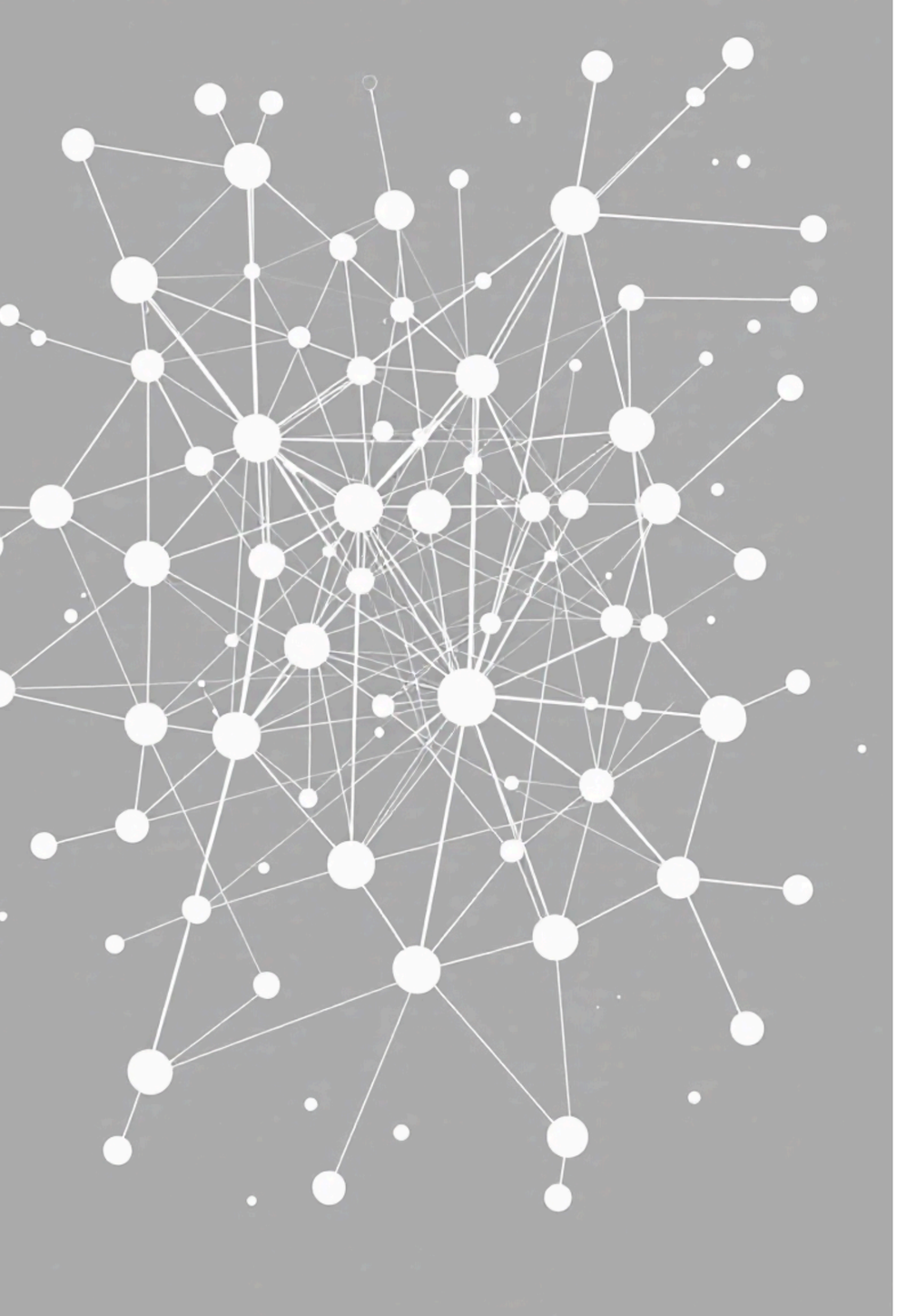
Temporal Stability and Equilibrium Dynamics in Online Social Interactions

Case Study of a Subreddit Activity Network

Complex Data Analysis 2025/26
Prof. André Falcão

Students:
M. João Vicente (44489), Pedro Fanica (54346) and Quentin Weiss (66292)



Objectives

- Temporal Network Analysis (2015-2018)
- Assess Stability & Cohesion
- Identify Structural Patterns

Data Lifecycle

01

Data collection

Downloaded dataset from Cornell's ConvoKit: user-post interaction in the r/Documentaries subreddit.

02

Data Cleaning

Parsed timestamps, excluded deleted users and invalid utterances. Applied sentiment analysis to each comment.

03

Network Construction

Monthly snapshots (04-2015 to 11-2018). Directed and undirected weighted graphs from user reply patterns.

04

Network Analysis

Computed Network metrics and applied Louvain algorithm to identify communities within each snapshot.

Network Analysis Results

Network Analysis

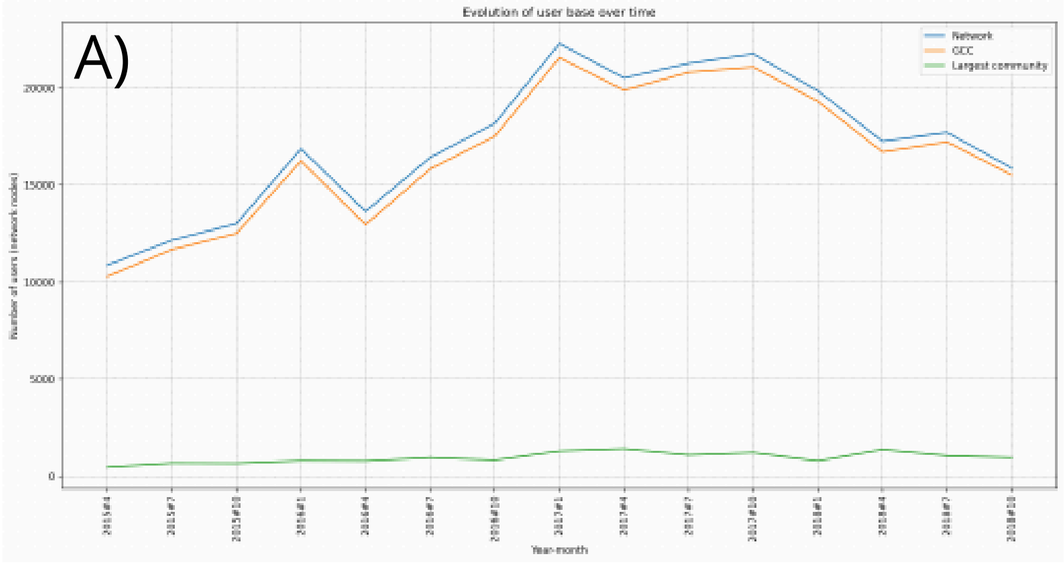
Network metrics **over time:**

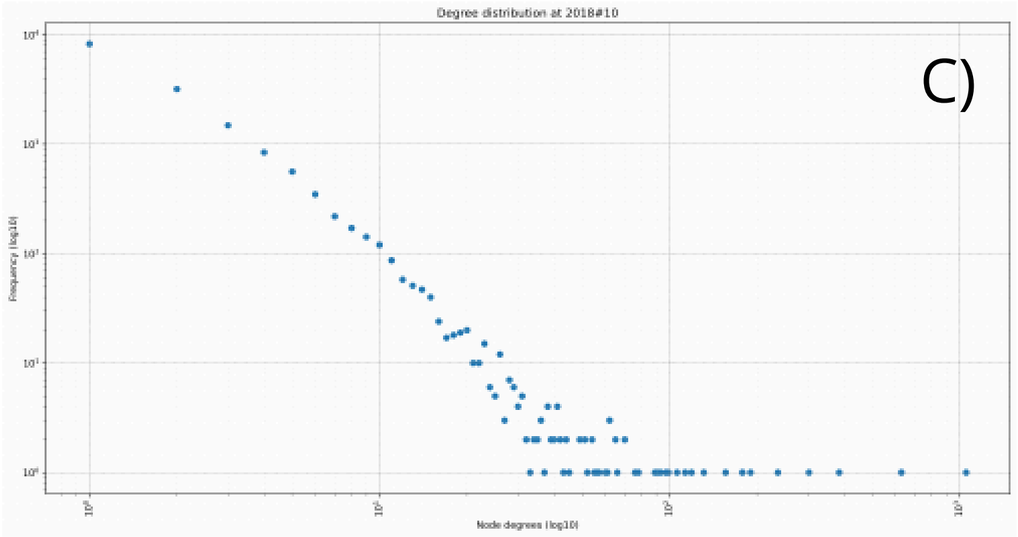
A) Network size (number of nodes)

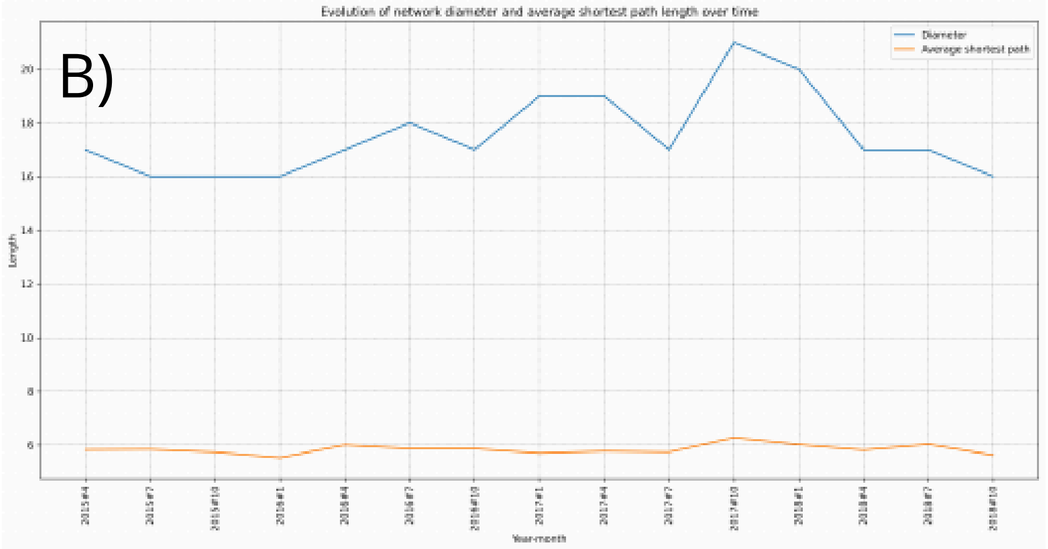
B) Average Path Length and Diameter (GCC)

Network metrics per **snapshot:**

C) Degree distribution in 11-2018 (log-log)

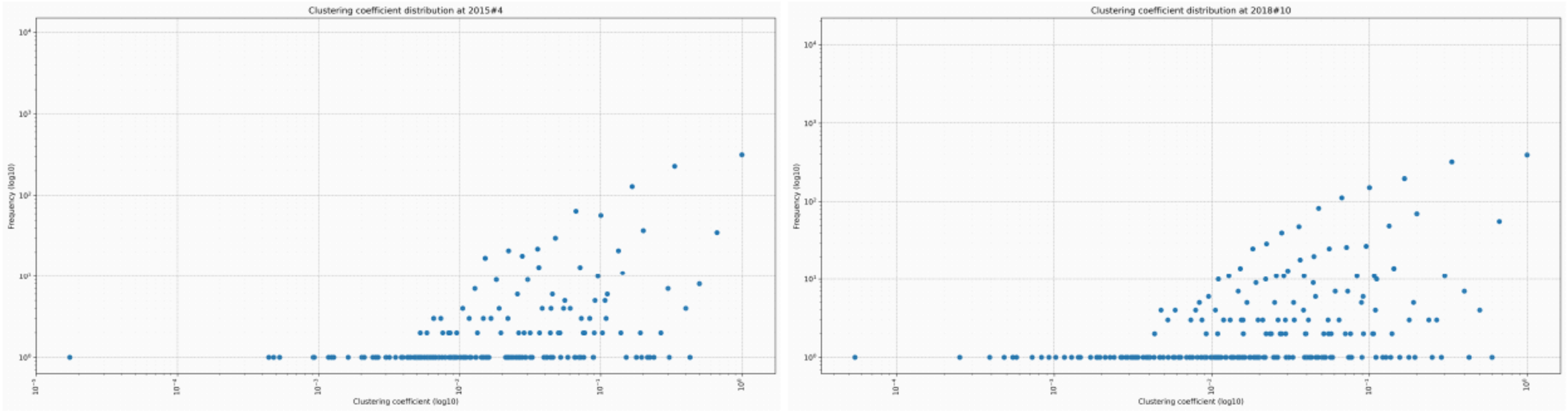






Network Analysis

- High **sparseness**
- Dominance of **weakly connected components**

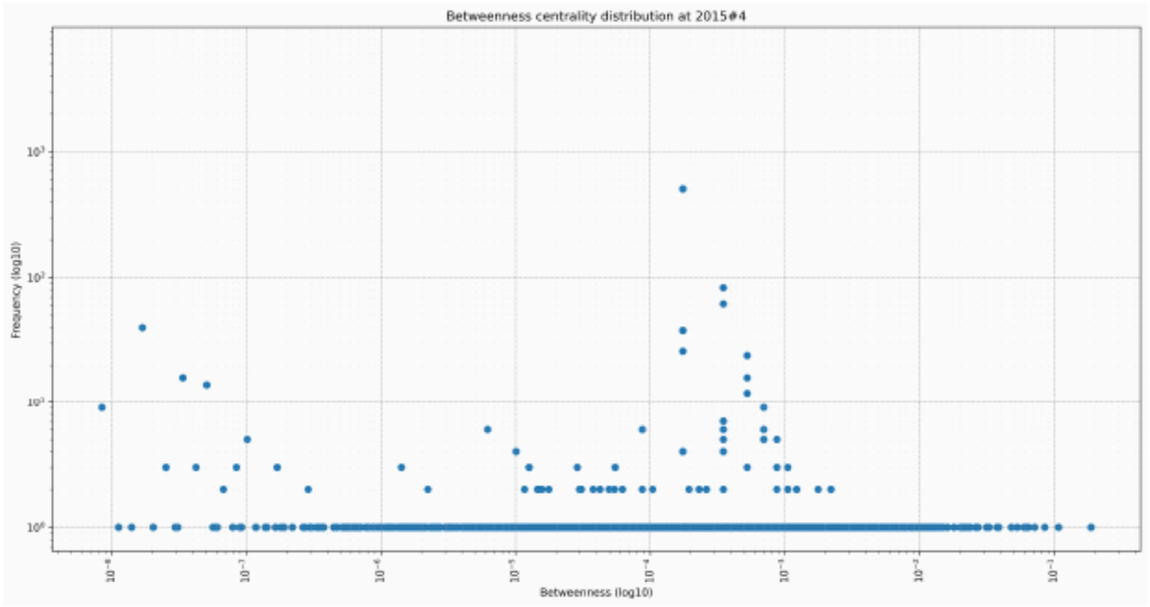


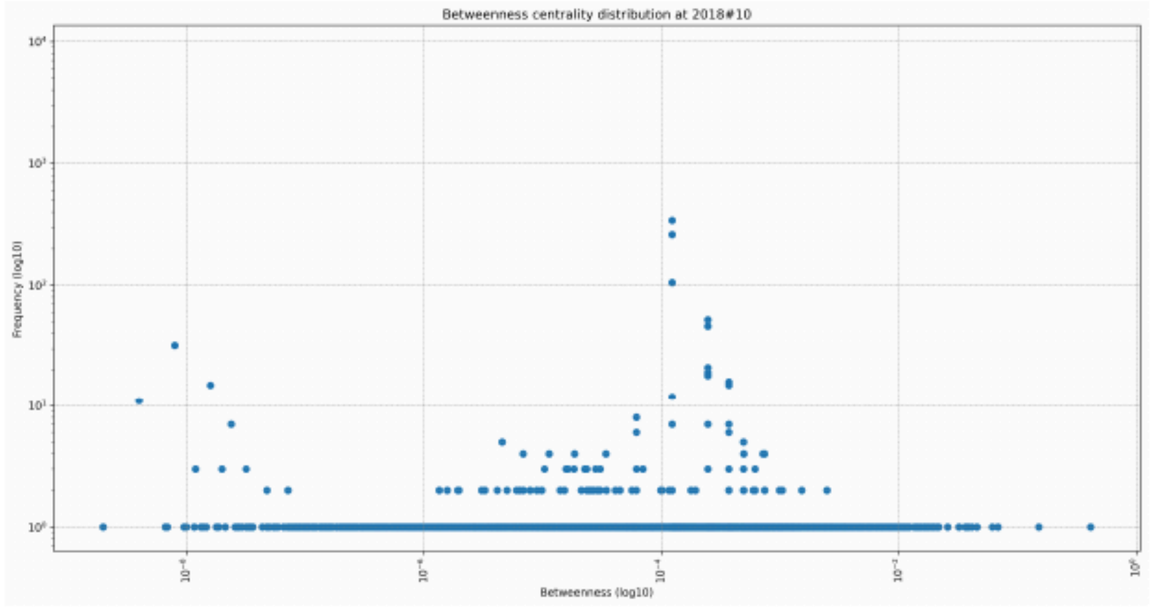
(a) Log-scaled distribution in the April 2015 snapshot (b) Log-scaled distribution in the October 2018 snapshot

Fig. 3. Network clustering coefficient distributions

Network Analysis

- Centrality metric: **Betweenness**
- Dominance of **users with low Betweenness** (not bridges)





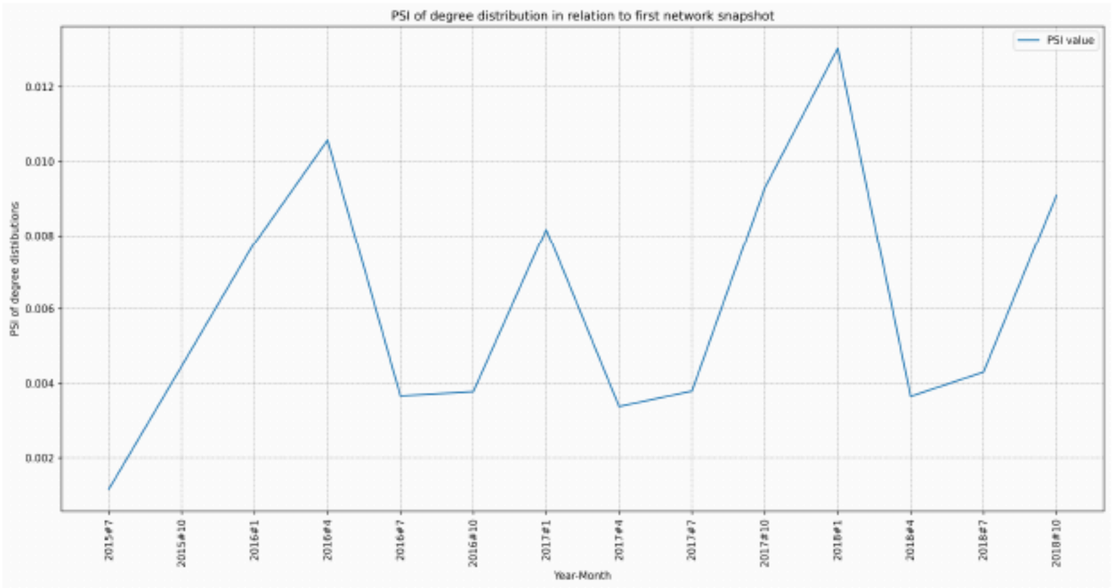
(a) Log-scaled distribution in the April 2015 snapshot (b) Log-scaled distribution in the October 2018 snapshot

Fig. 4. Network betweenness centrality distributions

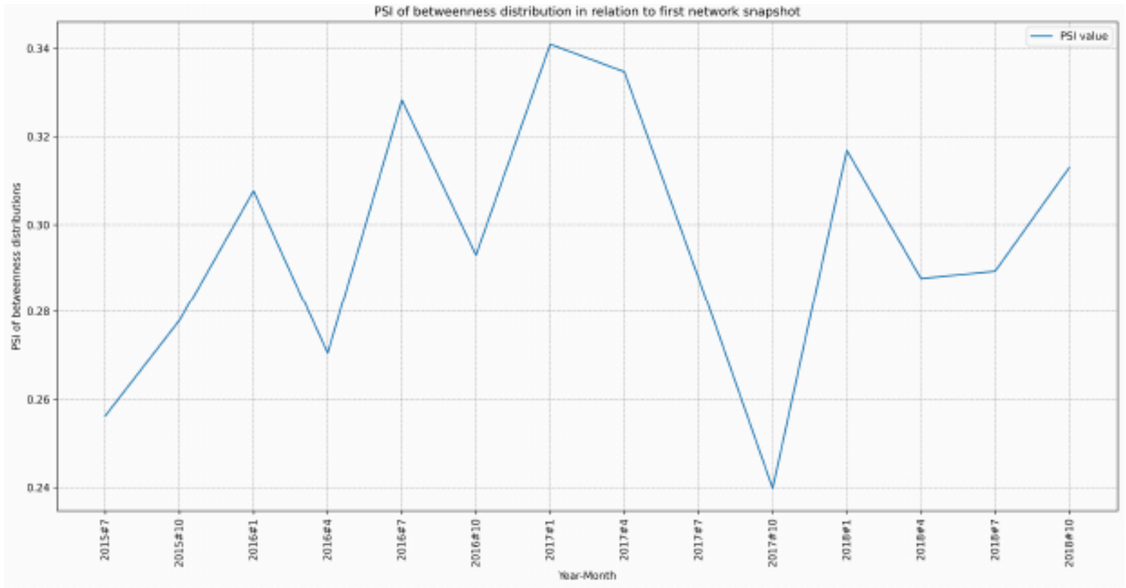
Network Analysis

Population Stability Index (PSI), 04-2015 as baseline:

- Degree distribution **approximately stable:** PSI < 0.2
- Betweenness centrality is **more unstable:** 0.24 < PSI < 0.34



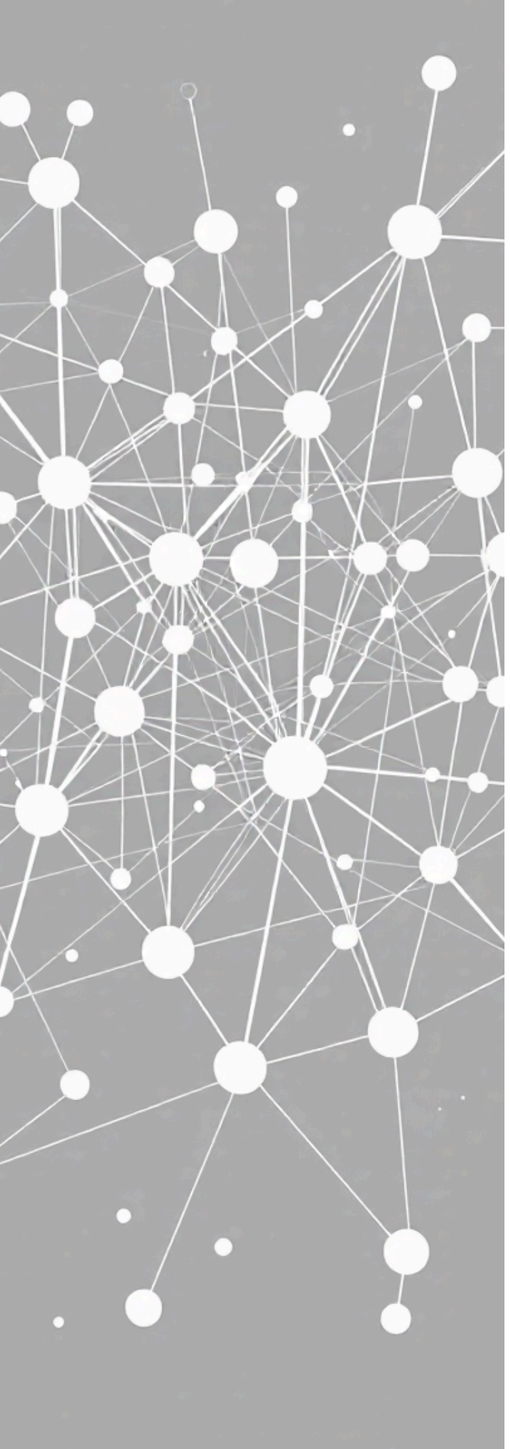
(a) Stability across degree distributions over time



(b) Stability across betweenness centrality distributions over time

Fig. 5. Population Stability Index (PSI) values for key network metrics

Community Analysis Results



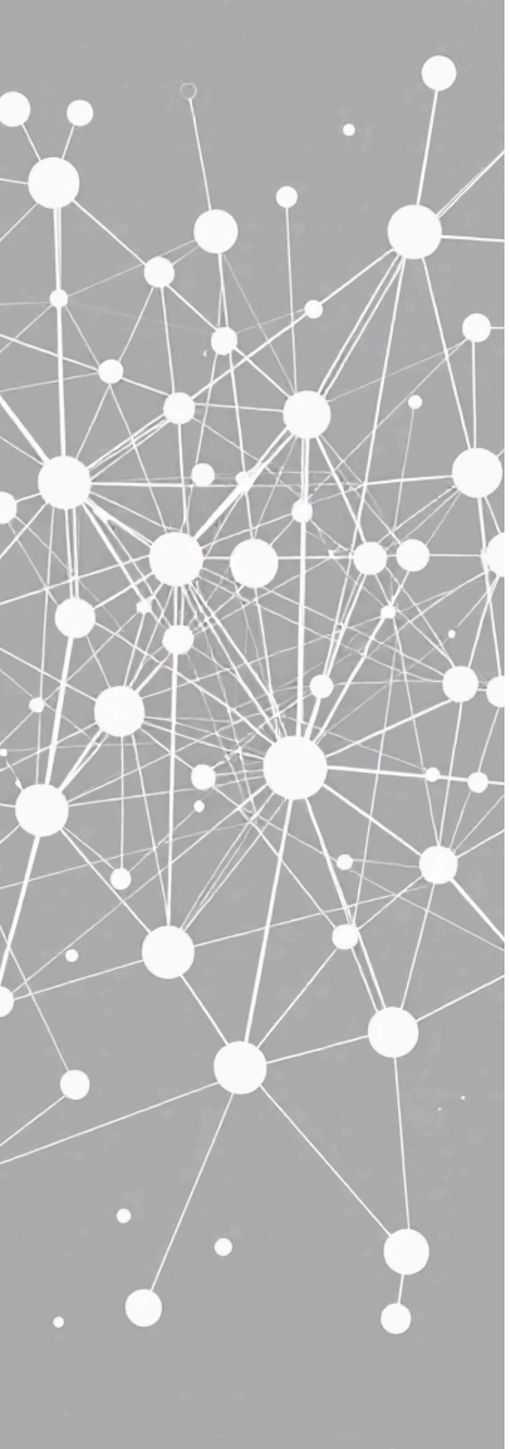
Community Analysis

"Liquid" Network characteristics:

- Extreme **instability** at the **microscopic** level:
  - **Population is in constant flux** (unstable PSI for Betweenness centrality)

versus

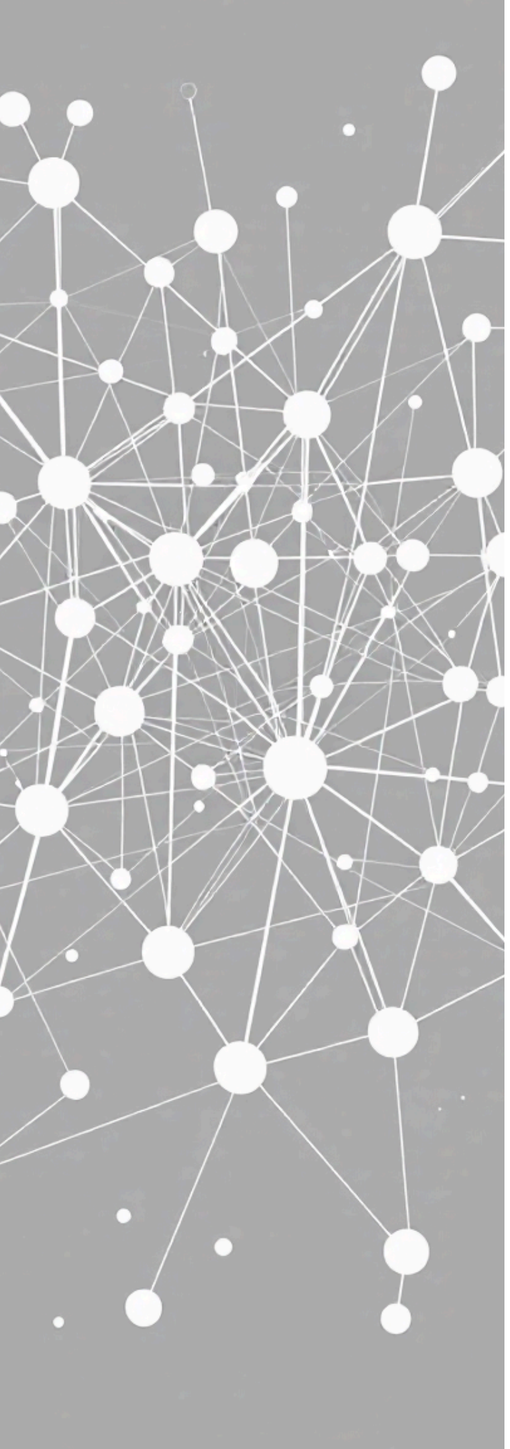
- **Stability** at the **macroscopic** level:
  - **Stable PSI for degree distribution**



Community Analysis

"15 Minutes of Fame" phenomenon:

- **High Turnover**
- **Hub Survival Rate** is **very low** (5%-8%)
- **Only Constant Hub:** AutoModerator
- No social leaders, **Influence** is driven by **viral content**, which fades quickly



Community Analysis

"Nomad" Effect:

- **Zero fidelity**
- **Communities** represents **Topics, NOT Cliques**
  - **Global Average Fidelity:** 0.004
  - **Switching Rate:** 99.6%-100%

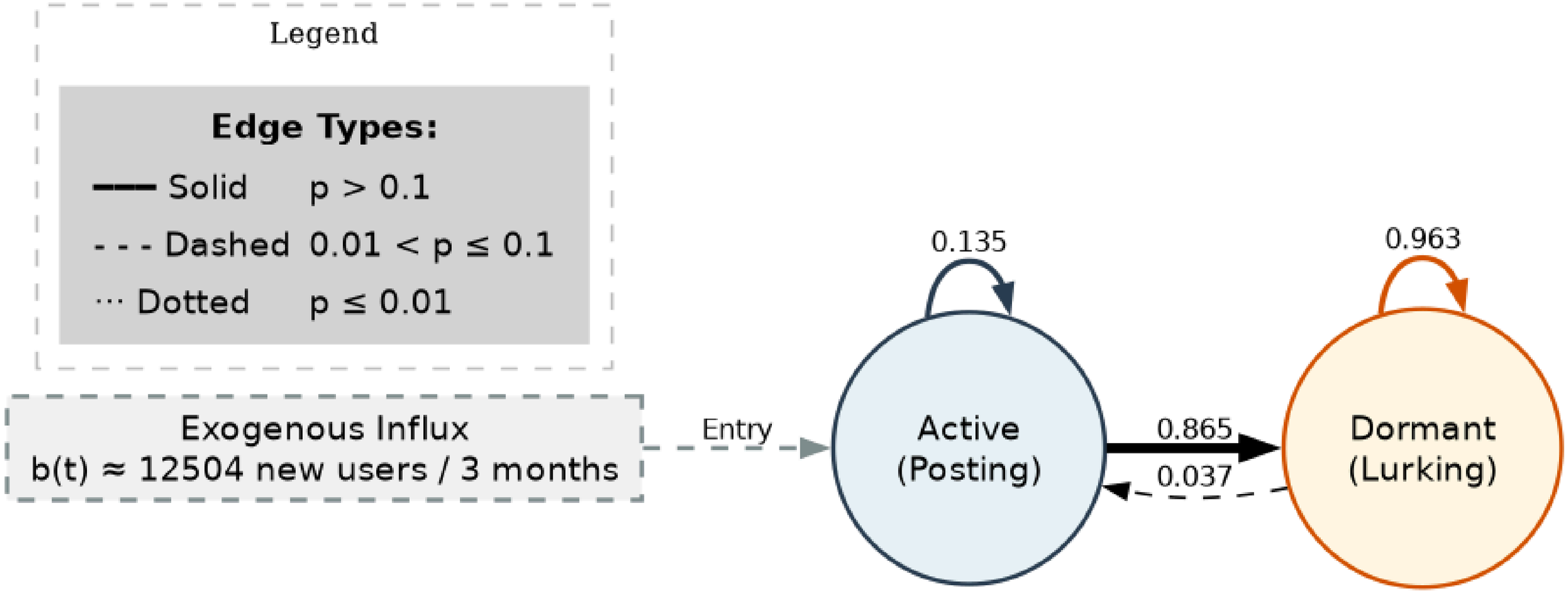
Theoretical Model

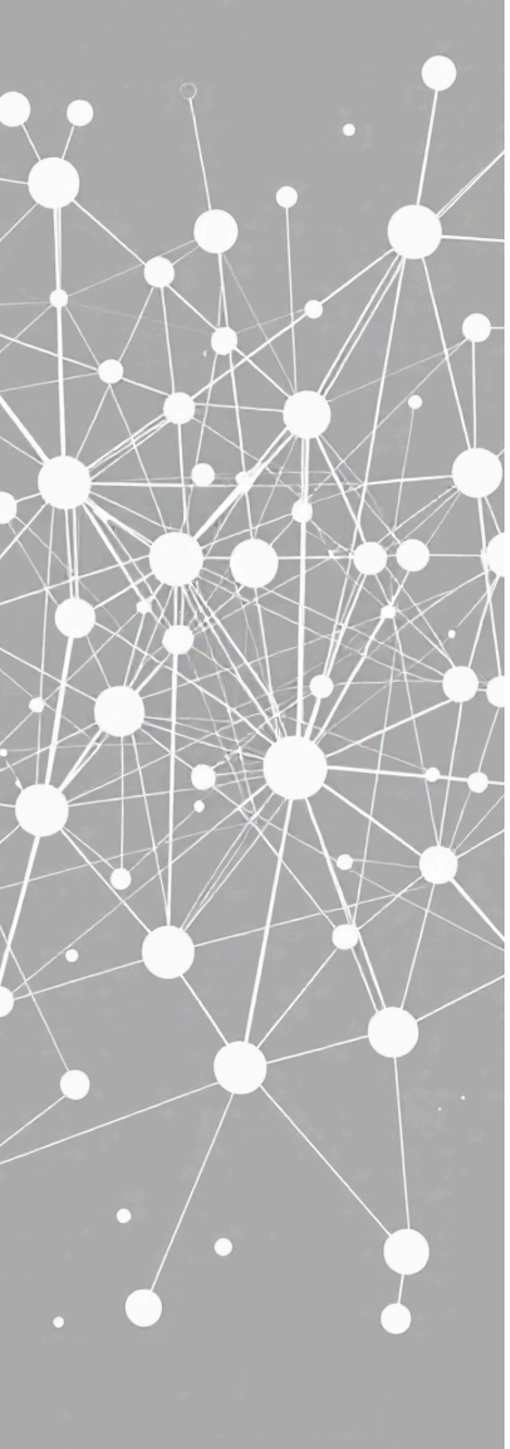
Model Considerations

- **Markov Model** for user lifecycle;
- Content-driven growth, **Fitness model**, for network generation;
- **Model training** and **performance analysis**

Markovian Lifecycle

P(‘going state S(t+1)’|’current state S(t)’)





Theoretical Model

Standard Barabási-Albert

Fails because it **favors older nodes,** and it is designed for **undirected acyclic graphs** and no node deletion, only edge rearrangement

versus

Hybrid Model:

- **Price's Model: directed** preferential **attachment**
- **Bianconi-Barabási (Fitness Model): new nodes** with high "**fitness"** (viral content) rapidly **overtake older nodes**
- Explicit **reciprocity:** user x replies to y, and then y to x

Model Performance and Validation

- **Power-Law Fit Quality (R^2 > 0.9)**

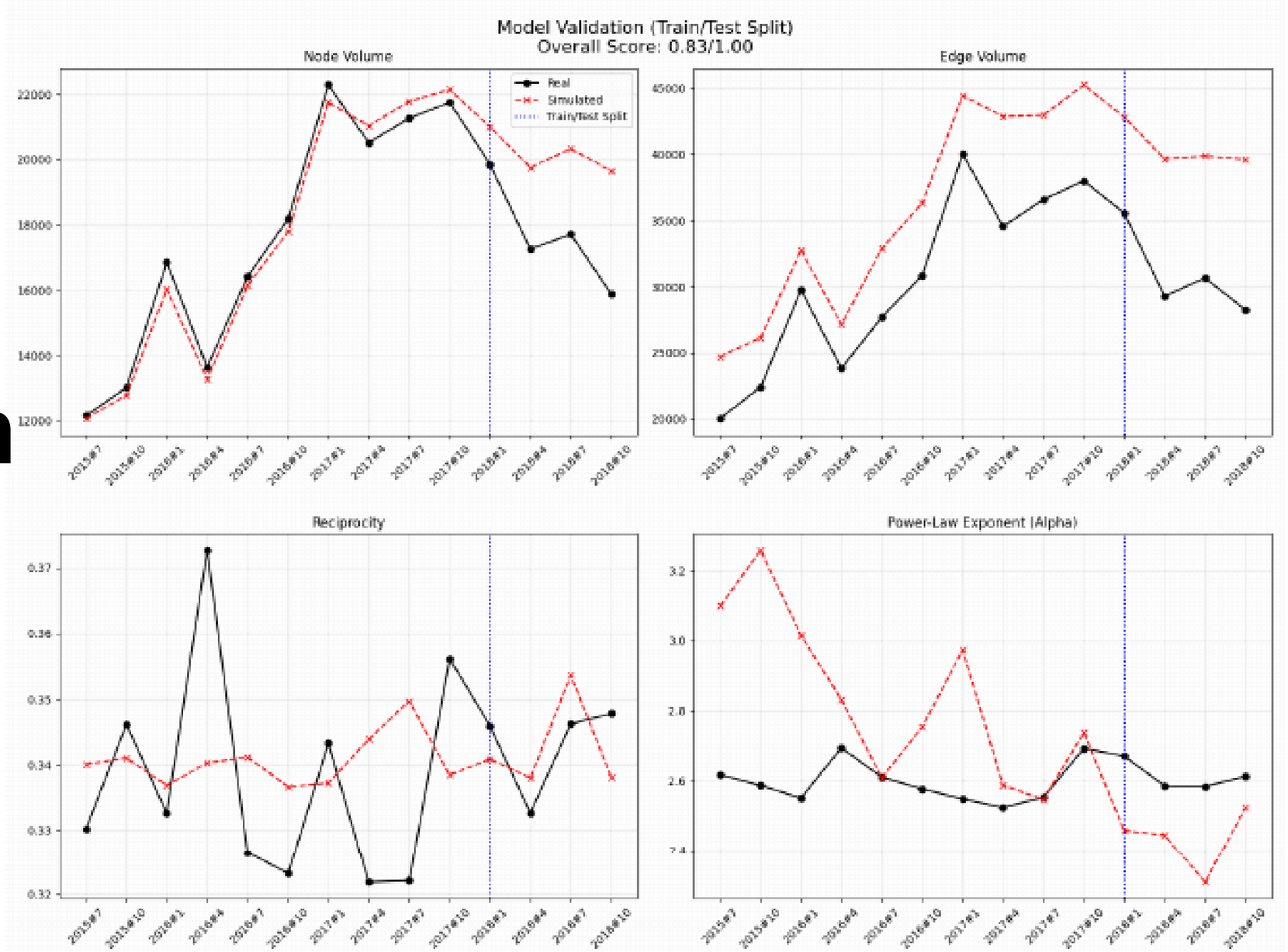
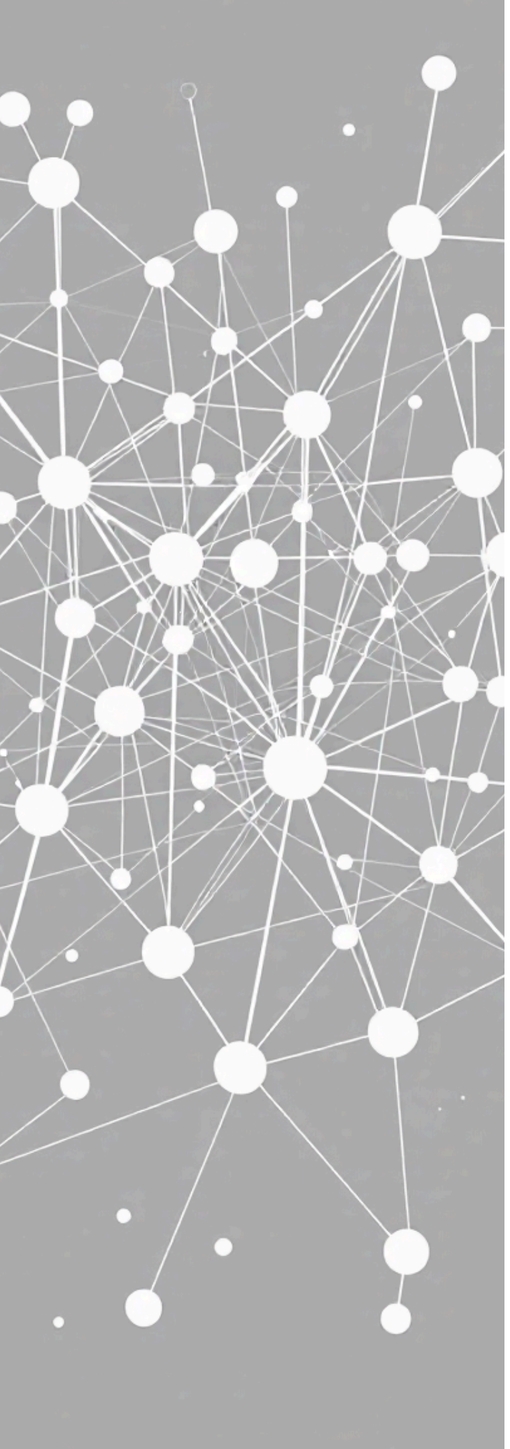
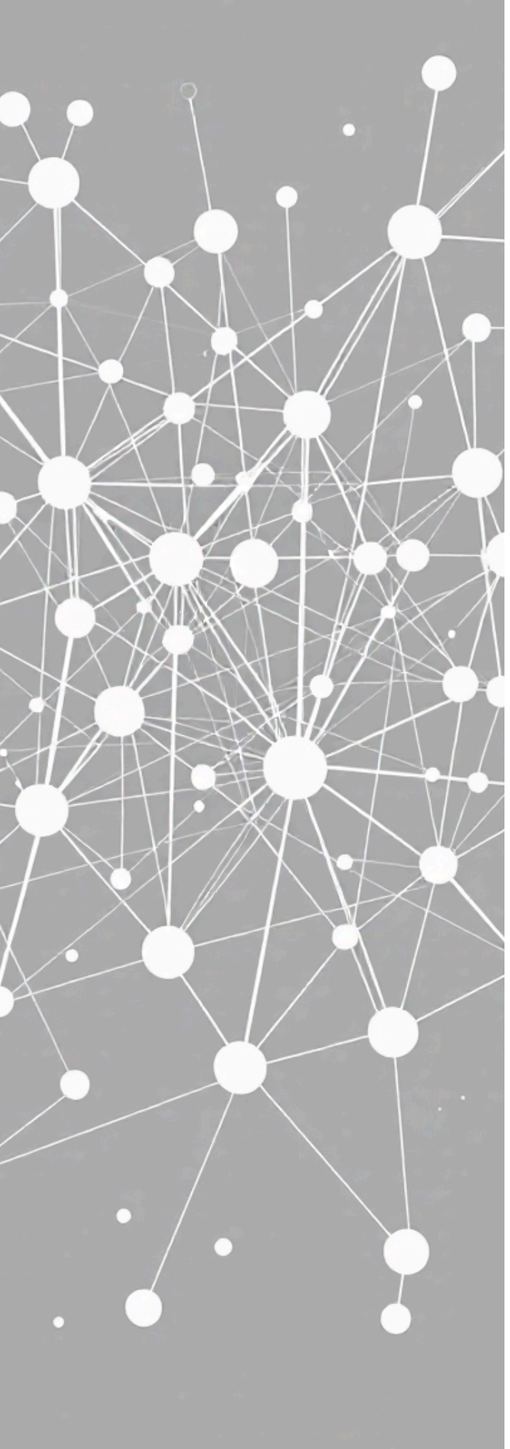


Fig. 7. Model Validation (Train/Test Split). The vertical blue line marks the split between the training phase (parameter estimation) and the testing phase (forecasting). The model successfully captures the trends in edge volume and reciprocity, and approximates the power-law exponent (α).



Conclusion

- **"Liquid" Stability:** The network functions as a "Liquid Network," where robust macroscopic stability persists despite extreme microscopic instability, resembling a transport hub rather than a community.
- **Transient Influence:** "Hub" status is temporary and content-driven, with the AutoModerator bot remaining the only permanent fixture.
- **Fitness-Driven Growth:** Structural evolution is best modeled by "Fitness" (Bianconi-Barabási) rather than age, as new viral content constantly displaces older nodes.



Future Considerations

- Integration of **sentiment scores** as a tool in **community identification**;
- Stricter **bot filtering** to isolate purely organic user behavior;
- Improve **parameter estimation** in **theoretical model** during training:
  - by **sampling** on distributions of the training data, for example, to determine the alphas for the degree distribution;
  - fit a **forecasting model** to predict number of nodes and the ratio of (num edges/ num nodes) over time