Misinformation Cascade Analysis on X (Twitter)

Ahmad Suleiman, Celine Taki. Pradyun Shrestha. Niharika Belavadi Shekar, Andy Mai, Jeremy Albios

INTRODUCTION

- Misinformation spreads rapidly on social media platforms.
- Understanding diffusion patterns is crucial for detection and mitigation.
- Need to identify key amplifiers and echo chambers.
- We need to compare fake vs. real news spread patterns.

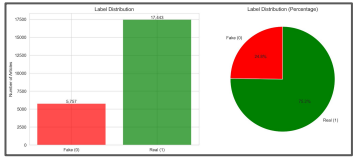
What we are doing:

- Analyzing complete life cycle of false information
- Identifying key amplifiers through retweets/replies
- Tracking temporal diffusion patterns
- Exposing underlying structures enabling viral spread

DATASET & METHODOLOGY

Dataset:

- FakeNewsNet from Arizona State University
- 23,196 articles (5,755 fake, 17,441 real)
- Sources: GossipCop (22,140) + PolitiFact (1,056)
- 1.9M synthetic users, 3.6M interactions



Methodology Pipeline:

1. Data Acquisition & Text Preprocessing
2. Graph Construction (Bipartite, Similarity, User Interaction)
3. Link Analysis (PageRank, HITS)
4. Community Detection (Louvain)
5. Cascade Modeling (Independent Cascade Model - ICM)

KEY RESULTS - NETWORK STATISTICS

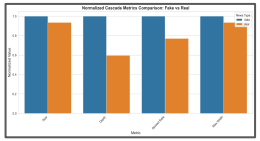
User Interaction Graph: 1.9M nodes, 3.6M edges
Bipartite Graph: Article-Tweet connections
Similarity Graph: Article-Article similarity
Communities Detected: 77,263

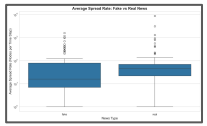
KEY FINDINGS - FAKE VS REAL COMPARISON

Engagement Patterns:

- Fake news: 132.8 tweets/article
- Real news: 74.8 tweets/article
- 77% higher engagement for fake news

Key Insight: Fake news spreads 30% faster than real news

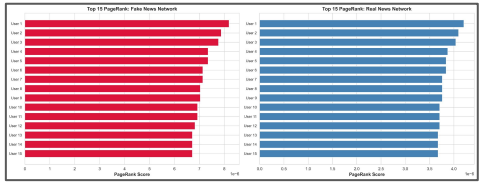




LINK ANALYSIS RESULTS

PageRank & HITS:

- Identified top influencers in misinformation network
- Authority scores: Users whose content is frequently shared
- Hub scores: Users who frequently share others' content

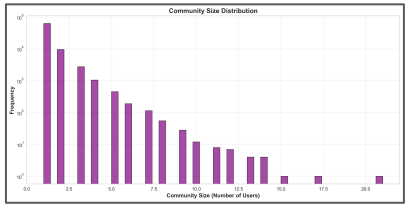


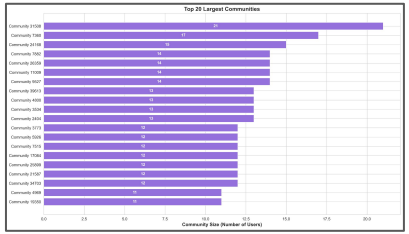
COMMUNITY DETECTION RESULTS

Louvain Algorithm:

- 77,263 communities detected
- Modularity: 0.999892 (very high - strong community structure)

Key Finding: Strong clustering suggests users form tight-knit groups sharing similar content.





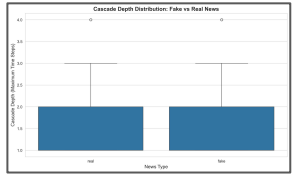
CASCADE MODELING RESULTS

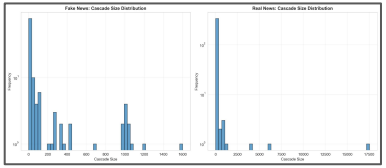
Independent Cascade Model (ICM):

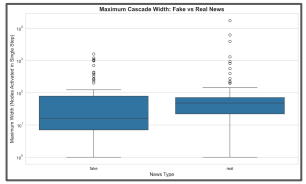
- Simulated information diffusion with activation probability 0.1
- Tracked cascade depth, width, and spread rate

Key Findings:

- Fake news cascades are shallower (avg depth: 0.22) but wider
- Real news has deeper cascades but slower initial spread
- Fake news reaches maximum width faster







CONCLUSION

Key Takeaways:

1. Fake news shows higher engagement and faster spread rates
2. Strong community structure detected
3. Distinct influencer patterns in fake vs. real networks
4. Cascade modeling reveals temporal differences in diffusion
5. Interactive tool enables real-time detection

Implications:

- Understanding these patterns can improve detection systems
- Community structure analysis helps identify echo chambers
- Cascade metrics can predict viral potential
- Real-time tools can help users fact-check information