CptS 591: Elements of Network Science
Spring 2024
Semester Project
Professor.Assefaw Gebremedhin

Sharath Kumar Karnati

March 2023

Project Title: Community Detection(Network Analysis) using NetworkX(python)

Abstract:

This project's goal is to investigate different facets of network analysis using Python's NetworkX library. Applications for network analysis can be found in many different fields, including infrastructure, biological, and social networks. A wide range of subjects will be covered in this project, such as community detection algorithms, link prediction, graph statistics, and visualization techniques.

Goals and Objectives:

- To understand different types of graphs and their properties.
- To learn how to visualize networks using different layout algorithms.
- To compute and interpret various graph statistics to gain insights into network structure.
- To implement and evaluate common link prediction techniques.
- To explore and compare community detection algorithms for partitioning networks

The ultimate objective of this paper is to meticulously compare and analyze diverse types of networks utilizing algorithms specifically tailored for comparison purposes.

Methodologies:

The project aims to accomplish its goals and objectives through the implementation of the following methodologies:

NetworkX Graph Types:
Firstly I will Implement undirected, directed, signed, weighted, multigraph, bipartite, and projected graphs using NetworkX to see how they works and Visualize each type of graph to understand its structure and properties.
Performing Graph Statistics:
Calculate and interpret various graph statistics including clustering coefficients, distance measures, connectivity, robustness, and centrality measures and then, I will utilize the Karate Network and other example networks that we used for assignments for demonstration.
Graph Link Prediction:
Implement and evaluate common link prediction features such as common neighbors, Jaccard coefficient, and etc to visualize potential connections between nodes using these techniques.
Algorithms Implemented:
Explore community detection algorithms including Girvan-Newman and Louvain, and Leiden algorithms and then will apply these algorithms to real-world networks and visualize the community partitions.
Commarison of the results that I will get from using above algorithms:
Compare the performance algorithms for community detection using the Stanford Network Analysis Project dataset that is provided to us for reference, and Highlight the advantages among the algorithms used and provide visualizations of community partitions.

Data Sets :

I will use publicly available datasets to evaluate the performance of the algorithms. The datasets provided by my professor in class will be my main source of data. Additionally, I can access these datasets online from sites like Kaggle, the Network Repository, and the Stanford Large Network Dataset Collection.
I will download the data files from the specified internet sources to retrieve the datasets. Afterwards, I will pre-process the data to eliminate any irrelevant information and transform it into a format compatible with my simulation framework. For example, I will convert the data into a graph structure, where nodes represent entities and edges signify connections between them.

Background work and Literature Review :

The proposed project on network analysis utilizing NetworkX draws upon fundamental concepts from graph theory and leverages the capabilities of the NetworkX library to explore and analyze complex networks. With a focus on graph types, visualization techniques, statistical metrics, link prediction methods, and community detection algorithms, this project aims to provide a comprehensive understanding of network structures and dynamics across diverse domains. By integrating theoretical principles with practical implementations, the project seeks to offer valuable insights into the complexities of real-world networks and facilitate informed decision-making processes.
The present work, along with several of the methods we employ in our research, draws inspiration from a number of influential historical works. Among them

are:
Network Community Detection: A Review and Visual Survey Bisma S.Khan1, Muaz A.Niazi,1
A. Clauset and C. Moore. How Do Networks Become Navigable? arXiv:cond-mat/0309415v2, 2003.
https://medium.com/@ds225229101/detecting-communities-in-large-networks-using-networkx-packages-1e7e9d73a5cd These provided a solid foundation on the work I am gonna present and they helped a lot to understand those concepts in depth.

Timeline :

Week 1: Literature Review and Data Collection: I intend to conduct a thorough literature review of the relevant research in network navigation and algorithmic analysis. Additionally, I will collect publicly available datasets for social and transportation networks to use in the subsequent stages of the project.

Week 2: Simulation Framework Development: Following the literature review and dataset collection, I will focus on developing a simulation framework capable of generating various types of networks with diverse topologies. Subsequently, I will proceed with implementing the algorithms and performance metrics within this framework.

Week 3: Algorithm Evaluation: My next step will involve evaluating the performance of the algorithms on the simulated networks. I will analyze their efficiency, accuracy, and scalability to determine their effectiveness in different scenarios and network structures.

Week 4: Analysis and Report Writing: In the final week, I will analyze the results obtained from the algorithm evaluations and draw conclusions based on the findings. I will then document the entire project, including the methodology, results, and conclusions, in a comprehensive report.

Conclusion :

This project proposal outlines a structured plan for exploring network analysis using the NetworkX library. By completing this project, I aim to develop a deeper understanding of network properties, visualization techniques, statistical measures, link prediction methods, and community detection algorithms. The acquired knowledge and skills will be valuable for analyzing and understanding complex networks in various domains.