Network Science - Final Project Proposal

Adnana Ivana, Jiatong Li, Nidal Ogur, Marvin Süss

14 November 2025

1 Motivation

User-provided content has become available on all social media platforms, which causes unchecked rumors spread [5]. While previous research focused on how these rumors diffuse, through network structures [4], emotional sources [3] or other mechanisms, we are motivated to study how different strategies, such as early correction [7] or targeting influential users can affect the structure of rumor propagation.

Thus, the research question is: How can different fact-checking strategies influence the spread of misinformation in online social networks?

2 Network Structure

In order to answer the research question, we intend to build a network in which nodes represent users in the social network who post, and interact through messages, and directed edges represent the flow of information between users. The provided data set is the PHEME Dataset for Rumour Detection and Veracity Classification. This dataset contains annotated rumors and non-rumors collected from Twitter during breaking news. It includes tweets, replies, retweets and annotations of veracity, which can be either true, false or unverified.

Misinformation will be analysed by tracking the propagation of information in the network, and keeping in account the different types of information, based on veridicity.

3 Methodology

3.1 Plan

We use the PHEME dataset to reconstruct time-stamped rumor conversation trees on Twitter (now, X) and, when available, retain stance annotations as features. Each thread is treated as a temporal cascade whose nodes are users and directed edges are either replies, retweets or quotes ordered by their timestamps. In addition, we derive an auxiliary network to study interaction between the users across threads. The purpose is to compute influence and bridge scores (degree/PageRank, betweenness, community membership). After standard cleaning (time normalization, deduplication, language filtering), we compute descriptive baselines, such as cascade size, depth, growth rate, and structural virality, to characterize difficulty and to calibrate models [1].

To study interventions, we simulate counterfactual fact-checking under two canonical diffusion competing mechanisms:

- Independent Cascade Model: In a competing Independent Cascade model, a contact from an M (misinformation-supporting) or F (fact-checked/denying) user activates a susceptible neighbor with probabilities λ_M and λ_F; states are exclusive, with optional small forgetting. [2, 6]
- Threshold Model: In a competing Threshold model, a user adopts M or F when the fraction of active neighbors exceeds θ_M or θ_F (typically $\theta_F \geq \theta_M$). Parameters are chosen by matching simulated and observed prevalence curves. [2]

We then seed corrections at time τ after the source tweet under alternative strategies—earliest responders (early broadcast), high-degree/PageRank hubs, high-betweenness bridges, and community-aware allocation—plus a random baseline, and replay the real temporal edge sequence to measure outcomes. In parallel, we run an observational check: using within-event matched comparisons or survival analysis, we estimate how the appearance of an early deny reply correlates with subsequent support in the same thread.

3.2 Evaluation

Performance of the network will be assesed through different techniques: final misinformation prevalence, reduction versus a no-intervention replay, area under the prevalence curve, and changes in peak timing/height and structural virality [1]. We also report cross-community containment and benefit per seed to capture cost-effectiveness. Models are validated with leave-one-event-out evaluation and multiple stochastic replicates to obtain confidence intervals, and robustness is assessed by shuffling timestamps, varying intervention latency and budget, and perturbing stance/unverified subsets. Results are reported per event and in aggregate with paired statistical tests and fully reproducible code and configurations.

4 Expected Outcome

We expect to obtain a quantitative and comparative understanding of how different fact-checking strategies shape the diffusion of misinformation in online social networks. Specifically, we anticipate that interventions targeting structurally influential users (e.g., high degree or high betweenness nodes) or well-positioned bridge users will outperform random seeding in reducing both final misinformation prevalence and structural virality. Early interventions are expected to produce the largest relative gains, particularly under competitive Independent Cascade dynamics, where small latency gaps can substantially shift cascade trajectories.

We also expect to identify situations in which community-aware strategies are essential, such as events whose diffusion is highly modular or where misinformation is concentrated within echo chambers. In contrast, strategies relying on early responders may be more effective in deep, fast growing cascades where rapid adoption dominates structural influence. The comparison across diffusion models should highlight which mechanisms are most sensitive to fact-checking timing and placement.

Overall, we aim to obtain a clear picture of when early correction, structural targeting, or community-based approaches are most likely to curb misinformation. The findings should connect diffusion outcomes to measurable network properties and provide general insights into how fact-checking can leverage network structure to reduce the reach of false information.

References

[1] Sharad Goel, Ashton Anderson, Jake Hofman, and Duncan J. Watts. The structural virality of online diffusion. *Management Science*, 62(1):180–196, 2016.

[2] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.

[3] Nicolas Pröllochs, Dominik Bär, and S. Feuerriegel. Emotions in online rumor diffusion. *EPJ Data Science*, 10, 2021.

[4] Nicolas Pröllochs and S. Feuerriegel. Mechanisms of true and false rumor sharing in social media: Collective intelligence or herd behavior? *Proceedings of the ACM on Human-Computer Interaction*, 7:1 – 38, 2022.

[5] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, F. Petroni, Antonio Scala, G. Caldarelli, H. Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113:554 – 559, 2016.

[6] Duncan J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.

[7] Lan Yang, Zhiwu Li, and A. Giua. Containment of rumor spread in complex social networks. *Inf. Sci.*, 506:113–130, 2020.