Predicting The High Proportion of People Whose Average Friend Has More Friends Than They Have (PLU)

Alec McGail am2873@cornell.edu
Scott Feld sfeld@purdue.edu

BACKGROUND

Feld (1991) showed that friends inevitably have more friends, than people do, on average, as a simple function of variation of degree.

The FRIENDHIP PARADOX was interpreted to predict that most people would find that their OWN friends would have more friends, on average, than they do.

Research on many friendship networks has shown strong consistent evidence that a large majority of people do find that!

But, the variation in degree is NOT the ONLY reason. We consider the additional impact of skewness of the degree distribution and of assortativity in the network.

THEORY

We theoretically predict that the Proportion Looking Up (PLU) will be well predicted simply from the distribution of degree.

We specifically describe how the variability and skewness of the degree distribution will affect the PLU.

Feld (1991) showed the mean number of friends of friends is:

M + V/M
M-Mean degree
V- Variance of degree

We expect that the proportion of people with degree less than this mean number of friends of friends will predict PLU, irrespective of the extent to which people with many friends tend to have friends with many friends (assortativity).

DATA

We used data from the network of 63,731 New Orleans Facebook friends.

From this network, we collected subsamples by choosing a random person, and their friends as well as their friends' friends. We only include such networks if they have between 100 and 2000 individuals. We took 3000 of these random samples and discarded all but 529 based on the above criteria, mostly for being too small.

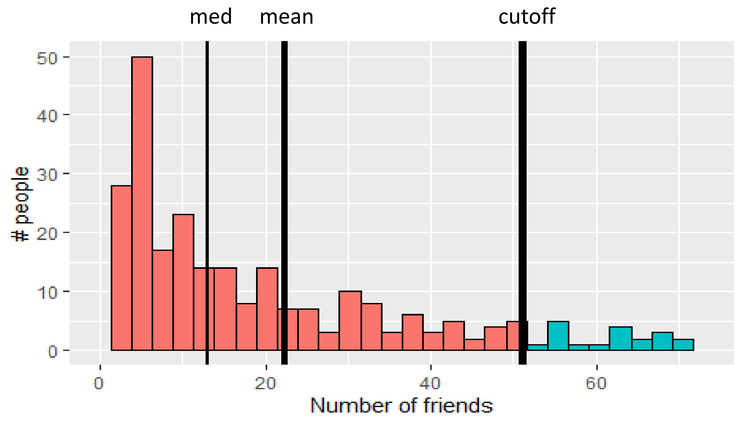


Figure 1 This plot shows the degree distribution for a single random subnetwork. The median, mean, and predicted cutoff (M+V/M) are labeled. People with degree less than the cutoff (those with red bars) are expected to be "looking up."

	variation	skewness	assortativity	our prediction	PLU
variation	1.00	0.31	0.01	0.40	0.45
skewness	0.31	1.00	0.01	0.44	0.49
assortativity	0.01	0.01	1.00	-0.48	-0.49
our prediction	0.40	0.44	-0.48	1.00	0.91
PLU	0.45	0.49	-0.49	0.91	1.00

Table 1 Correlation matrix among network characteristic of the subnetworks (n=529).

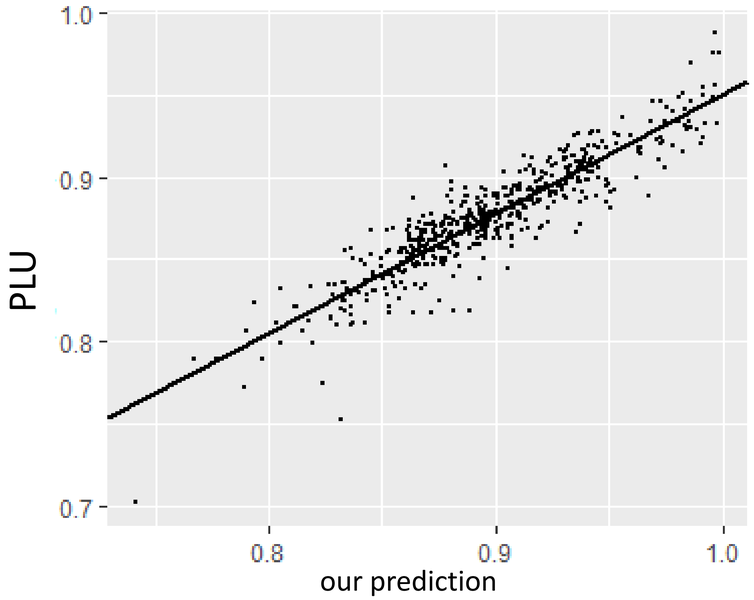
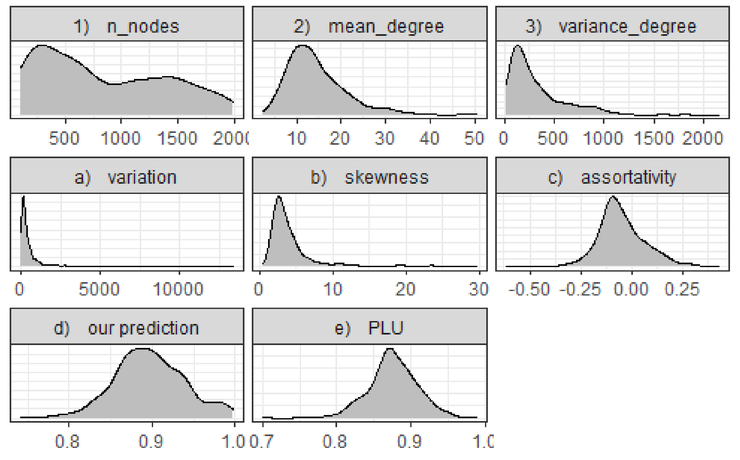


Figure 2 (above) is a scatterplot showing our predicted proportion and the actual proportion of individuals with less friends than their friends, on average. Each point on this plot represents a subnetwork of the New Orleans network. The solid black line shows the regression between these two quantities. The figure shows the high accuracy of the predictions (r=.91).

Figure 3 (below) shows that there was wide variation among the 529 subnetworks in each of the variables we analyzed.



DEFINITIONS

a) variation – (M/V)**2, where M and V are the mean and variance of the degree distribution, respectively

b) skewness – The difference between mean and median degree, divided by the mean degree.

c) assortativity – The correlation between degree and mean egonet degree. That is, between a randomly selected person and the average number of friends of their friends.

d) our prediction – The proportion of individuals with less than M+V/M friends.

e) PLU – The proportion of nodes whose friends have, on average, more friends than they do.