Mapping the Ramayana: A Network Science Approach to Character Interactions

Dynamical Processes in Complex Networks – Project P18

Authors:

Yashwanth Duggi (Roll No: 2022102051)

T S Mitra Somayaji (Roll No: 2022102061)

Chetan Krishna Vellanki (Roll No: 2022102047)

Sai Praneeth Bommana (Roll No: 2022101097)

IIIT Hyderabad

October 2025

Abstract

This project explores the *Ramayana* through the lens of network science, modeling characters and their interactions as a dynamic co-occurrence graph. By examining how relationships evolve across the six books of the epic, we aim to reveal structural and emotional shifts that shape its narrative. Sentiment and topic modeling further highlight how themes of love, exile, and devotion progress over time. Inspired by prior network analyses of the *Mahabharata*, this study extends such quantitative approaches to the *Ramayana*, offering fresh insights into one of India's foundational texts.

1. Introduction

The idea that stories can be studied as networks—of characters, themes, or motifs—has gained increasing traction in computational humanities. Network analysis enables quantitative exploration of narrative structures, revealing how relationships and themes evolve across time and context.

In ***A Quantitative Social Network Analysis of the Character Relationships in the Mahabharata*** by Eren Gultepe and Vivek Mathangi (2023), the authors constructed character co-occurrence graphs and analyzed network metrics such as centrality, clustering, and modularity to uncover social hierarchies within the epic. Their work suggested future extensions to other major Indian epics, notably the *Ramayana*, where similar techniques could illuminate narrative and cultural dynamics.

Despite a few exploratory attempts—such as Sumana Sridharan's ***Of Epics, Networks and Bots***, which created a preliminary Ramayana network using text-mined co-occurrences—there has been no rigorous academic study quantifying relationship evolution in the *Ramayana* through network and sentiment analysis.

This project aims to fill that gap, combining co-occurrence modeling, temporal analysis, and sentiment modeling to map the epic's relational and emotional trajectories.

Data Sources

We use publicly available digital versions of the *Valmiki Ramayana* to construct the dataset. These include both English translations and structured JSON or text datasets for computational processing.

- Kaggle: Ramayan Text Data
- GitHub: Valmiki Ramayan Dataset
- Kaggle: Valmiki Ramayana Dataset
- GitHub: Ramayana Book Repository

References

1. Sridharan, S. *Of Epics, Networks and Bots.* Medium Article. Repo: GitHub – su-mana-s/Ramayanam.

2. Gultepe, E., & Mathangi, V. (2023). *A Quantitative Social Network Analysis of the Character Relationships in the Mahabharata. Heritage*, 6(11), 366. Link.

3. Paul, A., Mishra, R., Padia, N., & Das, D. (2024). *Event Identification and Chronology Model of Ramayana.* IEEE Xplore. Link.

4. Beveridge, A., & Shan, J. (2016). *Network of Thrones. Math Horizons.*

5. Newman, M. E. J. (2018). *Networks: An Introduction.* Oxford University Press.

2. Objectives and Motivation

2.1 Objectives

- **Model Foundations:** Construct a co-occurrence network of characters from the *Ramayana*, using window-based proximity or canto-level co-appearance to form edges. The goal is to translate narrative co-presence into a measurable interaction graph.

- **Temporal Evolution:** Analyze how network properties—such as degree centrality, clustering coefficient, and modularity—change across the six books or grouped cantos. This temporal slicing allows us to track emerging alliances and rivalries through different narrative arcs.

- **Sentiment and Theme Modeling:** Perform sentiment analysis using VADER for polarity scoring and DistilBERT for context-aware sentiment classification. Thematic shifts will be captured via topic modeling to understand emotional and conceptual evolution.

- **Network Visualization:** Create static and dynamic visualizations using NetworkX and Gephi to show how relationships strengthen, weaken, or dissolve over the course of the epic.

- **Cross-Cultural Comparison:** Lay the groundwork for comparative analysis with prior network studies of the *Mahabharata* and other large narratives to identify universal storytelling structures.

2.2 Motivation

The *Ramayana* remains one of the most influential epics in world literature, yet it has rarely been explored through computational and quantitative lenses. Understanding its narrative through network analysis provides a novel way to map social hierarchies, emotional arcs, and thematic coherence. By bridging literary theory and complex network analysis, we aim to uncover deeper structural and cultural patterns that traditional readings may overlook.

3. Materials and Methods

Data Preprocessing

- Tokenize the epic into individual cantos or chapters.
- Perform Named Entity Recognition (NER) to extract and normalize character names and aliases. (Or simply extract the principal names from already existing index of Ramayana)
- Segment text into temporal slices—either six books or fixed sets of cantos—to capture progressive evolution.

Network Construction

- **Co-occurrence Graph:** Nodes represent characters; edges denote co-appearance within the same canto or textual window. Edge weights correspond to co-occurrence frequency.
- **Interaction Graph:** Derived from verbs and dialogue mentions to infer directed edges (who interacts with whom).
- **Thematic Graph:** Words or topics as nodes, connected by semantic proximity to capture conceptual relationships.

Each of these networks can reveal distinct narrative insights:

- *Co-occurrence Graphs* highlight social proximity and alliances.
- *Interaction Graphs* expose directionality—leaders, followers, antagonists.
- *Thematic Graphs* visualize transitions of major motifs such as duty, exile, or devotion.

Temporal and Sentiment Analysis

- Compute temporal metrics (degree, betweenness, community structure) per book or canto batch.
- Apply **VADER** and **DistilBERT** to measure sentiment polarity and contextual emotional shifts across narrative time.
- Map sentiment trends onto network visualizations for combined emotional–structural interpretation.

Visualization Tools

Interactive and static visualizations will be produced using:

- **NetworkX** for Python-based metric computation and temporal graph generation.
- **Gephi** for large-scale visualization and modularity-based clustering.
- **Plotly/D3.js** for interactive online displays of network evolution.

4. Initial Results

Dataset Used:
The analysis was performed using the publicly available Ramayana text dataset from Kaggle: `https://www.kaggle.com/datasets/ajayverma23/ramayan-text-data`.

4.1 Canto Distribution Across Books

The text of the Ramayana was divided into six books with the following number of cantos:

- Book I (1): 75 cantos
- Book II (2): 119 cantos
- Book III (3): 76 cantos
- Book IV (4): 67 cantos
- Book V (5): 56 cantos
- Book VI (6): 101 cantos

4.2 Character Mentions and Co-occurrences

The top 10 most frequently mentioned characters in the first 20 cantos are:

Character	Mentions
Ráma	1810
Lakshmaṇ	553
Sítá	462
Raghu	423
Bharat	365
Sugríva	325
Indra	322
Rávaṇ	300
Lanká	178
Báli	161

The top 10 co-occurring character pairs are:

Character Pair	Co-occurrences
Lakshmaṇ - Ráma	422
Raghu - Ráma	408
Ráma - Sítá	366
Indra - Ráma	294
Lakshmaṇ - Raghu	254
Lakshmaṇ - Sítá	252
Ráma - Rávaṇ	240
Raghu - Sítá	206
Indra - Raghu	196
Bharat - Ráma	194

Table 1: Summary statistics of the full Ramayana co-occurrence network

Metric	**Value**
Total nodes	359
Total edges	7,144
Average degree	39.8
Network density	0.111
Diameter	4
Average clustering coefficient	0.021
Degree assortativity	-0.414
Top 3 nodes by weighted degree	Ráma (4,979), Raghu (3,192), Lakshmaṇ (3,016)

4.3 Visualizations

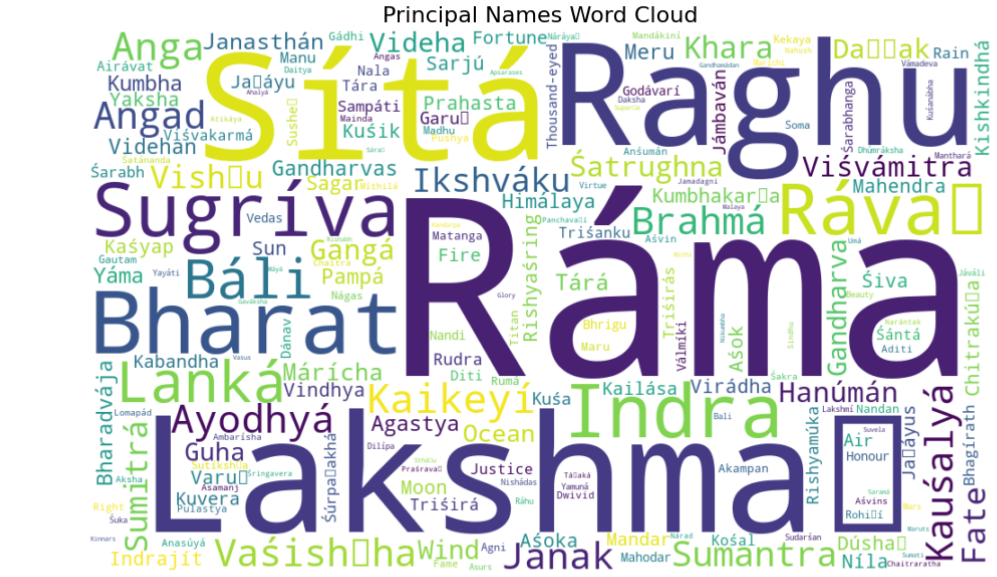

Figure 1: Word cloud of the principal names in the complete text.

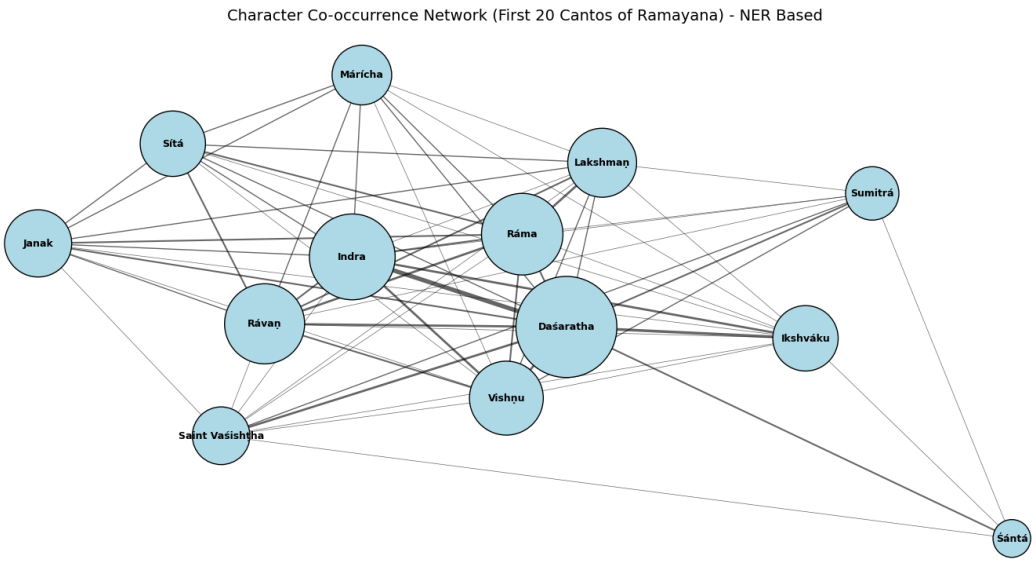

Figure 2: Co-occurrence network of characters in the first 20 cantos. Node size indicates total connections and edge width indicates co-occurrence frequency.

5. Conclusion

The initial analysis of the first 20 cantos of the Ramayana provides the following insights:

1. **Character Centrality**
   The network reveals that *Daśaratha*, *Ráma*, and *Indra* are the most central characters in these cantos, having the largest node sizes. This indicates their prominence in the early narrative, guiding the story's key events and interactions.

2. **Strong Character Relationships**
   The strongest co-occurrences are observed between *Ráma* and *Lakshmaṇ*, *Ráma* and *Sítá*, and *Ráma* and *Daśaratha*, highlighting familial and divine connections that drive the plot.

3. **Network Density and Early Narrative Focus**
   Although only 20 cantos are analyzed, the network is already densely connected, showing that major characters frequently appear together. This emphasizes the narrative's early focus on establishing key relationships and character interactions.

4. **Visualization for Further Analysis**
   The NER-based co-occurrence network provides a clear visualization of character prominence and interactions, enabling future analysis of narrative evolution, temporal patterns, and thematic clusters as more cantos are incorporated.

Key Insight:
Even in the first 20 cantos, the Ramayana shows a structured network where central characters like *Daśaratha* and *Ráma* form the core of the story, while other figures such as *Sítá*, *Lakshmaṇ*, and *Indra* act as strongly connected nodes supporting major plotlines. In short: *"Early character interactions define the narrative backbone before the epic expands into broader story arcs."*

Our further work will explore temporal analysis to track character prominence over the course of the epic, semantic understanding of cantos and events, construction of thematic and motif-based graphs, and more detailed network analyses as additional cantos are incorporated.