James Beavers

jjbeaver

CSC591

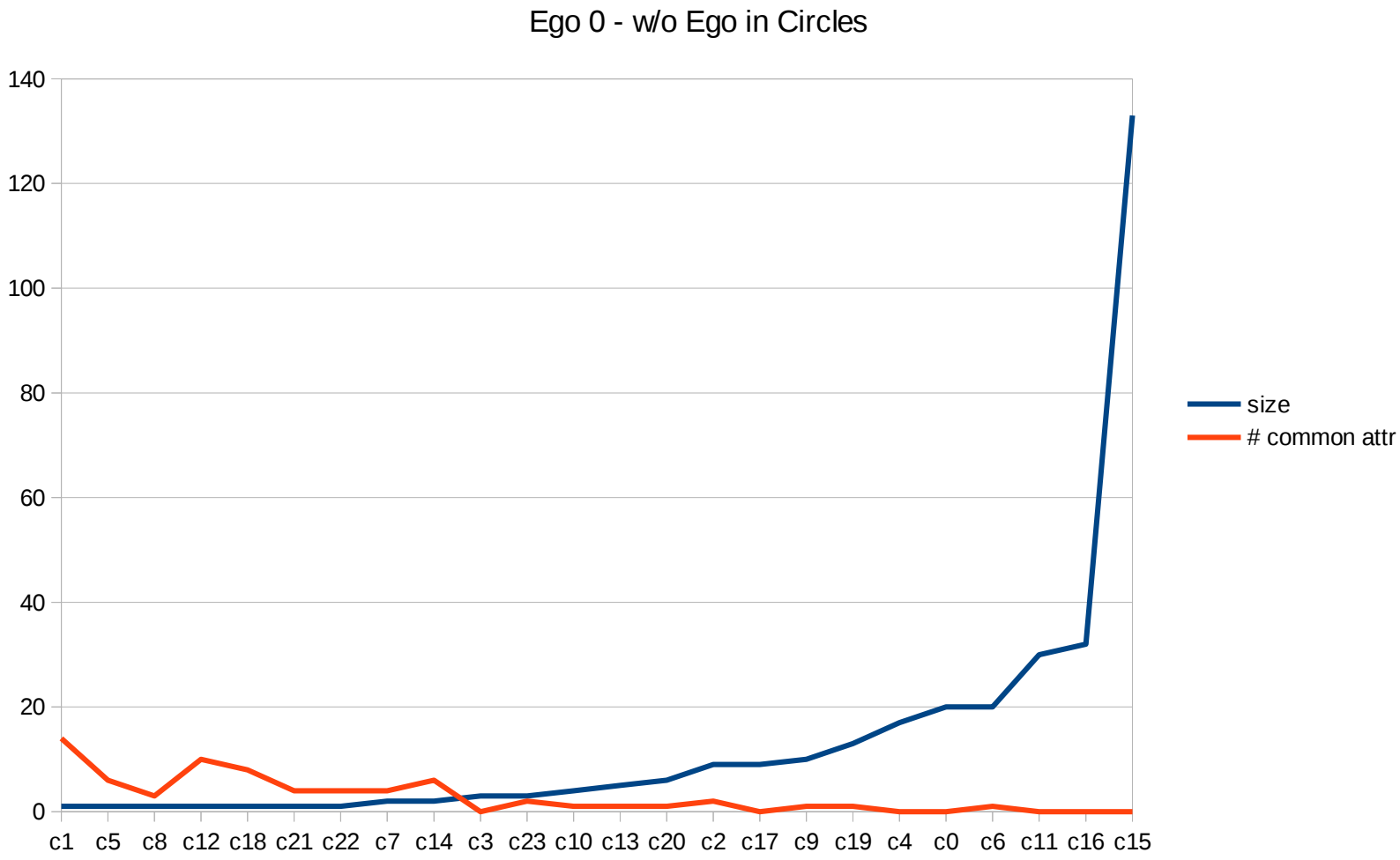
Project P1

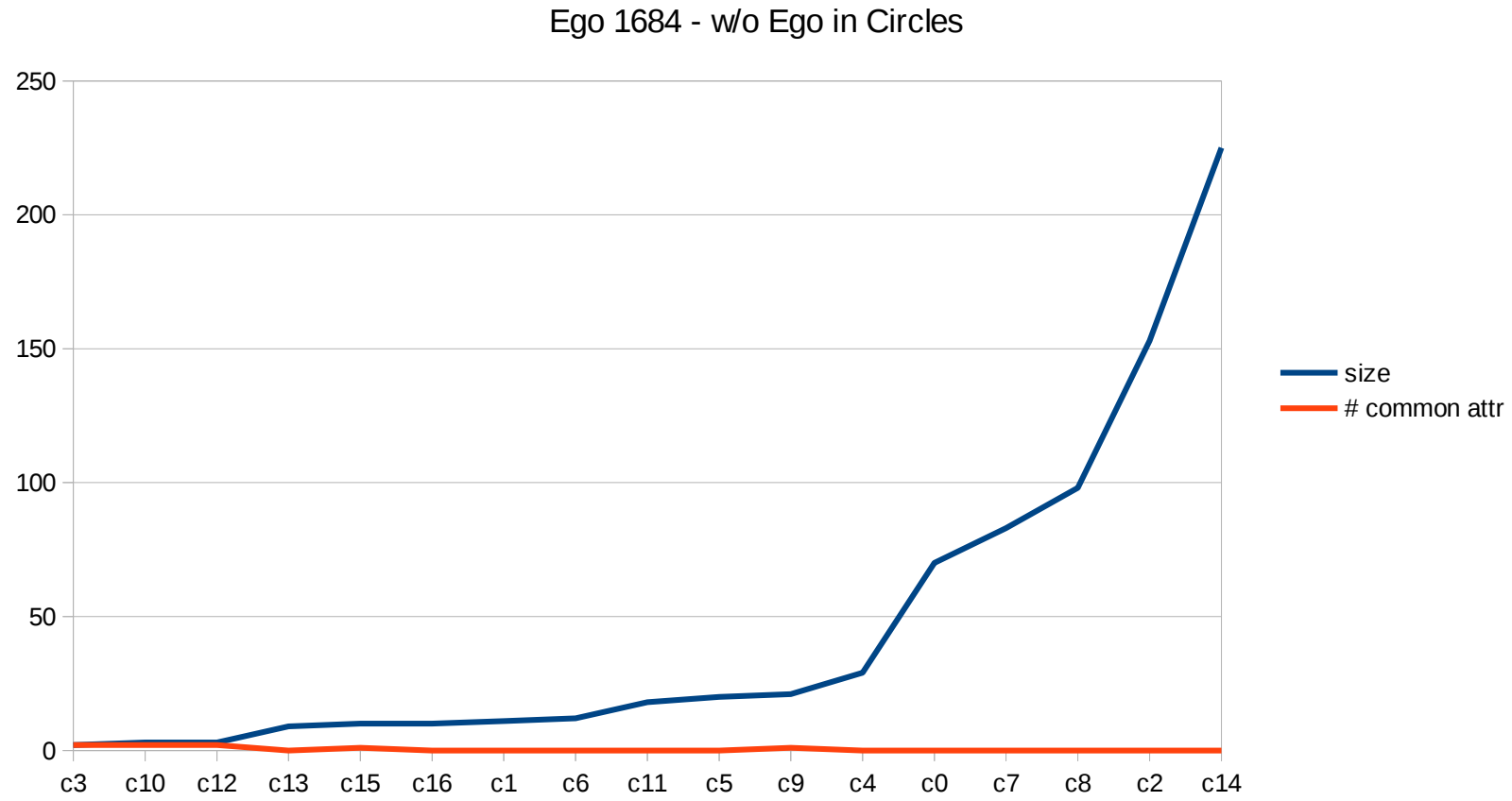
Due: 2/24/15

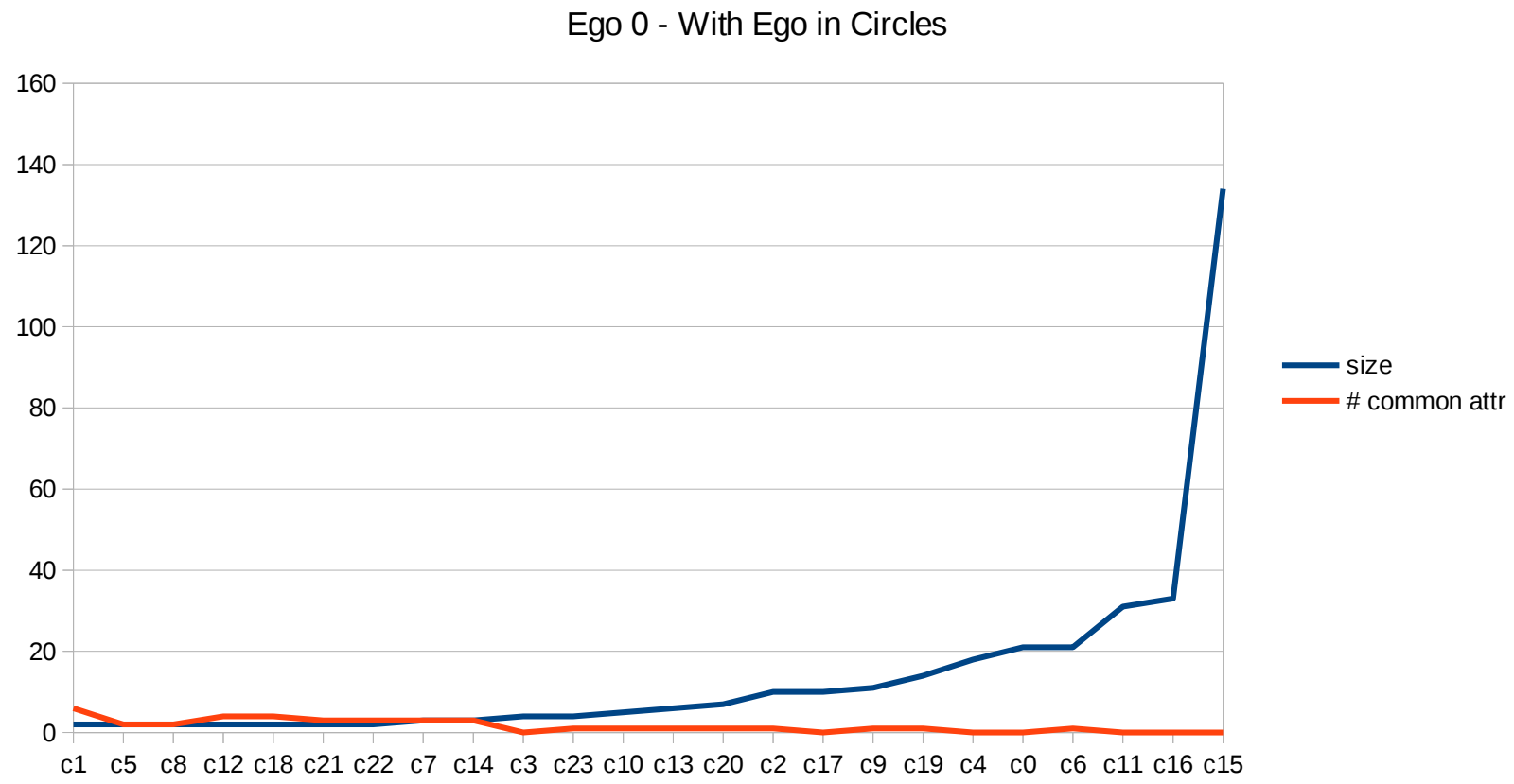
Experimental Results

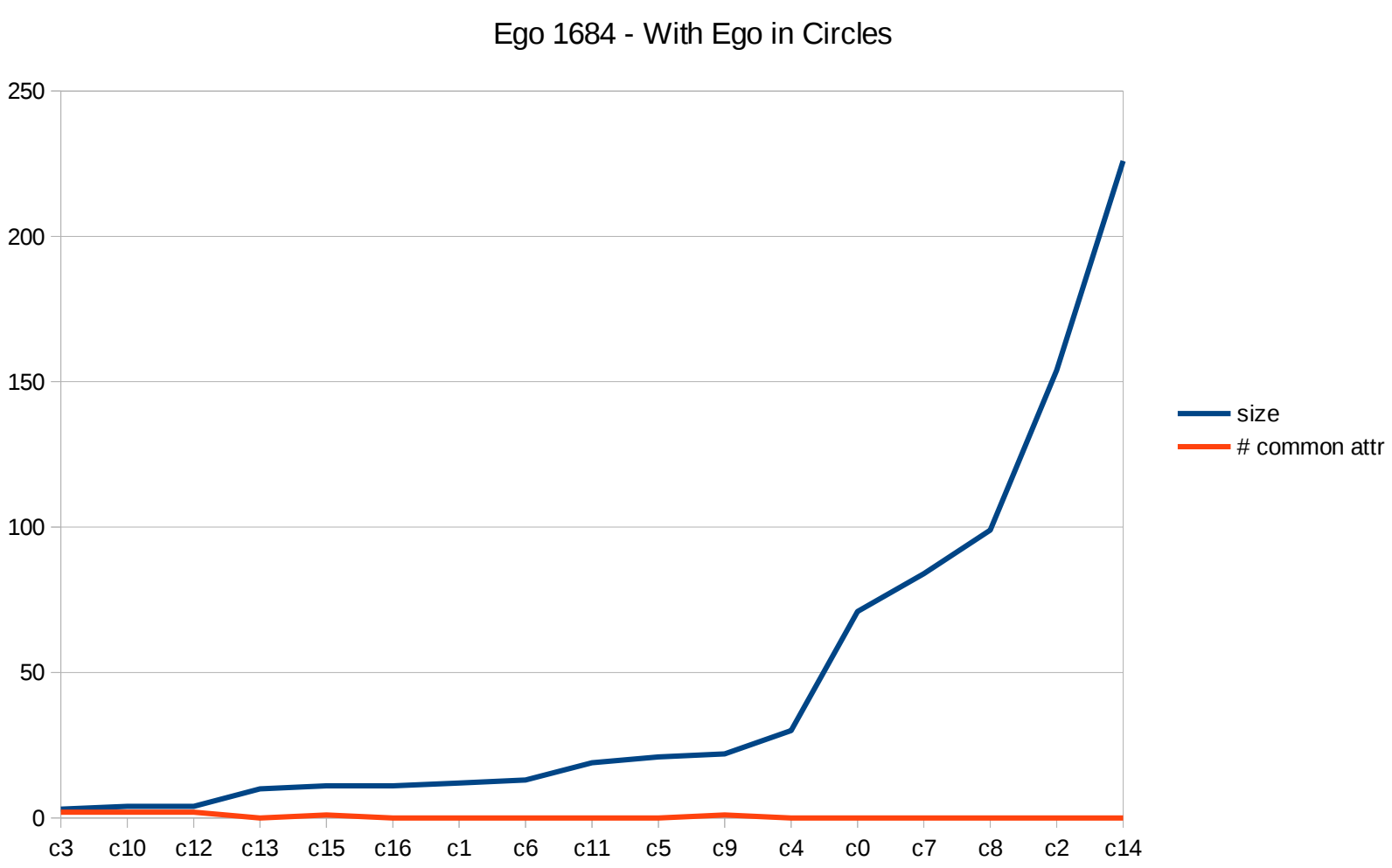
<amadan2> – <1> “On average, as the size of the circle increase, the common features between users in the circle decreases.”

i) Given the sample data and egos we used to test this hypothesis, it seems plausible that this hypothesis might be true. As the size of circles increased, the number of common attributes among the group seemed to decrease. See relevant graphs below:




Ego 1684 - w/o Ego in Circles
250
200
150
100
50
0
c3 c10 c12 c13 c15 c16 c1 c6 c11 c5 c9 c4 c0 c7 c8 c2 c14
size
common attr


Ego 0 - With Ego in Circles
160
140
120
100
80
60
40
20
0
c1 c5 c8 c12 c18 c21 c22 c7 c14 c3 c23 c10 c13 c20 c2 c17 c9 c19 c4 c0 c6 c11 c16 c15
size
common attr



From the above graphs, we can see that clearly, as the size of the circle increases, the number of common attributes among the circle approaches 0 very quickly.

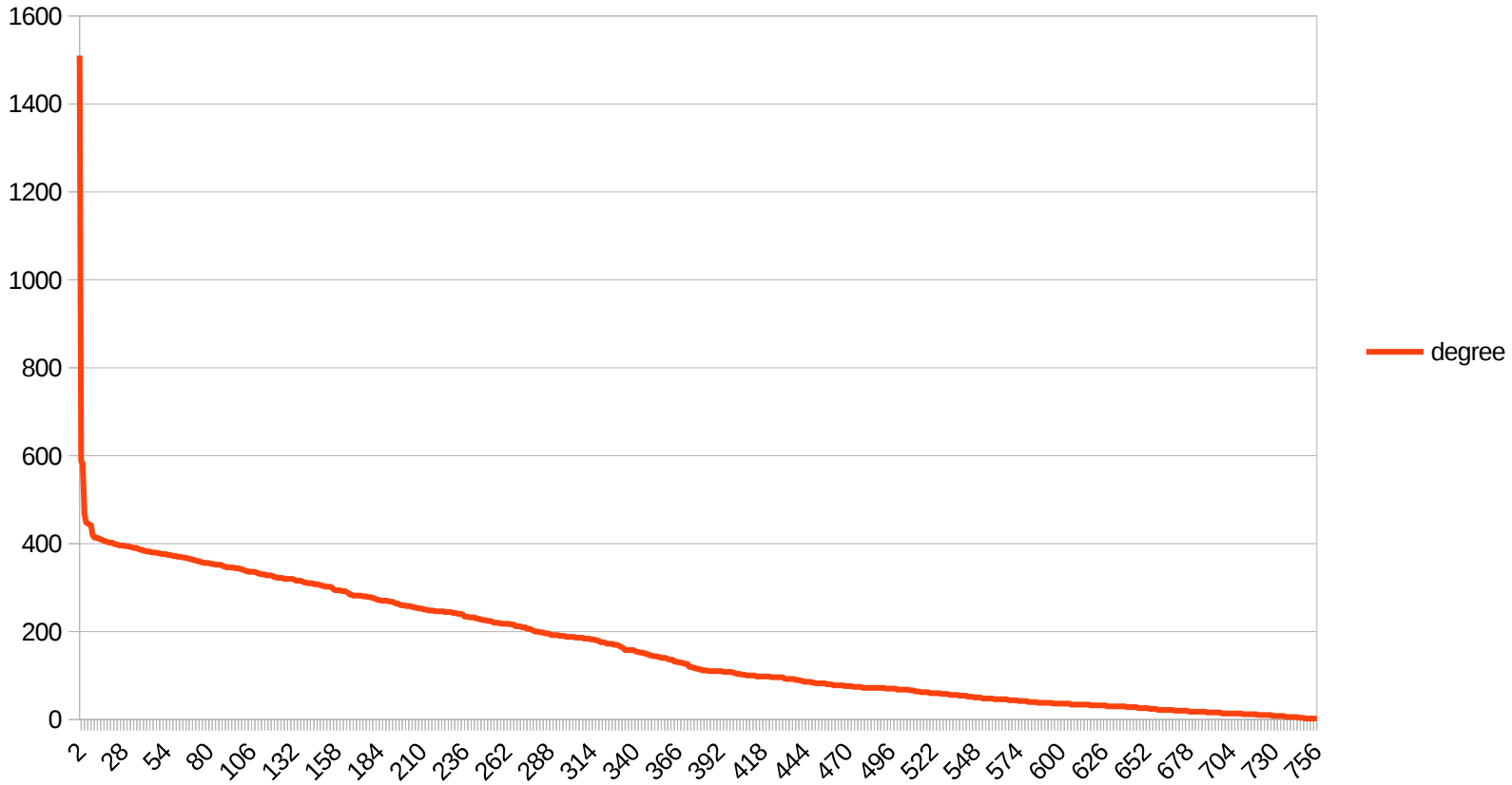
ii) In general, there were a lot of circles that had 0 common attributes. Additionally, each individual didn't necessarily have many features defined. Some users had as many as 32 features (out of the 0.circles users), while some only has 1 or 2 features. This means that any circle (regardless of its size) that contains a user of only 1 or 2 features, the resulting # common attributes is limited to a maximum of 1 or 2. This could be due to users not properly defining all of their profile attributes, etc.

Though this analysis seems somewhat sufficient, I would argue that a better analysis would be one that takes into account these users who have very few attributes defined. Maybe a better number (I'll call x) would be the number of common attributes with commonality of 70% or better (as opposed to 100% in our analysis). For example, if x is 10, this means that for the given circle, there were 10 common attributes among at least 70% of the circle. The percentage (70%) could be adjusted as needed.

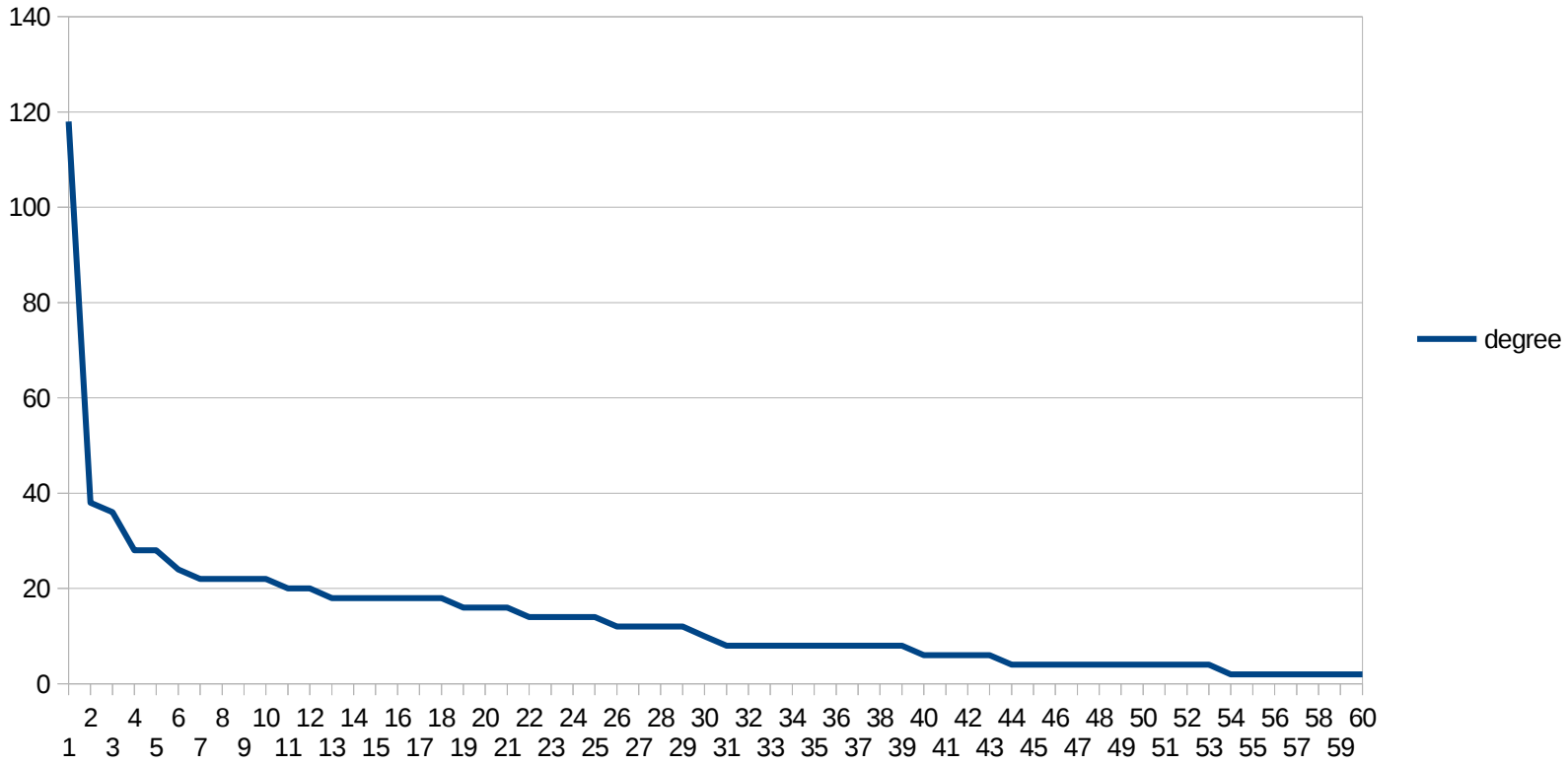
<mmashay2> - **<1>** “The node degree distribution follows so-called power-laws distribution (in which a few nodes have high degrees and many nodes have small degrees).

i) I disagree with this hypothesis. While it is true that there was an outlier of 1 very large degree node in each graph (below), we can see that there are many large degree nodes as well as small degree nodes. The data seems like it might follow more of a normal or binomial distribution than a power distribution.

Ego 1912 - Degree of Nodes Distribution



Ego 3980 - Degree of Nodes Distribution



ii) I believe this data clearly shows that the degrees of nodes do not follow a power law distribution. However, I also believe it wouldn't hurt to carry this out further with additional ego-networks. Two seems like a rather small sample size.