Learning from Networks Project Proposal INQ0091104

Authors

Albert Lalaian - 2106239
Ilyas Issa - 2109673
Thomas G. Ewers - 2134205

November 14, 2024

Contents

1 Introduction

The friendship paradox is a counterintuitive phenomenon observed in social networks. The phenomenon first was mentioned in a 1991 paper, when social networks were offline only, moreover, this paradox might be applied to modern Internet networks as well. According to Johan Ugander et al. in 2012, this phenomenon was confirmed for the social network Facebook, the authors analyzed 721 million users and the result was statistically significant [1]. Also, another study by Nathan O. Hodas, Farshad Kooti and Kristina Lerman in 2014 found that this statement is true for 98% of Twitter users [2]. However, social networks are rapidly developing and these days this statement may already be outdated, and in turn, we want to check whether this statement still holds for different social networks.

The goal of this project is to empirically test the friendship paradox using real social network data and analyze the relationship between node order and various centrality metrics. Examining whether individuals with high node order also have high centrality according to metrics such as betweenness centrality and closeness centrality will provide insight into the role that nodes play in a network and how they affect network connectivity and information flow.

2 Motivation

Understanding this paradox is critical because it affects network dynamics, information dissemination, and viral marketing strategies. We have two main motivations for this project. First, by studying the centrality in the context of the friendship paradox, we can contribute to the development of more sophisticated targeting methods in social networks. Effectively reaching the most influential and most relevant users is not only resource efficient marketing, but also more impactful.

3 Method

3.1 Problem

The primary objective of this project is to make an empirical analysis of the friendship paradox in the context of the Facebook social network to validate the paradox and explore the relationship between node degree and various centrality measures in order to access the network roles.

3.2 Data and Network Metrics

For our analysis, we decided to use the âThe Network Data Repository with Interactive Graph Analytics and Visualizationâ by Ryan A. Rossi and Nesreen K. Ahmed dataset available at the Network Data Repository. This dataset contains a network of Facebook friendships within a user cohort [3]. The dataset consists of 63,731 nodes and 1,269,502 edges, where each node represents a unique Facebook user and each undirected edge represents a friendship tie. Since the network is unweighted, all friendships are treated equally, without additional attributes denoting the strength or frequency of interactions.

3.3 Algorithm

In order to test the Friendship paradox for each node, we will calculate the node degree, obtain the list of neighbouring nodes with their degree value and calculate the average node degree for the connected nodes. By comparing the nodeâs degree with the found average value for all nodes in the network we can test the friendship paradox. For the further analysis of centrality measures roles, for each node, we will also compute the Betweenness Centrality using Brandeâs algorithm, closeness centrality and eigenvector centrality.

3.4 Analysis

In order to investigate how node degree correlates with each centrality measure, and how exactly it affects the node degree value, we will perform the correlation analysis using the Pearson correlation coefficient for each centrality measure. The coefficient will show the strength and vector of the relationship (positive or negative vector). After, the results will be tested on the significance using the p-value in the Z-test. Another experiment is analyzing centrality measure values between Low-Degree and High Degree nodes, to assess if there is any significant difference between the groups.

4 Contribution List

Name	Student ID	Contribution
Ilyas Issa	2109673	40%
Thomas Ewers	2134205	30%
Albert Lalaian	2106239	30%

Table 1: Contribution of Team Members

References

[1] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow, "The anatomy of the facebook social graph," *CoRR*, vol. abs/1111.4503, 2011.

[2] N. O. Hodas, F. Kooti, and K. Lerman, "Friendship paradox redux: Your friends are more interesting than you," *CoRR*, vol. abs/1304.3480, 2013.

[3] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *AAAI*, 2015.