



Project Title:

Friend Recommendation in Social Networks Using Common Neighbors

Group Members:

1. Shaheer Mumtaz [22K-4389]
2. Shaheer Beig [22K-4321]
3. Samroze Ali [22K-4187]

Project Overview:

This project implements a friend recommendation system using the common neighbors metric, a fundamental concept in graph theory. It processes a Facebook ego network dataset, evaluates recommendations using key performance metrics, and visualizes subgraphs for better understanding.

Tools and Technologies Used:

- Python
- NetworkX
- Matplotlib
- Scikit-learn
- Kaggle

Introduction:

Social networks thrive on connections, and recommending friends is a vital feature for enhancing user engagement. Our project focuses on developing a friend recommendation system based on shared connections, implemented using graph theory. Specifically, we use the common neighbors metric, which identifies potential friends by analyzing the number of mutual connections between users.

The dataset used in this project is the Facebook ego network dataset, where nodes represent users, and edges represent friendships. The system processes this dataset to recommend friends for sampled users, tests the model's effectiveness, and visualizes the network structure.

Implementation Details:

1. Dataset Loading:

The dataset is loaded using the `load_graph` function, which parses the edge list file into a graph structure. Each node represents a user, and edges indicate friendships.

2. Friend Recommendation Algorithm:

The core of the project is the `recommend_friends` function. For each node, this function:

- Identifies its direct neighbors.
- Computes the common neighbors score for every other node in the graph.
- Ranks potential friends based on the score and returns the top-k recommendations.

3. Testing and Evaluation:

Using the `test_model` function, the system:

- Randomly samples some nodes from the graph.
- Simulates a ground truth by assigning each node a set of randomly selected friends.
- Compares the algorithm's recommendations with the ground truth to calculate key performance metrics:
 - Accuracy: Measures overall correctness.
 - Precision: Measures how many recommended friends are correct.
 - Recall: Measures how many actual friends were successfully recommended.
 - F1 Score: Balances precision and recall.

Results:

The performance of the friend recommendation system was evaluated using the sampled nodes and the simulated ground truth. Below are the metrics calculated for the system:

- Accuracy: [0.75]
- Precision: [0.75]
- Recall: [1.00]
- F1 Score: [0.86]

These metrics indicate the system's ability to recommend meaningful friends in the context of the dataset.

Visualization:

To provide a deeper understanding of the social network's structure, a subgraph of 20 sampled nodes was visualized. The subgraph highlights the connections between users and the relationships influencing recommendations. This visualization is crucial for analyzing network patterns and validating the algorithm's results.

Conclusion:

Our project demonstrates how graph theory can be applied to solve real-world problems like friend recommendations. The common neighbors metric effectively identifies potential friends based on mutual connections. While the algorithm performs well in this context, further improvements such as incorporating weighted edges or additional similarity measures could enhance its accuracy and applicability to larger datasets.

This project helped us explore key concepts in graph theory, algorithm development, and data analysis, providing a solid foundation for further research and practical applications.