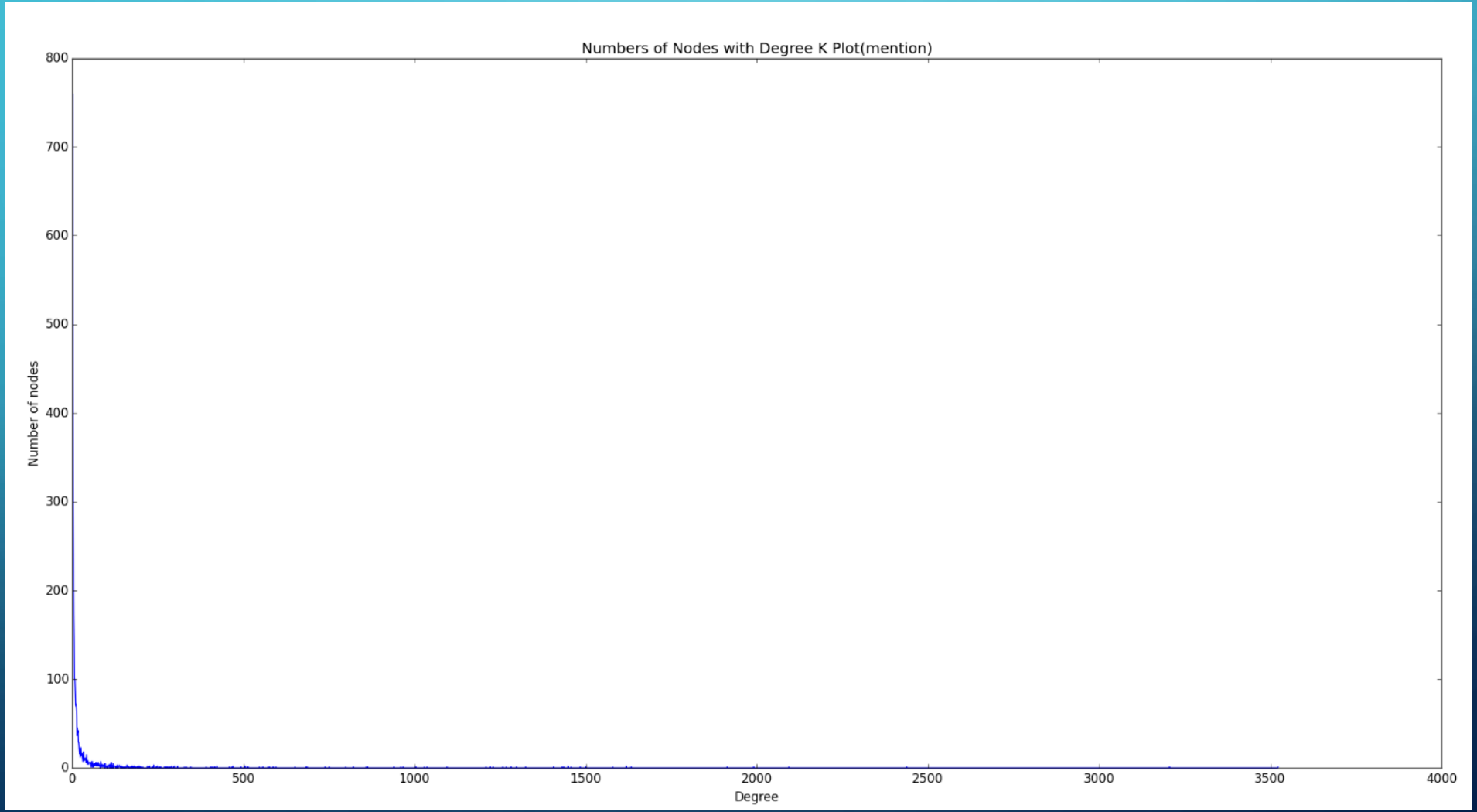


number of components in network: 80
the size of giant components: 3332/3513
ratio: 94.8%

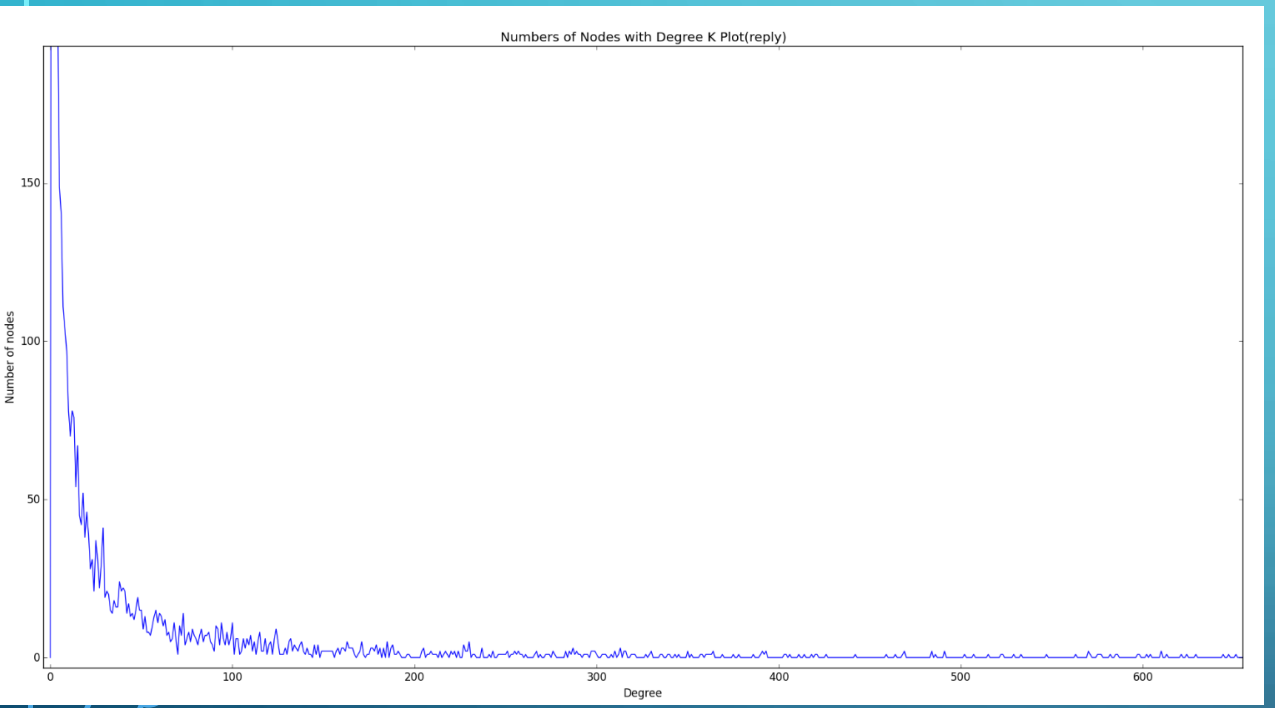
DEGREE DISTRIBUTIONS IN DIGRAPHS

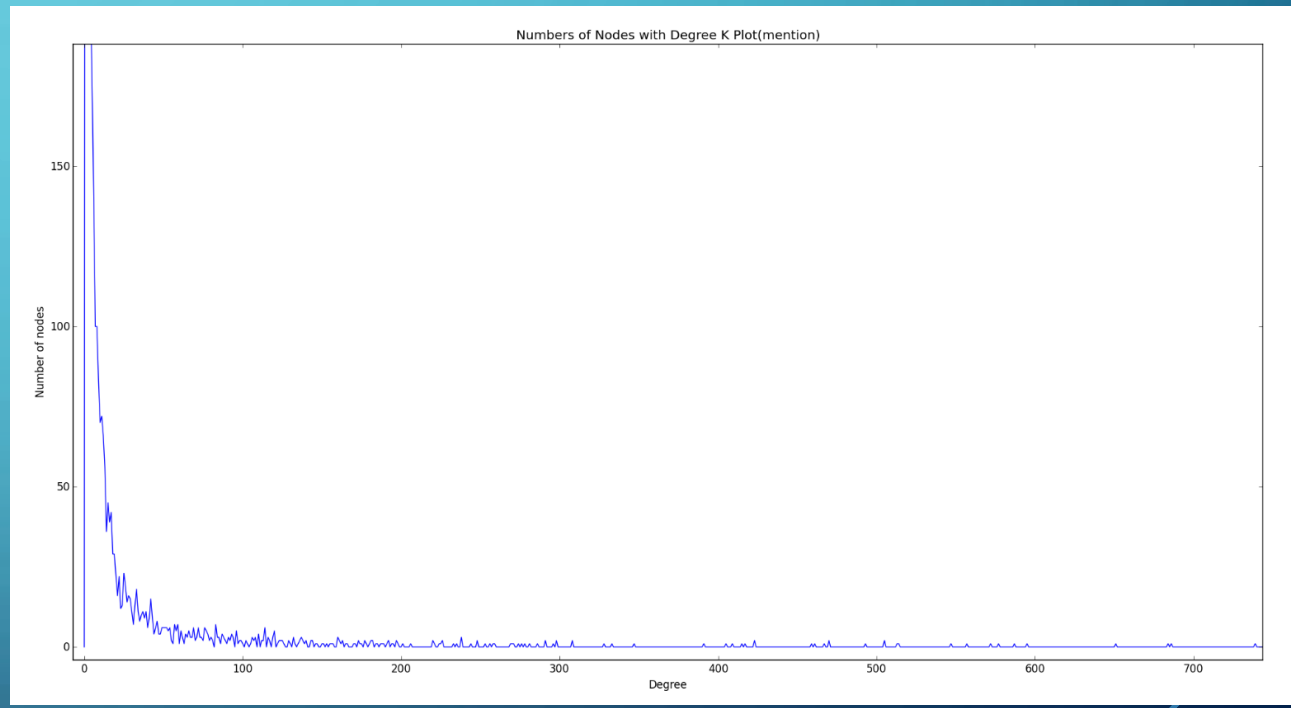


DEGREE DISTRIBUTIONS IN DIGRAPHS



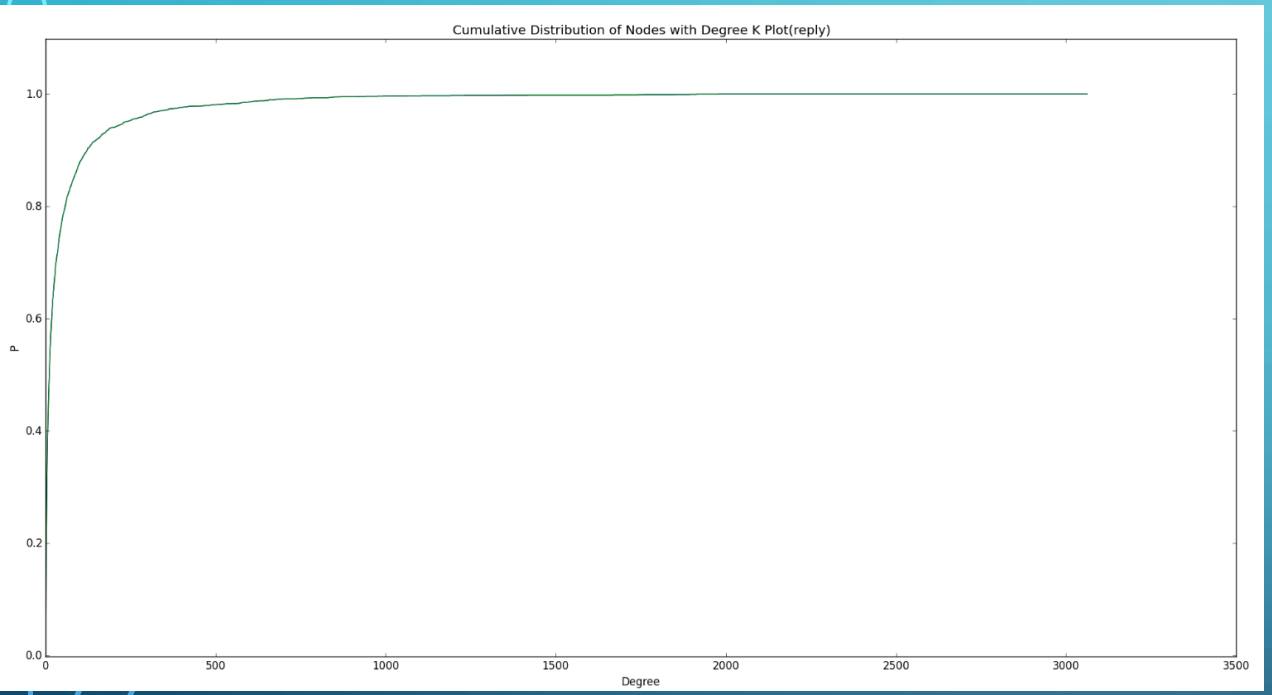
DEGREE DISTRIBUTIONS IN DIGRAPHS

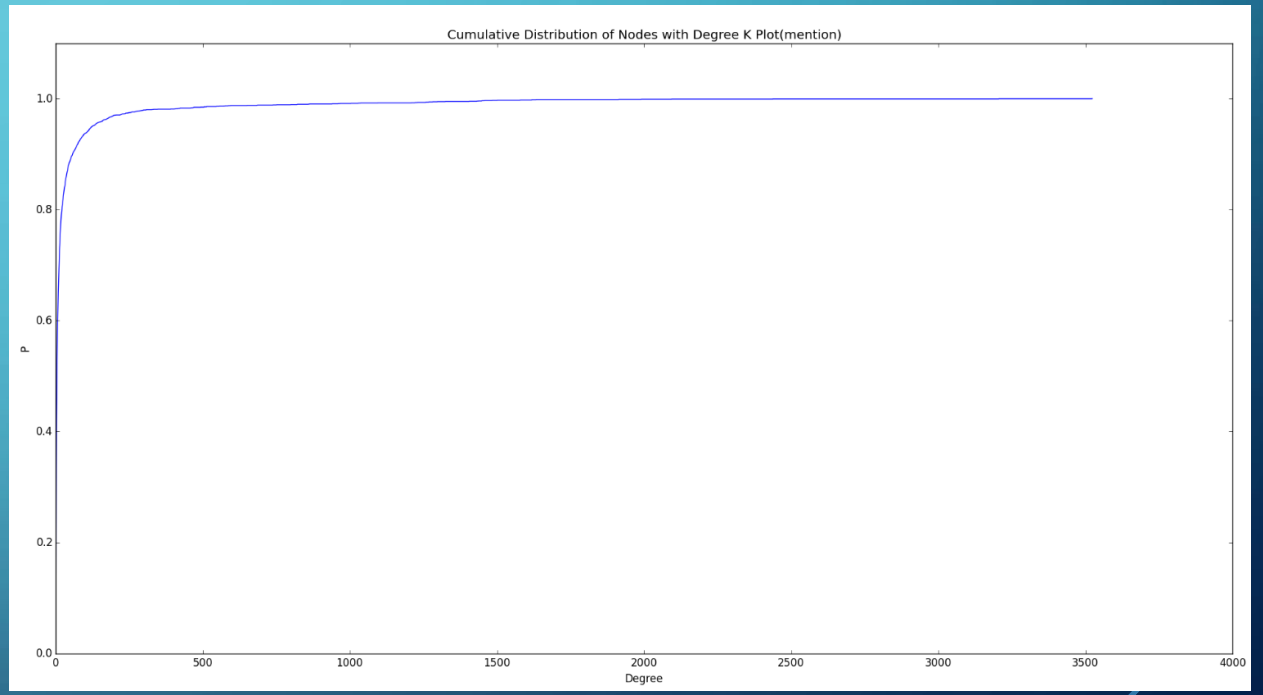




The numbers of nodes with the degree within [20, 75] in reply network are more than those in menti

DEGREE DISTRIBUTIONS IN DIGRAPHS

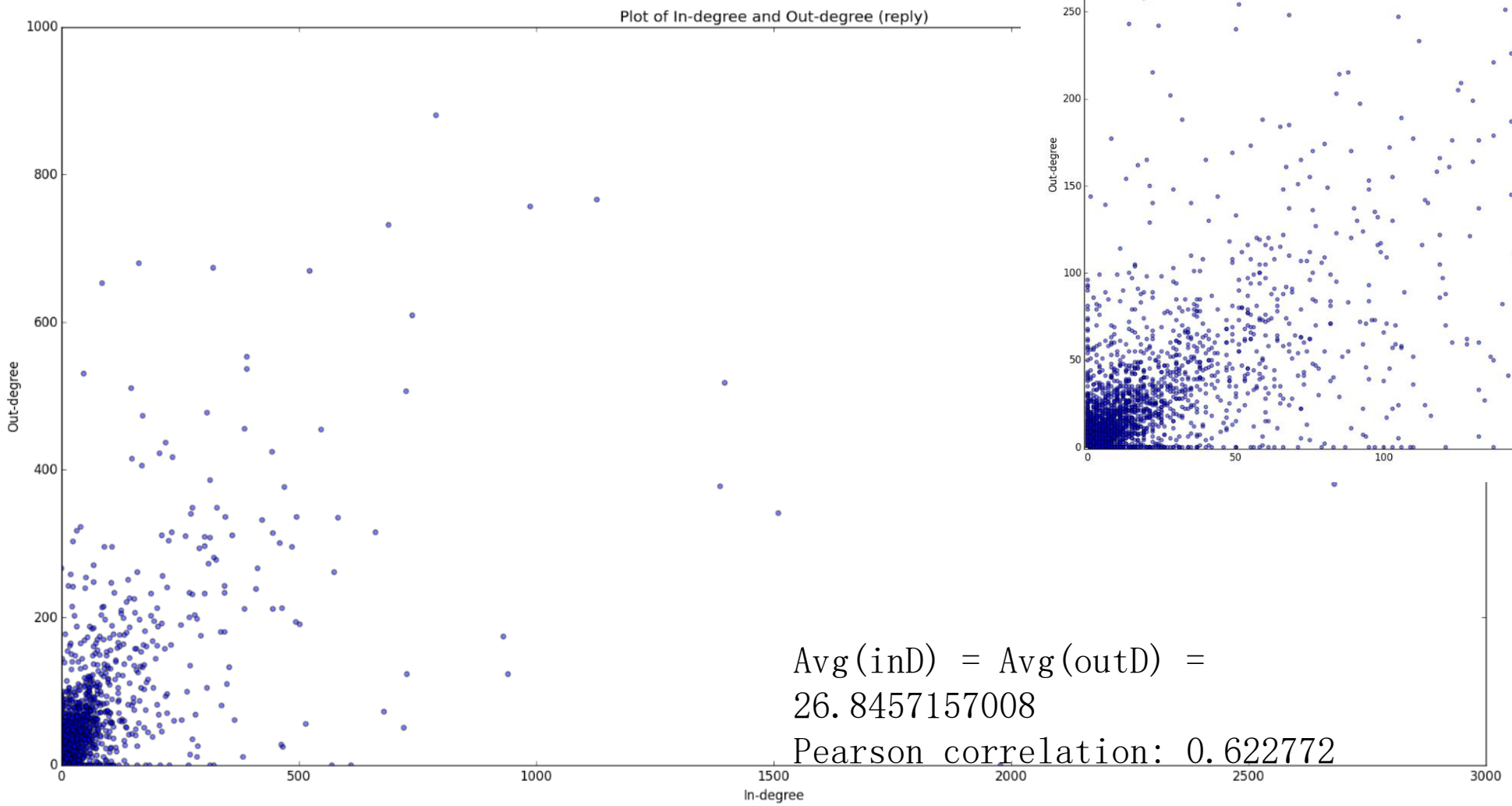




Hence, the cumulative probability of mention network reaches its peak in the smaller range of degrees t

→ How are their tail distributions?

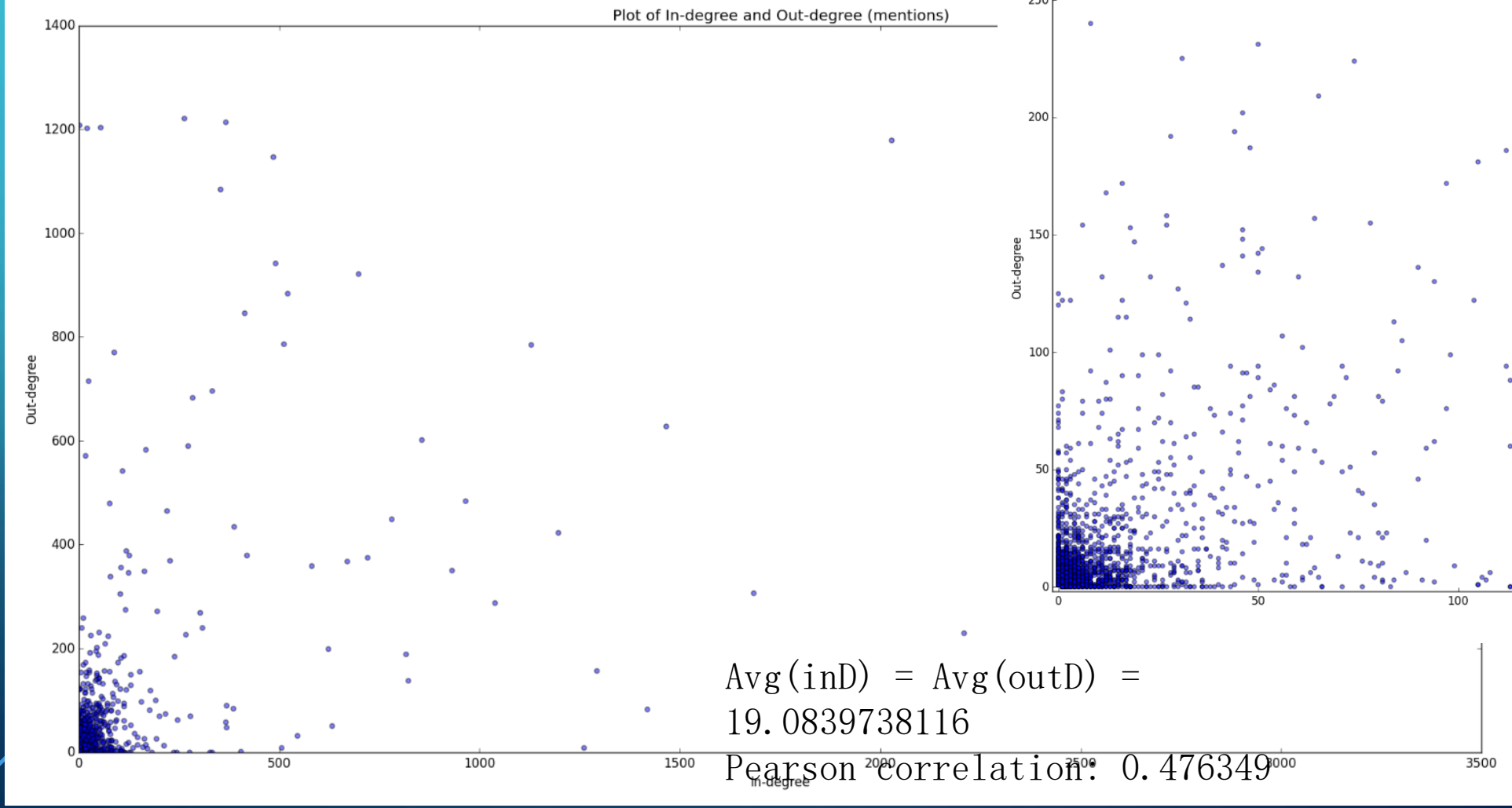
RELATIONSHIP BETWEEN REPLY INDEGREE AND OUTDEGREE



Avg(inD) = Avg(outD) = 26.8457157008

Pearson correlation: 0.622772

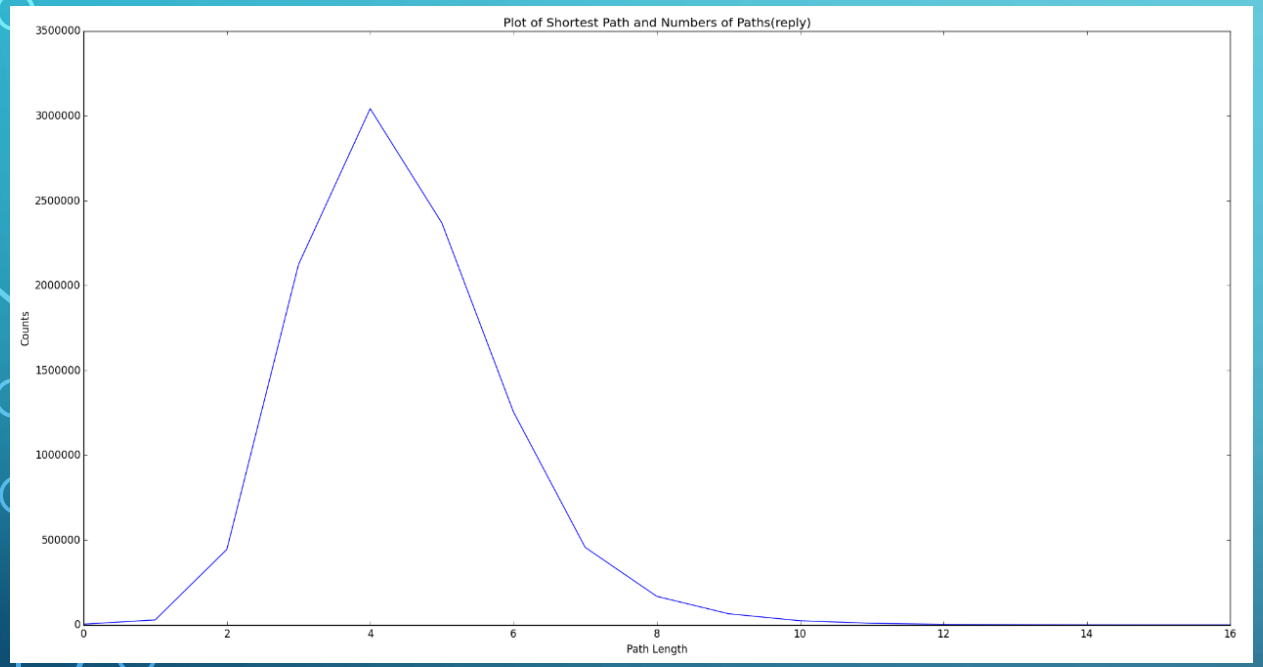
RELATIONSHIP BETWEEN MENTION INDEGREE AND OUTDEGREE



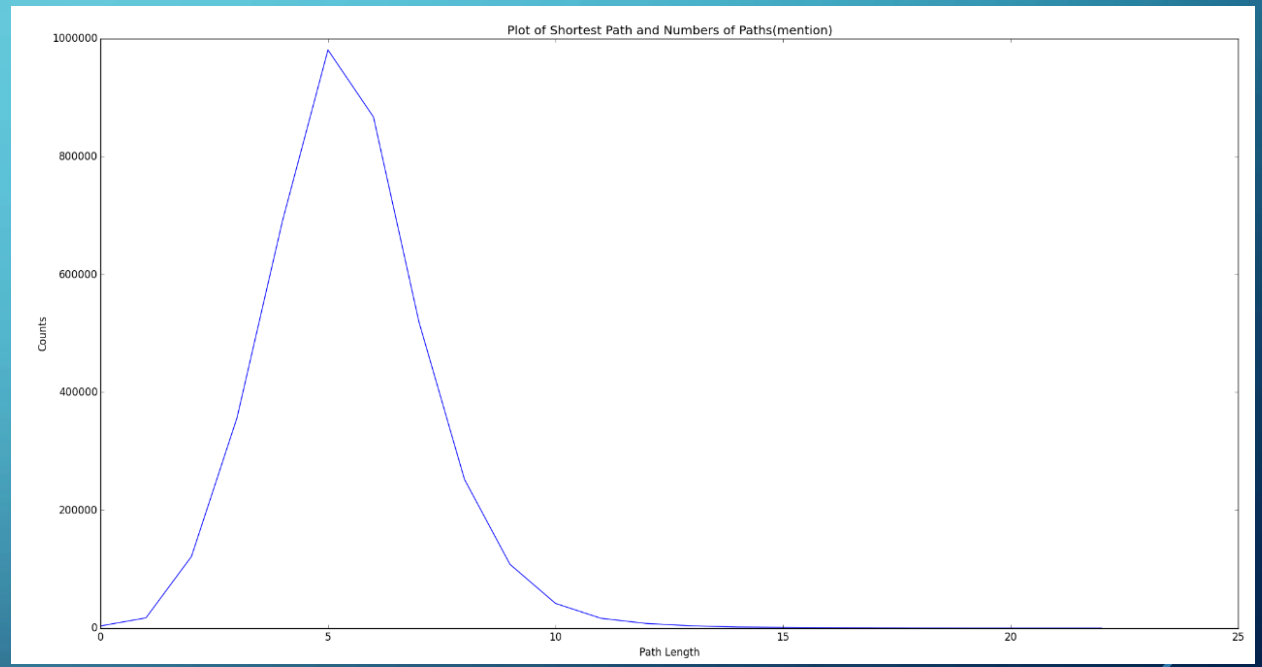
Avg(inD) = Avg(outD) = 19.0839738116

Pearson correlation: 0.476349

DIRECTED PATH DISTRIBUTION OF USERS



The mean of path length is 4



The mean of path length is 5

The path of mention is generally longer than that of reply

→ the reachability of reply is typically higher than