CSCE 590 Individual Project Proposal

Sai Siva Rohith Tirumalasetti
ULID: C00576124

Project Title:

Extraction and Analysis of an Online Network

Project Description:

Analysis of DBLP Computer Science Bibliography using graphs and evaluation of statistical properties of the same repository.

The DBLP Computer Science Bibliography is a comprehensive database of bibliographic information for major computer science journals and conferences. This project aims to analyze DBLP using graph theory, evaluating statistical properties to better understand the structure and relationships within the database.

Project Goals and Objectives:

The DBLP repository provides bibliographic data on computer science publications. The primary goals of this project are to:

1. **Develop a Command Line Interface (CLI)** and **Graphical User Interface (GUI)** to allow users to:
    - Give author names as input.
    - Generate a graph-based network that represents collaborations between authors based on shared journals and proceedings.
    - Treat each author as a node and at least one collaboration between two authors as an edge in the graph.
2. **Evaluate statistical properties** of the DBLP repository as a whole to gain insights into its structure and relationships (e.g., centrality measures, clustering coefficients, etc.).

Anticipated Project Deliverables:

1. **Source Code**: Complete source code for backend, CLI and GUI.
2. **Deployable Bundle**: A deployable solution that works with both the Command Line Interface (CLI) and the Graphical User Interface (GUI).

3. **Documentation**: A comprehensive document detailing the background, design, functionality, and usage of the project.

Project Weekly Plan:

- **Week 3**: Explore the DBLP repository and gather some knowledge about its contents.
- **Week 4**: Identify required software/tools (programming languages, frameworks etc.) and develop a prototype.
- **Week 5**: Define and document a set of functional and non-functional requirements.
- **Week 6**: Define use cases and the technical design for the system.
- **Week 7**: Setting up a development environment.
- **Week 8**: Develop a set of RESTful APIs to search / filter data.
- **Week 9**: Develop a set of RESTful APIs to generate graphs.
- **Week 10:** Develop a Command Line Interface (CLI) to interact with the graph service and display network details.
- **Week 11**: Develop a Web UI for interacting with the graph data.
- **Week 12**: Integrate of the CLI, UI, and graph service.
- **Week 13**: Identify key statistical properties to be evaluated from the DBLP data.
- **Week 14**: Develop code to evaluate Statistical Properties.
- **Week 15**: Work on completing the documentation**.**
- **Week 16:** Demonstrate the final product.