LinkedIn Network Analysis Report

Ravi Rajput

B.Tech in Computer Science

1. Introduction

This report analyzes the LinkedIn network of **28,492 students**, highlighting influential nodes, connectivity gaps, and strategic opportunities for improved engagement.

2. Data Cleaning

Unstructured data was cleaned and standardized. Noisy filenames were corrected, duplicates removed, and irrelevant nodes filtered to ensure integrity.

3. Key Findings

3.1 Degree Centrality

Top connectors:

- **Rohit Malviya:** 4248 connections

- **Ravi Rajput:** 4060 connections

- **Manoj Dewda:** 3960 connections

These individuals are central to the network's communication and opportunities.

3.2 Random Walk Insights

Walk paths between top 20 students revealed:

- **Avg. Raw Length:** 145.4 steps   (Pruned: 11.1)

- **Max Pruned Path:** 56 steps, **Min:** 2 steps

- **92.4% paths shortened after pruning**

- **22.6% walks included loops**

Interpretation: While long paths exist, meaningful connections form quickly when redundant nodes are removed. A few key nodes drive reachability.

3.3 Graph Partitioning

- **Sitare Subgraph:** 126 nodes – well-connected

- **Non-Sitare Subgraph:** 28,366 nodes – fragmented

Insight: Sitare is cohesive and effective; Non-Sitare needs connection strategies to boost network impact.

4. Dominating Set

4.1 Sitare vs Non-Sitare Comparison

Metric	Sitare	Non-Sitare
Total Students	126	28,366
Dominating Set Size	9	28,366
Coverage (%)	6.35%	100.00%
Connected Graph?	Yes	No (isolated nodes)

4.2 Sitare Core Members

Bhagwan Singh Rawat, Chiranjeet Biswas, Karanpal Singh Ranawat, Kuldeep Saraswat, Maneesh Sakhwar, Manoj Kharkar, Rahul Kumar Verma, Uppara Sai Maithreyi, Gaurav Khainwar

Insight: Just 9 students cover the entire Sitare subgraph—showing strong internal collaboration. The Non-Sitare group lacks cohesion and requires structural improvements.

5. Strategic Insights

- **Graph Density:** 0.0002 – sparse yet functional.

- **Core Influencers:** Central nodes (Rohit Malviya, Ravi Rajput, Manoj Dewda) drive reach and influence.

- **Efficient Subgraph:** Sitare's 6.35% dominates full reach—model of efficiency.

- **Weak Zone:** Non-Sitare is fragmented—opportunity to foster stronger links.

- **Fast Info Flow:** Median pruned paths 10 steps. Great potential for idea exchange.

6. Conclusion

This LinkedIn student network reveals that a few well-connected individuals maintain strong internal communication, especially within Sitare. The rest of the network is under-connected but holds massive untapped potential. Strategic efforts to improve non-Sitare engagement and empower core influencers will unlock widespread collaboration, opportunity, and impact.

Thank you