Mathematics of Fake News Propagation: A Probabilistic and Network-Based Approach

Author: Neelu Khandelwal, M.Sc. Mathematics, Assistant Professor, Department of Mathematics, Vikrant Institution of Technology and Management, Indore

Abstract

The rapid spread of misinformation, popularly known as fake news, has become a significant concern in the digital age. Unlike traditional information, fake news is designed to be emotionally appealing, making it more viral and difficult to control. This paper explores a mathematical framework for analyzing fake news propagation by combining concepts from probability theory, network science, and epidemiological modeling. Social media networks are modeled as graphs, where users represent nodes and connections represent potential channels of information diffusion. We adapt the classical Susceptible–Infected–Recovered (SIR) model, originally used in epidemiology, to study how fake news spreads through a population. The proposed model accounts for the probability of individuals sharing fake news, the rate of fact-checking interventions, and the role of influential nodes in accelerating dissemination. Analytical results show that fake news tends to spread faster than verified information due to its higher transmission rate. The study highlights the importance of mathematical modeling in understanding the dynamics of misinformation and provides insights for policymakers, social media platforms, and educators in designing effective countermeasures.

1. Introduction

The digital revolution has transformed the way information is consumed and shared. Platforms such as Facebook, Twitter, and WhatsApp allow news to reach millions within seconds. While this enhances connectivity, it also facilitates the rapid dissemination of misinformation, commonly referred to as fake news. The impact of fake news is profound—it can influence elections, incite violence, and create widespread confusion in public health crises.

Traditional studies on fake news often emerge from communication studies, psychology, or political science. However, these perspectives may not fully capture the complex mechanisms governing its spread. Mathematical modeling, with its ability to abstract, simulate, and predict dynamic systems, provides a powerful framework for understanding misinformation diffusion.

This paper aims to address the following research questions:

1. How can mathematical tools such as probability theory and graph models be applied to fake news propagation?
2. What factors accelerate the spread of fake news compared to true information?
3. How can these models inform strategies for controlling misinformation?

2. Literature Review

The spread of information has often been studied using models from epidemiology.

The SIR model (Kermack & McKendrick, 1927) is widely used to describe disease transmission, where individuals transition between susceptible, infected, and recovered states. In recent years, researchers have applied this framework to study the diffusion of rumors and viral content online.

Daley and Kendall (1965) developed one of the earliest mathematical models for rumor spreading, where the probability of transmission depended on contact between individuals. More recently, Nekovee et al. (2007) applied epidemic models to information flow in technological and social networks. However, fake news presents unique challenges compared to ordinary rumors. Its emotionally charged and sensational nature increases the likelihood of sharing, making its effective transmission rate higher than truth-based information.

Although several studies exist on rumor and epidemic models, few focus specifically on fake news propagation with unique characteristics such as virality, emotional bias, and the role of influencers. This gap motivates the need for a refined mathematical framework.

3. Mathematical Modeling Framework

3.1 Graph-Theoretic Representation

We model social media as a graph G(V,E), where V represents users and E represents social connections. Each node has a probability of encountering and spreading news. Influential nodes (users with high degree centrality) play a critical role in accelerating spread.

3.2 Probabilistic Model

Let pf represent the probability of a user sharing fake news and pr represent the probability of sharing real news. Empirical studies suggest pf > pr, reflecting the greater virality of misinformation.

3.3 Adapted SIR Model

We adapt the SIR epidemic model as follows:

- Susceptible (S): Users who have not yet encountered the fake news.
- Infected (I): Users who believe and share fake news.
- Recovered (R): Users who stop spreading fake news after verification or disbelief.

The dynamics are described by differential equations:

$$dS/dt = -\beta SI$$
$$dI/dt = \beta SI - \gamma I$$
$$dR/dt = \gamma I$$

Here, β is the rate of fake news transmission, and γ is the correction rate (fact-checking, awareness campaigns). If $\beta > \gamma$, fake news spreads rapidly; if γ is high, misinformation decays quickly.

4. Detailed Example and Analysis

Consider a population of 1,000 users with initial conditions S(0)=999, I(0)=1, R(0)=0. Let $\beta=0.6$ and $\gamma=0.3$. The basic reproduction number is $R0 = (\beta/\gamma)S(0) \approx 2$, meaning each infected user initially creates 2 more. Numerical simulation shows:

- Peak infected fraction: ~15.4% (≈154 users).
- Time to peak: ~22.5 units (days if rates are per day).
- Final affected population: ~797 users (≈80%).

This illustrates that fake news spreads quickly and widely under these conditions. If interventions increase γ to 0.6 or reduce β to 0.3, $R0$ falls below 1, preventing widespread outbreaks. Thus, fact-checking and content moderation are effective strategies.

5. Applications and Discussion

The proposed model provides practical insights:
- Policy Making: Identifying thresholds where misinformation becomes uncontrollable.
- Social Media Platforms: Targeting influential nodes for early intervention.
- Education: Promoting critical thinking to reduce transmission rates.
- Public Health: Predicting and managing misinformation during crises.

These results highlight how mathematics bridges theoretical understanding with practical solutions.

6. Conclusion and Future Work

This paper presented a mathematical approach to understanding fake news propagation, combining probability theory, graph models, and epidemiological dynamics. The analysis demonstrates that fake news spreads faster than truth due to its higher transmission probability and the role of influential nodes.

Future work may include machine learning for parameter estimation, testing the model with real social media data, and using neutrosophic logic to capture uncertainty in user beliefs. Interdisciplinary collaboration will be essential.

References

- Daley, D. J., & Kendall, D. G. (1965). Stochastic rumours. Journal of the Institute of Mathematics and Its Applications, 1(1), 42–55.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London. Series A, 115(772), 700–721.
- Nekovee, M., Moreno, Y., Bianconi, G., & Marsili, M. (2007). Theory of rumour spreading in complex social networks. Physica A, 374(1), 457–470.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151.
- Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on Twitter. Proceedings of the 7th Workshop on Social Network Mining and Analysis, 1–9.