THE ANATOMY OF THE FACEBOOK SOCIAL GRAPH

PAPER REVIEW - PETAR PETROV

AGENDA

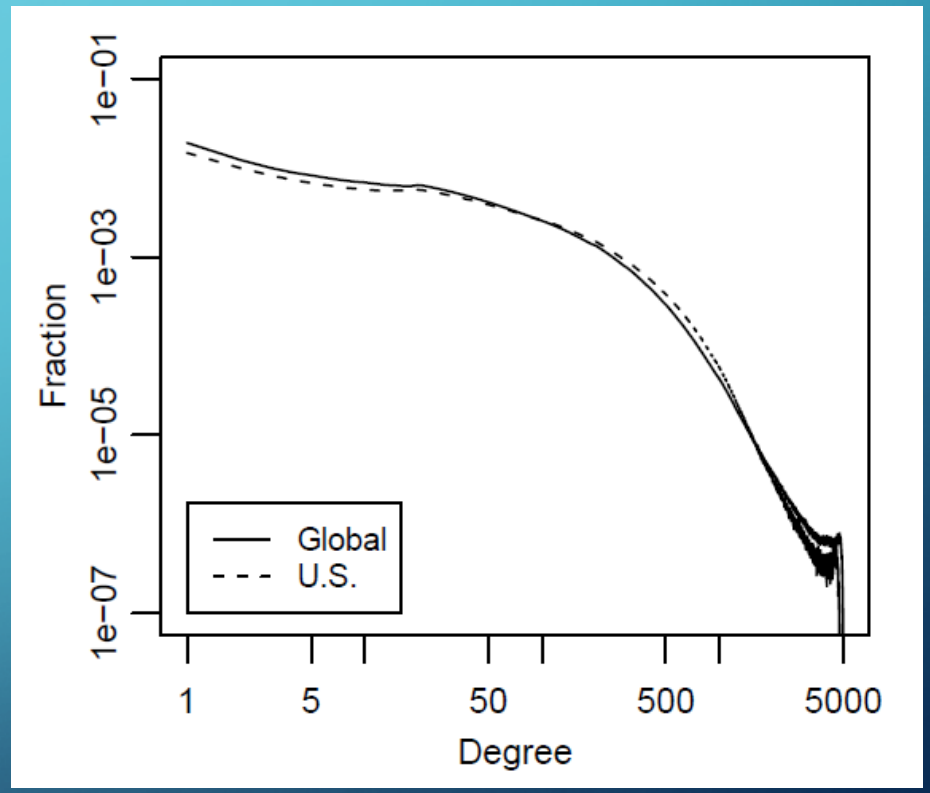
- 1. Method of Analysis – Requirements
- 2. Introduction – Small World Problem
- 3. Results – The Facebook Graph
- 4. Conclusion
- 5. Q&A

METHOD OF ANALYSIS

- N= 721 Million Active Users (May 2011) ~ 10% of World Population
- 68.7 Billion Friendship Edges = 190 Friends on Average
- Subgraph of 149 Million US Facebook Users
  - 15.9 Billion Friendship Edges = 214 Friends on Average
- For analyzing of Network Neighborhoods – 5000 users randomly selected
- Calculations were performed on a Hadoop cluster with 2,250 machines

DEGREE DISTRIBUTION

- Right-Skewed Distribution
- Not a Power Law Distribution
- High Variance
- Clear Cutoff at 5000 Friends
- Most individuals <200 Friends
- Median Nr. Of Friends – 99



SMALL WORLD PROBLEM

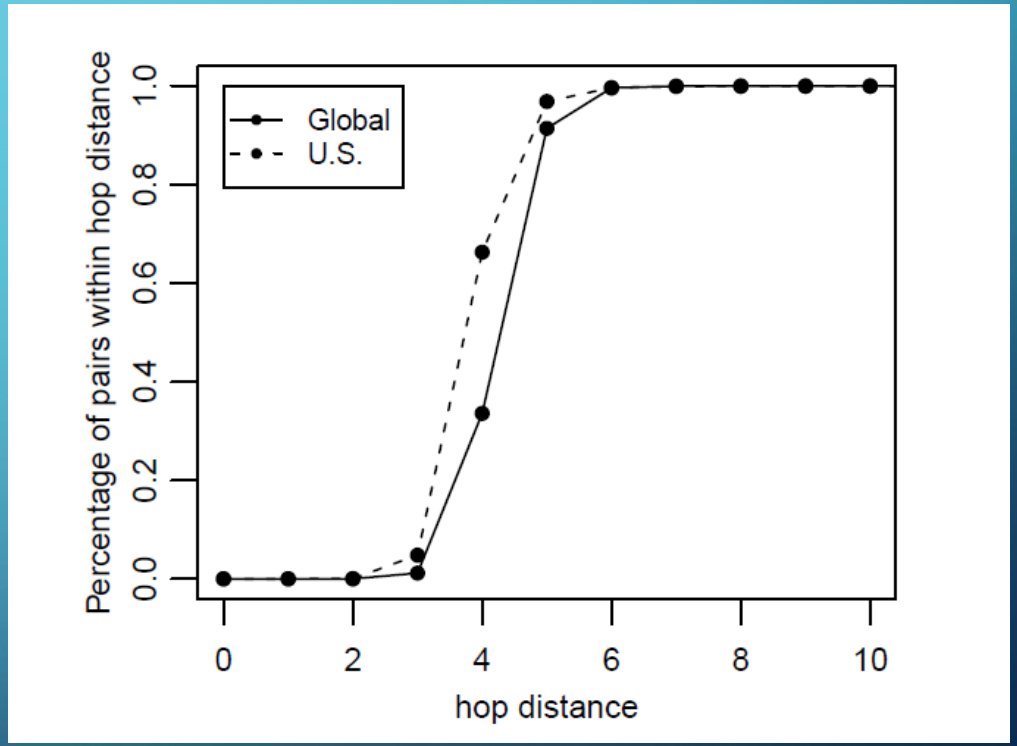
- We live in a “small world” where people are connected by “six degrees of separation”
- The Small World Method – Milgram 1967
  - Biased – Depending on social class
  - 217 chains started, 64 completed = success rate of 29% of started chains
- Facebook – Greatly improving the process

HOP DISTANCE

- Diameter – Less relevant than the Full Neighborhood Function
- Average Distance: Worldwide – 4.7

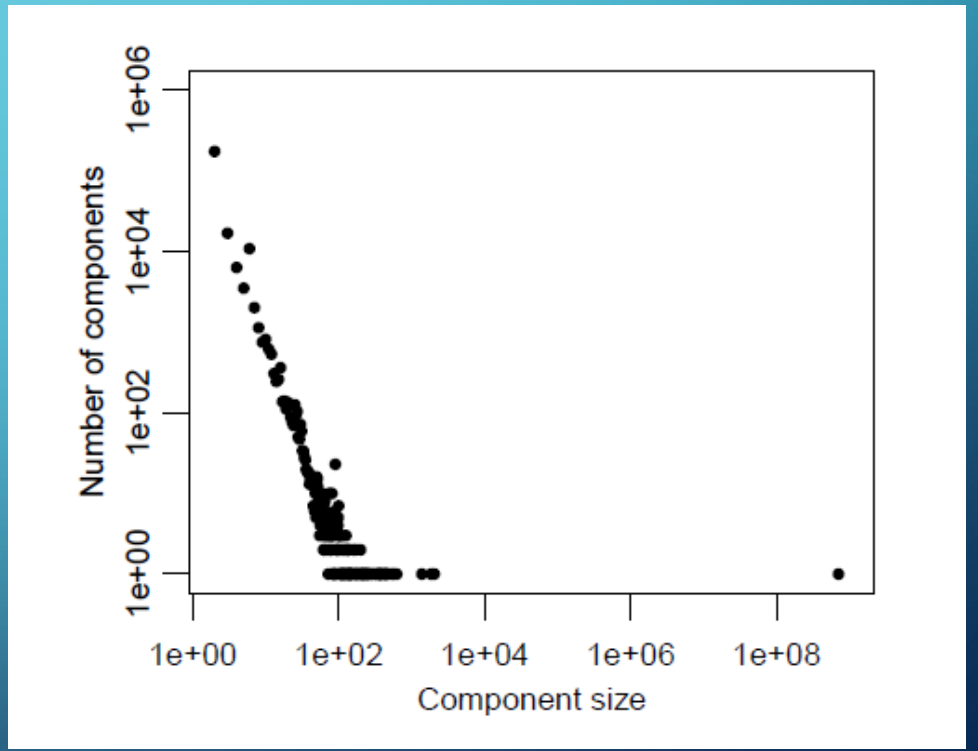
  US – 4.3
- 6 Degrees of Separation: World – 99.6%

  US – 99.7%



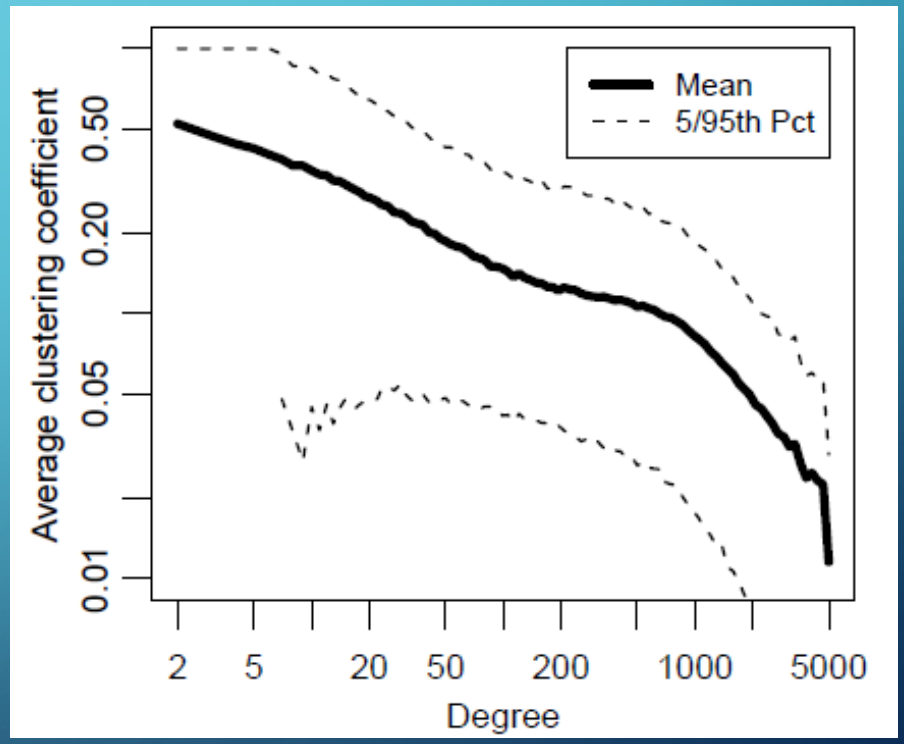
COMPONENT SIZES

- Newman-Zipf (NZ) Algorithm
- Log-log scale
- Many small components (2nd largest has just over 2000 individuals)
- 1 outlier – 99.91% of the Network



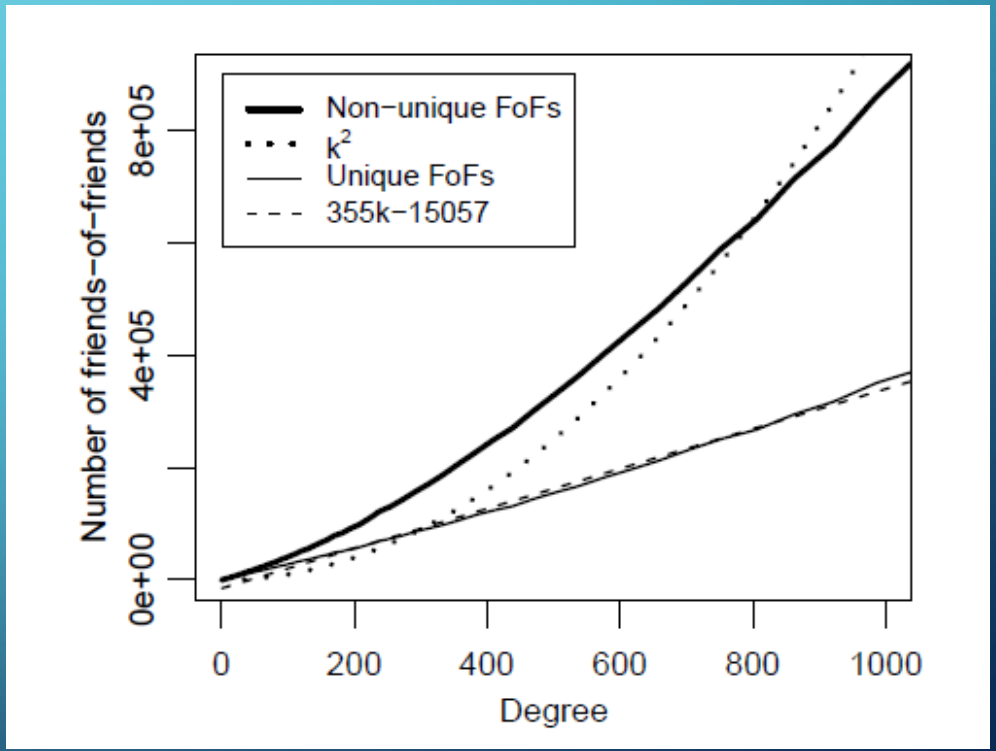
LOCAL CLUSTERING COEFFICIENT

- Very large regardless of the degree
- Decreases Monotonically with degree
- Huge drop at 5000 friends
  - Friending more indiscriminately
  - Less coherently social purposes
- Users with 100 friends - Average LCC = 0.14



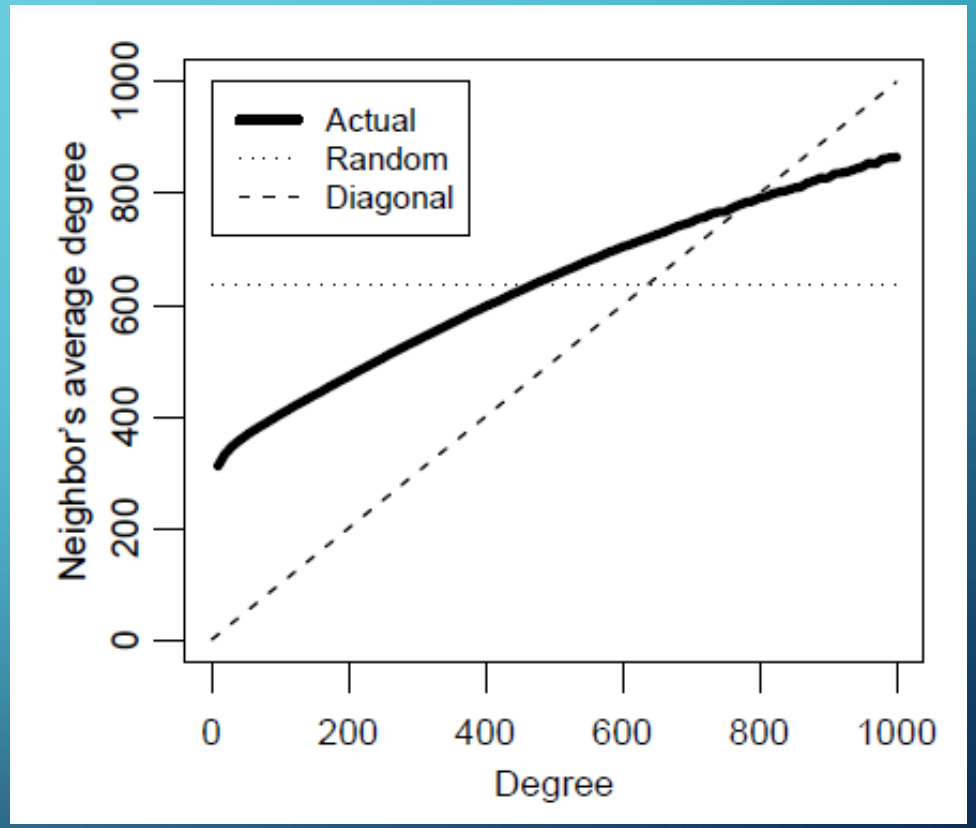
FRIENDS OF FRIENDS

- User with k friends has ~ k^2 **non-unique** friends-of-friends
- Non-Unique – Moderately faster than linear fit
- 355 **unique** friends-of-friends per additional friend
- **Unique –** Very close to linear



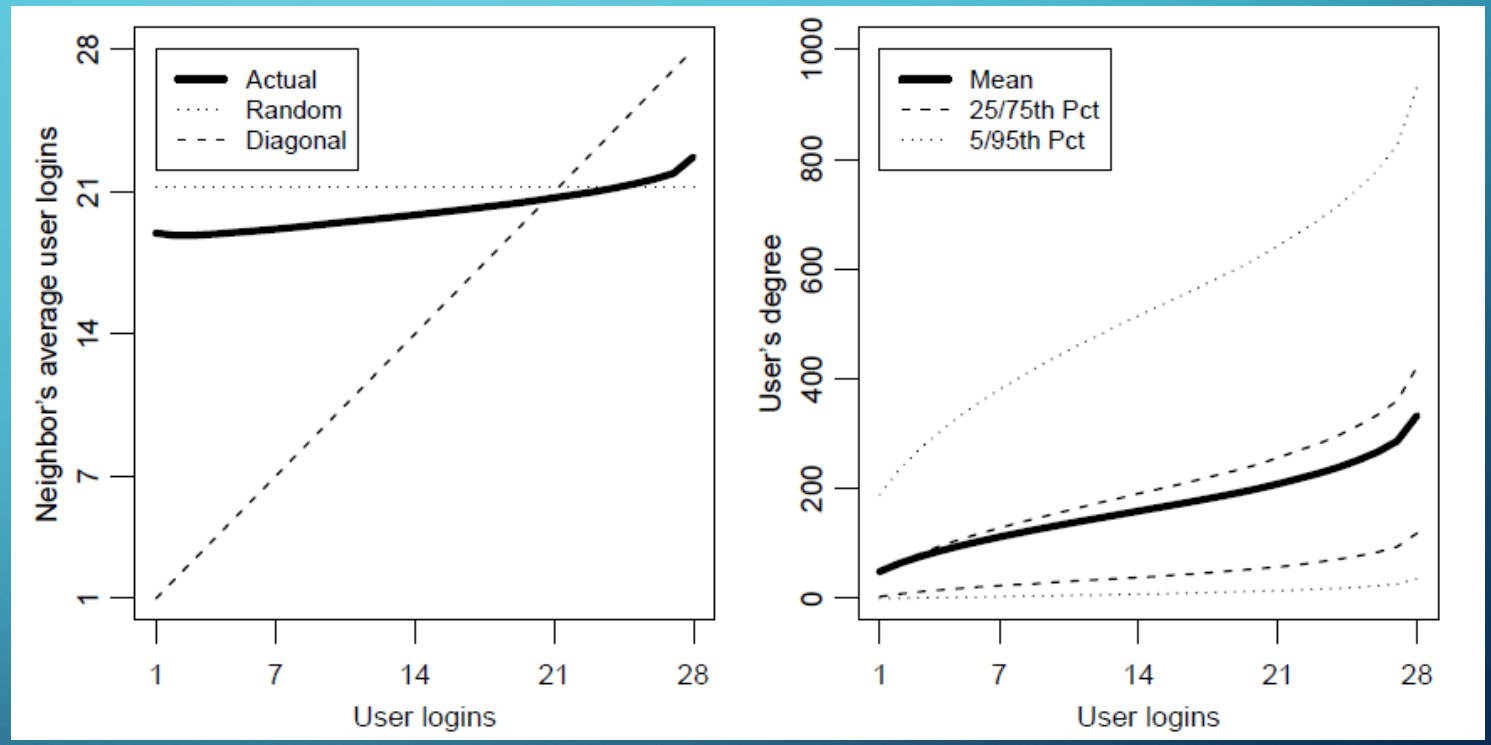
DEGREE CORRELATION

- Pearson coeff. = 0.226 – Positive Assortativity
- Below 700 friends: ‘Your friend has more friends than you.’
- 83.6% of users have less friends than the median friend count of their friends.



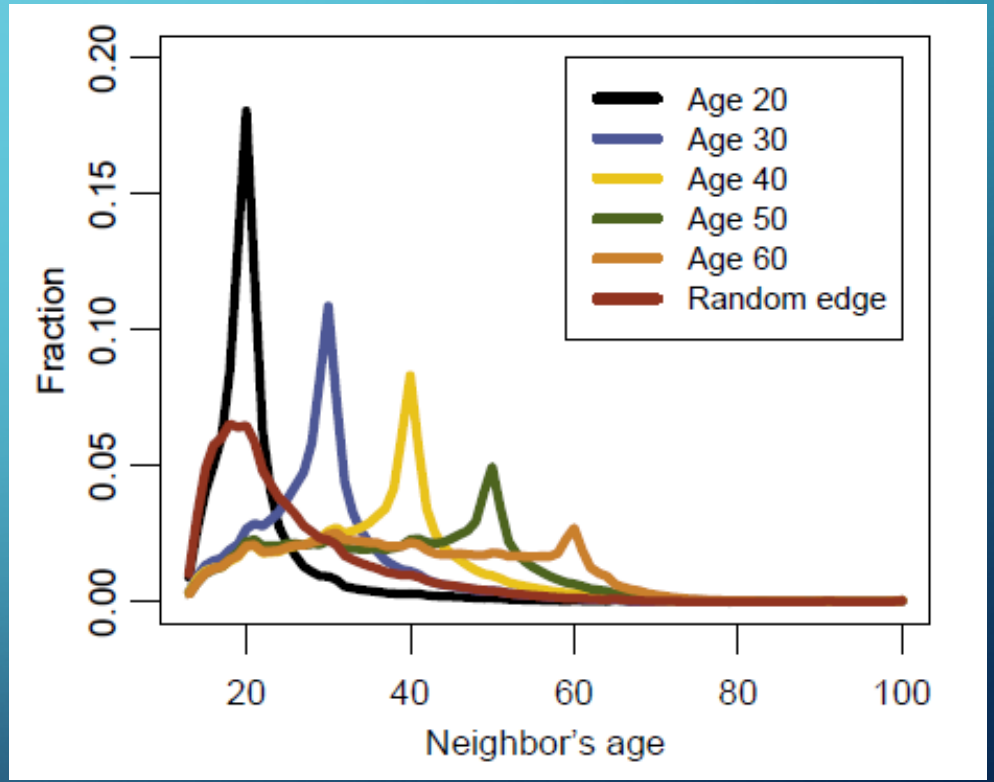
LOGIN CORRELATION

- Up until you log in around 70 % of days in a month, your friends login more
- Users with higher degree login more often



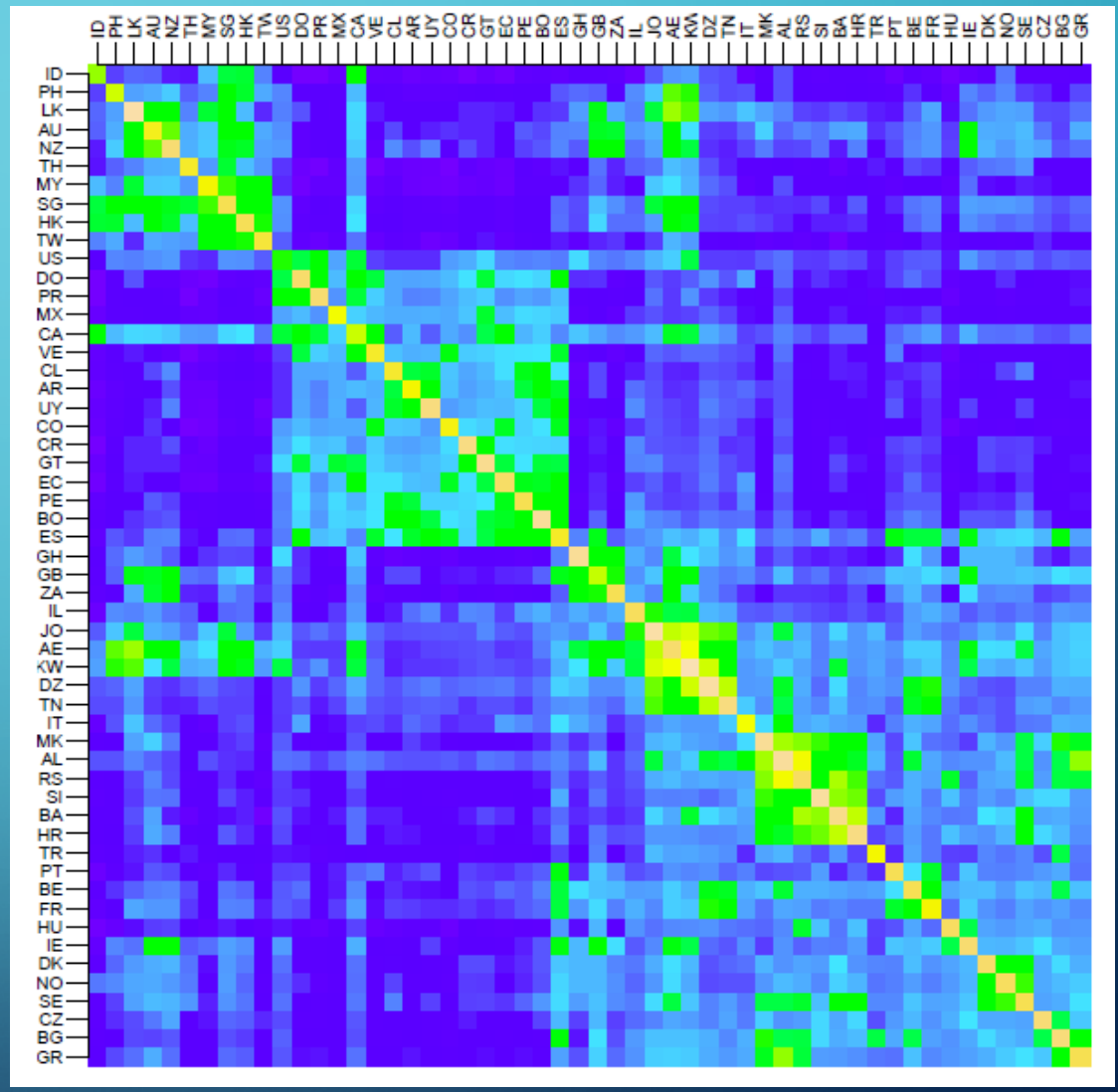
NEIGHBOR'S AGE DISTRIBUTION

- Random neighbor – Most likely to be the same age as you
- Young Individuals – Small age range
- Older Individuals – Wider range
- Variance increases with age



COUNTRY ADJACENCY MATRIX

- 54 countries with active Facebook population over 1 Million
- 84.2% of edges are within countries
- Also possible grouping by:
  - Geographical Distance
  - Historical Ties



SUMMARY

- Confirmed the Six Degrees of Separation Theory
- Almost the entire network is connected
- The Degree of a User is Correlated to the Degree of its Friends
- The Number of Logins is Dependent on Number of Friends
- Friend Preferences for same Age and Country

THANK YOU!