Network Analysis Report

February 26, 2025

Contents

Introduction

This report presents an analysis of a network dataset, focusing on key structural characteristics, group collaboration dynamics, and interdisciplinarity.
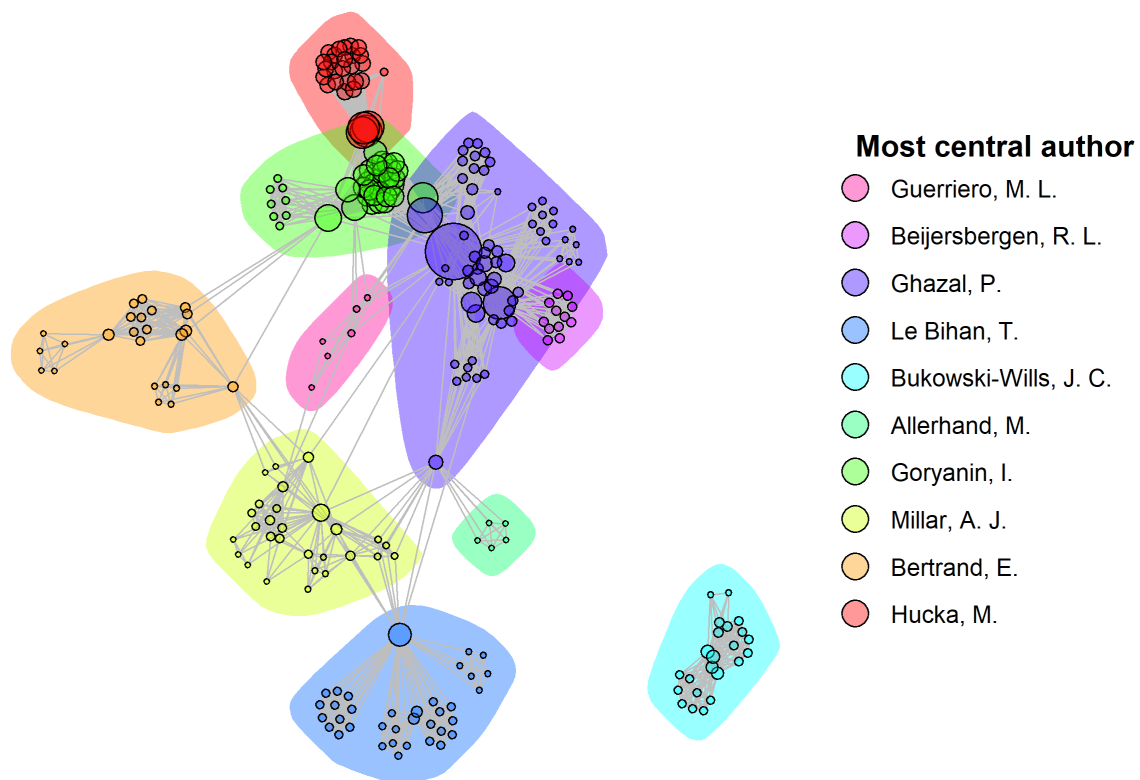
Figure 1: Visualisation of the Author network

Data filtering

All papers in the dataset had less than 50 authors and were included in the network analysis.

42 authors were removed from the network because they appeared in less than 5 papers. 268 authors were included in the network analysis.

Network Type

- **Undirected**: The network is undirected, meaning all connections are mutual. This is common in co-authorship or collaboration networks. In an undirected network, connections do not have a direction (e.g., if A is connected to B, then B might also be connected to A by definition).
- **Weighted**: The network is weighted, meaning that connections between the nodes (authors) vary in strength or intensity. This result indicates that the network has weighted edges, meaning that the relationships between nodes are assigned numerical values. These weights could represent the

strength or frequency of interactions, such as the number of co-authored papers.

Cohesion Metrics

- **Total Dyads**: Dyads, representing all possible pairs of nodes (authors connection), were calculated as (n*n-1) = 35778
- **Actual Edges**: The network contains 2466 actual connections (edges), showing the level of interconnectedness among nodes (authors). This is the count of actual connections (edges) present in the network, indicating how interconnected the nodes are
- **Number of Nodes**: The network consists of 268 nodes (authors), representing the total entities analysed.

Density and Transitivity

Density

Density measures the proportion of realised connections compared to all possible connections. It is calculated as:

```
Density = Number of Edges/(Number of Possible Edges) = Actual Edges/[nx(n-1)/2]
```

- **Value**: 6.89%
- **Interpretation**: Values of Density range from 0 to 1. High Density: Close to 1, meaning most nodes are connected. This indicates a highly cohesive network. Low Density: Close to 0, meaning most nodes are not connected. Sparse networks are common in real-world systems.

Transitivity

Transitivity, or the clustering coefficient, measures the tendency of nodes to form triangles (e.g., if A → B and B → C, then A → C). (In other words, if two nodes are connected to a common third node, they are also likely to be connected to each other). It is calculated as:

```
Transitivity = Number of Triangles/Number of Connected Triplets
```

- **Value**: 74%
- **Interpretation**: Values range from 0 to 1. High Transitivity: Close to 1, indicating a high level of clustering. Nodes tend to form tight-knit groups. Low Transitivity: Close to 0, indicating a lack of clustering. A higher Transitivity value suggests that network tend to collaborate in groups, forming cohesive communities.

Node Centrality and Centralization

Degree Centrality

Degree centrality measures the number of direct connections each node has in the network. Nodes with higher degree centrality are likely to be influential or well-connected individuals/entities.

- **The least connected node has (Min)**: 5
- **The most connected node has (Max)**: 117
- **On average, nodes have about (Mean)**: 18
- **Half of the nodes have more connection than (Median)**: 12

Betweenness Centrality

Betweenness centrality quantifies how often a node acts as a bridge in the shortest paths between other nodes. Nodes with high betweenness are critical connectors in the network. Removing them could significantly disrupt the flow of information.

- **Min**: 0
- **The node with the highest value is a critical bridge in the network, controlling the flow of information. with Max**: 10414
- **Mean**: 232
- **Median**: 4

Closeness Centrality

Closeness centrality measures how quickly a node can access other nodes in the network. Nodes with high closeness centrality are in advantageous positions to disseminate information quickly.

- **The least central node has a (Min) avarage path to reach other nodes**: 0.00069
- **Max**: 0.03704
- **Mean**: 0.00392
- **Median**: 0.00142

Network diameter

The network diameter is the maximum distance (in terms of edges or steps) required to connect any two nodes in the network through the shortest possible path.

- In our network: **the diameter is 9**, it indicates that the farthest two nodes in your co-authorship network can be linked by the shortest path of 9 steps.
- In practical terms, if two authors in the network are the most distantly connected, they are separated by 9 intermediaries (co-authors).
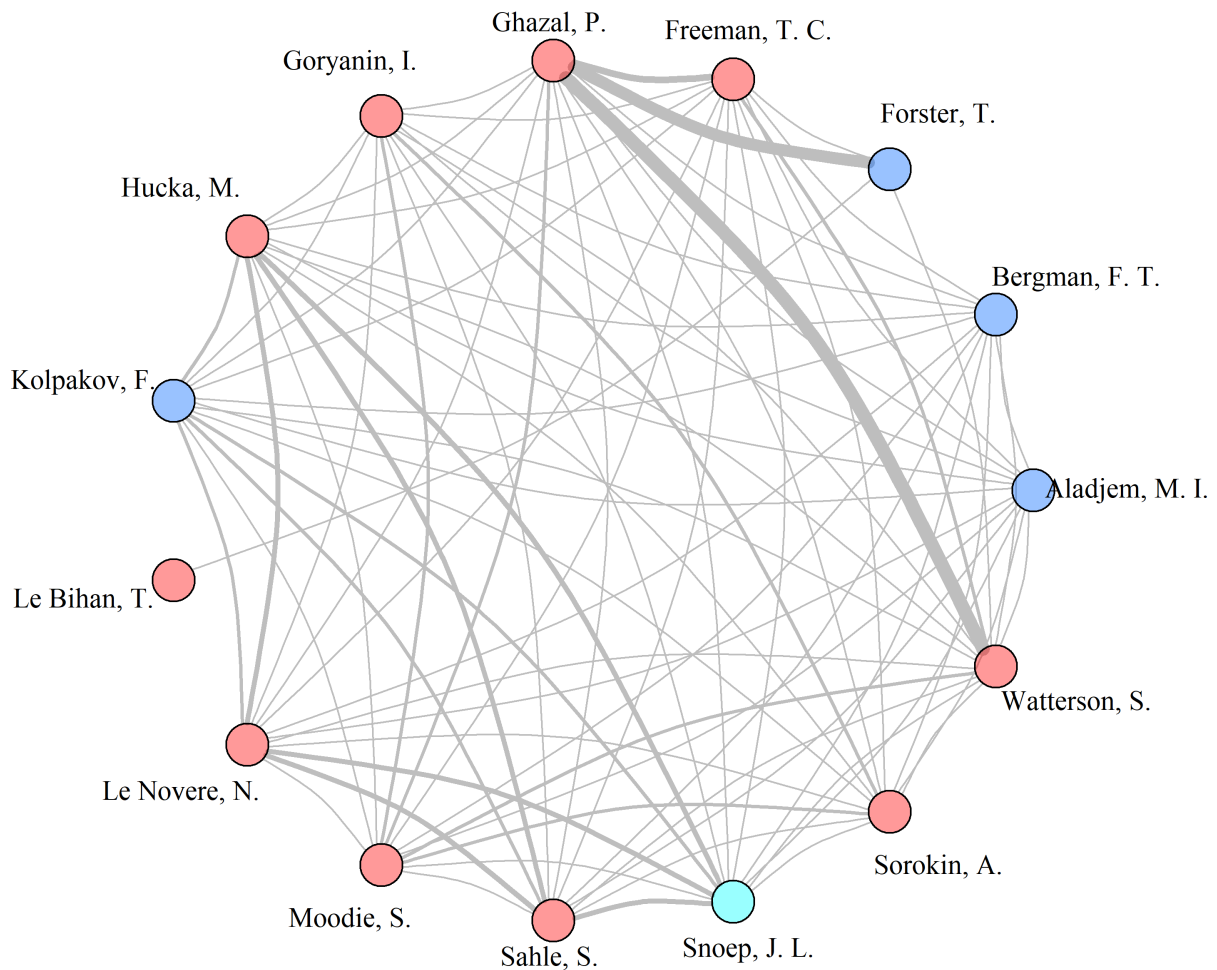

Figure 2: Direct connections between the 15 most central authors