MODELING AND ANALYSIS OF DIFFUSION PROCESSES IN NETWORK STRUCTURES: AN OPERATIONS RESEARCH APPROACH

Dr. Harsh Vardhan
Assistant Professor, Faculty of Education, Teerthanker Mahaveer University, Moradabad, U P, India

Dr. Rajeev Kumar
Assistant Professor, Head Department of Mathematics, Devta P G College Morna, Bijnor, U P, India

Abstract

Diffusion processes in networks are fundamental to understanding phenomena such as the spread of information, innovation, diseases, and behaviors in complex systems. This paper presents a mathematical and operations research-based framework for analyzing diffusion in deterministic and stochastic network models. We examine both threshold and cascade-based models, incorporating probabilistic and optimization perspectives to assess the speed and coverage of diffusion under varying network topologies. Furthermore, we investigate control mechanisms, such as seed node selection, to maximize influence under resource constraints. Real-world applications include marketing, epidemiology, and infrastructure resilience. The study is grounded in mathematical rigor and substantiated by recent literature in the fields of network science and operations research.

Keywords: Diffusion process, network model, operations research, influence maximization, graph theory, epidemic modeling, optimization

1. Introduction

The study of diffusion processes in networks lies at the intersection of operations research (OR), computer science, sociology, and epidemiology. As our world becomes increasingly interconnected—digitally, socially, and physically—the mechanisms by which information, influence, or contagion spread across networks have become central to a wide range of applications (Newman, 2010). Whether it is the viral spread of content on social media, the propagation of power failures in electrical grids, or the transmission of infectious diseases, the underlying phenomenon is governed by a networked diffusion process (Pastor-Satorras et al., 2015).

In the context of operations research, diffusion processes are particularly relevant due to the optimization and decision-making challenges they pose. For instance, selecting a minimal set of individuals in a network to maximize information spread (the influence maximization problem) or to contain an outbreak (the containment or immunization problem) are typical OR formulations (Kempe, Kleinberg, & Tardos, 2003; Chen, Wang, & Yang, 2009). These problems often involve complex trade-offs between efficiency, cost, and reach, and require a mathematical understanding of both the network structure and the dynamics of the diffusion process (Jackson, 2008).

Two widely studied classes of diffusion models are the threshold-based and probabilistic cascade-based models. The Linear Threshold Model (LTM) assumes that a node becomes active if a certain proportion of its neighbors are already active, reflecting social reinforcement (Granovetter, 1978; Valente, 1996). The Independent Cascade Model (ICM) considers a stochastic process where each activated node has a single chance to influence its neighbors with a given probability (Kempe et al., 2003). These models capture different real-world scenarios, from peer pressure-driven behavior adoption to viral marketing and epidemic spread (Domingos & Richardson, 2001; Centola, 2010).

Despite their simplicity, these models can yield complex emergent behavior depending on the topology of the underlying network—whether it is scale-free, small-world, random, or modular (Barabási & Albert, 1999; Watts, 2002). Consequently, a rigorous mathematical and algorithmic framework is essential to predict and control diffusion phenomena effectively.

This paper aims to:

- Develop a mathematical formalization of diffusion processes in networks.
- Analyze key parameters affecting diffusion, such as threshold distributions, activation probabilities, and network structure.
- Investigate optimization techniques for seeding strategies and containment.
- Review practical applications and real-world case studies.
- Provide simulations and visualizations to support theoretical insights.

By integrating tools from graph theory, probability, and optimization, we contribute to a deeper understanding of how to model, analyze, and strategically intervene in diffusion processes in complex networks. This is increasingly important not only for theoretical advancement but also for addressing urgent real-world challenges in public health (Kermack & McKendrick, 1927), marketing (Goyal, Lu, & Lakshmanan, 2011), cybersecurity (Liu, Slotine, & Barabási, 2011), and infrastructure resilience (Liu & Zhang, 2020).

2. Mathematical Foundations

Understanding diffusion processes in networks requires a solid mathematical foundation grounded in graph theory and probability. We define a network as a graph $G = (V, E)$, where V is the set of nodes (or vertices), and $E \subseteq V \times V$ is the set of edges representing interactions or connections between nodes. The state of each node is modeled as a binary function $f : V \rightarrow \{0, 1\}$, where $f(v) = 1$ indicates an active (influenced or infected) node and $f(v) = 0$ indicates an inactive node.

Diffusion unfolds over discrete time steps $t = 0, 1, 2, \ldots$, with the state of each node at time $t + 1$ depending on the configuration of its neighbors at time t. Two canonical models form the basis of our analysis: the Linear Threshold Model (LTM) and the Independent Cascade Model (ICM). These models offer complementary perspectives—deterministic and stochastic—on how influence or contagion spreads.

2.1 Linear Threshold Model (LTM)

The Linear Threshold Model reflects situations where individuals require social reinforcement to adopt a new behavior, product, or idea.

- Threshold Distribution: Each node $v \in V$ is assigned a threshold $\theta_v \sim U(0, 1)$, drawn from a uniform distribution, representing the minimum cumulative influence required for activation.
- Edge Weights: Each incoming edge $(u, v) \in E$ carries a weight $w_{uv} \in [0, 1]$, with the condition that for each node v, the total incoming weight satisfies $\sum_{u \in N(v)} w_{uv} \leq 1$, where $N(v)$ denotes the set of in-neighbors of v.
- Activation Rule: At each time step t, a node v becomes active if the sum of weights from its active neighbors A_t satisfies:

$$\sum_{u \in A_t \cap N(v)} w_{uv} \geq \theta_v$$

Once activated, a node remains active for the duration of the process (i.e., activation is irreversible).

This model captures behavioral diffusion driven by cumulative influence, such as the adoption of innovations or voting behavior.

2.2 Independent Cascade Model (ICM)

The Independent Cascade Model emphasizes probabilistic influence, modeling one-time attempts at activation without reinforcement.

- Activation Probability: Each directed edge $(u, v) \in E$ is associated with a fixed probability $p_{uv} \in [0, 1]$, representing the likelihood that node u, once activated, will activate node vv.
- One-Time Attempt Rule: When a node u becomes active at time t, it has one chance to activate each inactive neighbor $v \in N(u)$ in the next time step $t + 1$. Activation succeeds with probability p_{uv}.
- Propagation: The process continues until no new activations occur.

Formally, for an active node u at time t, and each inactive neighbor v, activation follows a Bernoulli trial:

$$\Pr(f_{t+1}(v)) = 1 \mid f_t(u) = 1, f_t(u) = 0) = p_{uv}$$

This model is particularly suited for modeling the viral spread of content, malware, or infectious diseases where exposure alone can be sufficient for adoption or infection.

2.3 Model Comparison and Implications

Feature	Linear Threshold Model (LTM)	Independent Cascade Model (ICM)
Type of Process	Deterministic (with probabilistic thresholds)	Stochastic
Activation Trigger	Cumulative influence exceeds a threshold	Single random attempt by neighbors
Influence Accumulation	Required	Not required
Use Cases	Behavioral diffusion, social pressure	Epidemics, viral marketing, malware spread
Activation Duration	Permanent	Permanent
Influence Propagation Timing	Synchronous across all nodes	Asynchronous, round-by-round attempts

These foundational models form the basis for the optimization and control techniques explored in the following sections. They are flexible enough to be extended with features such as time delays, adaptive probabilities, or multi-state dynamics, which are useful for real-world scenarios with higher complexity.

3. Optimization Objectives in Diffusion

At the heart of many diffusion-based applications lies a central question: How can we strategically intervene in a network to maximize (or minimize) the spread of influence, information, or contagion? In operations research (OR), this is formalized as a class of optimization problems with significant theoretical and practical implications.

3.1 Influence Maximization Problem

One of the most prominent OR formulations in network diffusion is the Influence Maximization problem. The goal is to select a subset of nodes (called seed nodes) whose initial activation results in the largest expected spread across the network. Mathematically, the problem is stated as:

$$\max_{S \subseteq V, |S| \leq k} E[|\sigma(S)|]$$

Where:

- SS is a subset of nodes (the seed set), constrained by $|S| \leq k$,
- $\sigma(S)$ is the random set of nodes activated by the end of the diffusion process starting from S,
- $E[|\sigma(S)|]$ is the expected number of nodes influenced.

This problem is both combinatorial and submodular in nature. A function $f : 2^V \rightarrow R$ is *submodular* if it exhibits diminishing returns — that is, for any $A \subseteq B \subseteq V$ and $v \notin B$, the marginal gain from adding v is higher for the smaller set:

$$f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B)$$

Kempe, Kleinberg, and Tardos (2003) showed that under both the Independent Cascade (ICM) and Linear Threshold (LTM) models, the influence function $E[|\sigma(S)|]$ is monotonic and submodular. This allowed them to apply a greedy algorithm that selects nodes one at a time based on marginal gains. Despite the NP-hardness of the problem, the greedy approach provides a (1 – 1/e) ≈ 63% approximation guarantee — a strong result in approximation algorithms.

3.2 Extensions and Constraints

In real-world settings, additional constraints complicate the optimization:

- Cost-aware influence maximization: Each node has a different activation cost. The problem becomes a budgeted version:

$$\max_{S \subseteq V, c(S) \leq B} E[|\sigma S|]$$

where $c(S) = \sum_{v \in S} c_v$ and B is the budget.

- Time-constrained diffusion: Optimization targets rapid diffusion within limited time steps, emphasizing early adopters.
- Competitive influence: Multiple actors seed the network (e.g., rival companies in marketing), leading to game-theoretic models (Goyal et al., 2011).
- Dynamic or streaming networks: Networks evolve over time, requiring real-time or adaptive seeding strategies (Liu & Zhang, 2020).

- Community-aware strategies: Influence maximization is performed with awareness of community structure to enhance spread efficiency (Wang et al., 2010).

These variations demand tailored algorithms—ranging from modified greedy heuristics (Chen et al., 2009) to metaheuristics (e.g., simulated annealing, genetic algorithms) and learning-based methods (e.g., reinforcement learning for seed selection).

3.3 Containment and Immunization

The dual of influence maximization is containment—minimizing the spread of harmful contagions, such as misinformation or disease. The objective is to select a subset of nodes to immunize or block to reduce the expected spread:

$$\min_{S \subseteq V, |S| \leq k} E[|\sigma_{Blocked} S|]$$

This problem is particularly relevant in epidemiology, cybersecurity, and critical infrastructure protection. Techniques include:

- High-degree heuristics: Target nodes with the most connections.
- Betweenness and centrality metrics: Focus on nodes that bridge communities (Borgatti, 2005).
- Spectral methods: Use eigenvalue analysis of adjacency or Laplacian matrices to identify critical spreaders (Liu, Slotine, & Barabási, 2011).

3.4 Submodularity and Algorithmic Guarantees

The submodular structure of these problems is critical. It allows the application of greedy algorithms with provable bounds:

- Greedy selection yields at least $(1-1/e)$ of the optimal influence spread (Kempe et al., 2003).
- CELF++ (Cost-Effective Lazy Forward) algorithm optimizes performance by reducing redundant computations (Goyal, Lu, & Lakshmanan, 2011).
- Real-world systems often require trade-offs between approximation quality and computational efficiency, especially for large-scale networks.

Optimization in diffusion processes is a rich area in operations research that blends combinatorics, probability, and graph theory. Whether the goal is to spread influence or suppress harmful contagion, these problems drive algorithmic innovation and have profound implications for domains such as marketing, public health, information security, and social behavior.

4. Simulation and Analysis Techniques

Accurately evaluating diffusion dynamics in complex networks often requires computational simulation, as analytical solutions become intractable for large, heterogeneous structures. In this section, we outline key methods and metrics used for simulating and analyzing diffusion processes, with a focus on estimating the influence spread, evaluating algorithmic strategies, and understanding the impact of network topology.

4.1 Monte Carlo Simulations

To estimate the expected spread $E[|\sigma(S)|]$ for a given seed set SSS, we use **Monte Carlo (MC) simulations**—a widely accepted technique for approximating stochastic outcomes.

The process involves:

- Repeating the diffusion process (e.g., under the ICM or LTM) over multiple random trials.
- Averaging the number of nodes activated across these runs to obtain an empirical estimate.

$$E[|\sigma S|] \approx \frac{1}{N}\sum_{i=1}^{N}|\sigma_i(S)|$$

Where N is the number of simulation runs, and $\sigma_i(S)$ is the spread in the i-th trial.

Kempe et al. (2003) demonstrated that accurate estimation requires a large number of trials, especially in networks with high variance in node degree or influence probability. More recent work (Tang et al., 2014) introduced reverse influence sampling to reduce the simulation cost while maintaining approximation guarantees.

4.2 Algorithm Evaluation: Greedy and Beyond

The performance of influence maximization algorithms—greedy, CELF++, TIM/TIM+ (Tang et al.), IMM (Tang et al., 2015)—is evaluated on two fronts:

- **Approximation quality**: Comparison to the optimal influence spread (approximated via exhaustive simulation).
- **Computational efficiency**: Measured by time complexity and scalability on networks with millions of nodes.

Heuristic methods (e.g., degree discount, PageRank) are also tested for their trade-off between speed and accuracy.

4.3 Metrics for Diffusion Analysis

Beyond spread size, several other metrics provide insights into diffusion behavior:

- **Diffusion Rate**: The number of time steps required to reach a specified proportion of active nodes. Faster diffusion is critical in scenarios like viral marketing or epidemic outbreaks.
- **Saturation Level**: The final proportion of nodes activated. This reflects the maximum reach of a diffusion process and is influenced by network structure and seeding strategy.
- **Activation Curve**: Plots the number of active nodes vs. time steps. These curves help identify tipping points or plateaus in diffusion.
- **Robustness to Perturbations**: Simulations can model random or targeted node/edge removals to assess how robust a diffusion process is to structural changes—relevant for network resilience analysis (Albert et al., 2000).

4.4 Network Structure and Its Influence

Network topology plays a critical role in diffusion:

- **Scale-free networks** (Barabási & Albert, 1999) exhibit hub-dominated spread, where targeting a few high-degree nodes can trigger widespread diffusion.

- **Small-world networks** (Watts & Strogatz, 1998) allow rapid spread through short paths.
- **Modular or community-structured networks** may require multi-seed strategies to bridge structural bottlenecks.

Analyzing how different models perform across these topologies helps generalize results and guide practical deployment.

4.5 Visualization Tools

Effective diffusion analysis also involves visual tools:

- **Cascade graphs** showing activation sequences.
- **Heat maps** representing spread intensity over time.
- **Temporal snapshots** of node states during simulations.

These visualizations aid in interpreting diffusion dynamics, especially for non-technical stakeholders in policy, marketing, or public health.

5. Applications and Case Studies

Diffusion processes modeled over networks are not just theoretical constructs—they are central to solving real-world problems across industries and disciplines. Below are key application areas where mathematical diffusion models inform strategy, policy, and system design.

5.1 Marketing: Viral Campaigns and Influence Propagation

In marketing, the goal is often to maximize product adoption or brand awareness with minimal advertising spend. Social networks serve as channels where consumers influence one another, making them fertile ground for viral marketing strategies.

- Use of Models: Both the Independent Cascade Model (ICM) and Linear Threshold Model (LTM) are used to model how consumers adopt new products due to peer influence or exposure. Marketers identify "influencers"—nodes with high degree centrality or betweenness—to serve as initial adopters.
- Optimization Goal: Influence maximization (Kempe et al., 2003) helps companies choose a small seed set of individuals (e.g., early adopters, bloggers) to trigger wide adoption through word-of-mouth.
- Real-World Case: Online platforms like Facebook, Instagram, and Twitter use these models to optimize advertisement placements and referral strategies. Companies like Amazon and Netflix use diffusion-based algorithms to model recommendations and user influence.

5.2 Epidemiology: Modeling Disease Spread on Contact Networks

Epidemiological models, particularly in public health and medicine, use diffusion processes to predict and contain disease outbreaks.

- Use of Models: Variants of the SIR (Susceptible-Infected-Recovered) model are mapped onto network structures where individuals (nodes) interact (edges). The ICM aligns with infection probabilities, and threshold models can capture behavioral responses (e.g., mask adoption, vaccination decisions).

- Optimization Goal: Identify key nodes to vaccinate or isolate to minimize disease spread—an immunization problem that mirrors the inverse of influence maximization.

- Real-World Case: During the COVID-19 pandemic, contact tracing and targeted quarantine were implemented based on network-level analyses. Simulation tools like GLEaM and EpiModel rely on diffusion theory to model and forecast infection rates.

5.3 Infrastructure: Cascading Failures in Power Grids and Transport Systems

In physical infrastructure networks, failure of one component can trigger failures in others—a classic case of cascading diffusion.

- Use of Models: Threshold models capture the tipping point behavior seen in power grids, where a node (e.g., substation) fails once its load exceeds capacity due to failures in neighboring nodes. These models are also used to simulate congestion in transport networks.

- Optimization Goal: Enhance resilience by identifying weak points or high-impact nodes and designing redundancies to prevent cascades.

- Real-World Case: The 2003 Northeast blackout in the U.S. is a prime example of a cascading failure, where the shutdown of a single power line led to a chain reaction. Diffusion models are now embedded in grid simulation tools to prevent such occurrences.

5.4 Cybersecurity: Malware Spread and Containment Strategies

In cybersecurity, computer viruses, ransomware, and worms spread across digital networks much like diseases in human populations.

- Use of Models: The ICM and variations of the SIR model are used to simulate the propagation of malicious software. Each node represents a device or user, and edges reflect communication channels or software dependencies.

- Optimization Goal: Identify critical systems to patch or monitor to contain malware early. This includes identifying hubs in software dependency graphs or social engineering targets in email networks.

- Real-World Case: The 2017 WannaCry ransomware spread globally by exploiting vulnerabilities in unpatched Windows systems. Organizations now simulate such spread using diffusion models to harden their networks preemptively.

6. Discussion and Future Directions

The modeling of diffusion processes in networked systems continues to evolve, driven by the expanding complexity of real-world applications and the rapid growth of available data. While foundational models such as the Linear Threshold Model (LTM) and Independent Cascade Model (ICM), as presented by Kempe, Kleinberg, and Tardos (2003), have shaped the influence maximization problem in social networks, the assumptions behind these discrete-time frameworks are increasingly being challenged by dynamic, real-time scenarios. Continuous-time diffusion models (Newman, 2010; Jackson, 2008) have emerged as a powerful alternative, offering more realistic representations of asynchronous and heterogeneous interactions, especially in domains like epidemiology, information spread, and infrastructure management. The next frontier in diffusion research lies in the integration of strategic behavior and game theory, where multiple agents operate with conflicting objectives—spreading their own influence while limiting that of competitors. This is particularly relevant in political campaigns, marketing rivalries, and adversarial cybersecurity. Easley and Kleinberg (2010) and Rogers (2003) laid the theoretical groundwork for analyzing strategic decisions in networked environments. These competitive settings call for equilibrium-based analysis, enabling prediction and control under scenarios where agents respond rationally to each other's actions.

Another major shift is the movement toward machine learning-based approaches to model and predict diffusion. Rather than relying purely on static probability models, data-driven techniques now leverage large-scale social, behavioral, and transactional datasets to learn influence parameters, infer latent network features, and adapt to changing environments. Domingos and Richardson (2001) were among the first to mine influence directly from customer data. This line of work has since expanded to include graph neural networks, reinforcement learning, and streaming algorithms (Goyal, Lu, & Lakshmanan, 2011; Liu & Zhang, 2020), enabling real-time decision-making in evolving social and communication networks. These learning-based models complement and sometimes outperform classical approaches, especially when diffusion behavior is nonstationary or influenced by contextual variables. Realistic diffusion modeling also demands attention to behavioral complexity and heterogeneity, beyond simple activation rules. Threshold distributions, social reinforcement, and contagion types vary significantly across domains. Granovetter's (1978) threshold model of collective behavior and Centola's (2010) experiments on behavior spread in online networks demonstrated the critical role of network structure and peer influence strength. Valente (1996) further highlighted how individual susceptibility differs based on social position and prior exposure, advocating for nuanced models that reflect real adoption pathways. Meanwhile, Borgatti (2005) emphasized how structural centrality affects flow and control, underscoring the need for combining behavioral dynamics with topological analysis.

As networks become more intertwined and complex, resilience and robustness have also come into focus. Whether it is a power grid, a transportation network, or a supply chain, a failure in one part can cascade throughout the system. Models of cascading failures (Albert, Jeong, & Barabási, 2000; Barabási & Albert, 1999) have shown that scale-free networks are robust to random attacks but vulnerable to targeted ones. This insight is especially critical in cybersecurity, where malware can spread along communication or software dependency graphs. Liu, Slotine, and Barabási (2011) advanced this work by addressing controllability—the ability to steer or stabilize a network's dynamics—offering strategies to intervene efficiently in critical nodes or clusters. In public health, the diffusion of disease and information often overlaps, requiring hybrid models. The classical SIR model developed by Kermack and McKendrick (1927) remains a cornerstone in epidemiological modeling, but its integration with network structure is key to modern outbreak response strategies. Pastor-Satorras et al. (2015) reviewed how complex networks alter the thresholds and reach of epidemic processes. Today's models include individual behavior, information feedback, and targeted immunization—making them both more realistic and operational. Banerjee et al. (2013), in their study on microfinance diffusion, revealed how structural positions in the network influence access to opportunity, adding a layer of social justice to diffusion analysis.

Applications in marketing, cybersecurity, and infrastructure have also spurred research into scalable algorithms and community-aware strategies. Wang et al. (2010) and Chen et al. (2009) proposed greedy and community-based heuristics to identify influential nodes quickly in large-scale social networks. While exact influence computation is computationally hard, these approximations—especially when guided by submodular properties—offer practical tools for organizations aiming to drive adoption or mitigate threats under constraints. In the coming years, we expect hybrid models—combining real-time data, behavioral heterogeneity, structural dynamics, and strategic interactions—to dominate diffusion research. Ethical considerations will also gain prominence, especially in areas like misinformation suppression, targeted advertising, and algorithmic fairness. As models increasingly guide public policy and platform decisions, transparency and accountability will become non-negotiable.

7. Conclusion

Understanding and optimizing diffusion processes in networked systems is a critical challenge in today's interconnected world. This study presented a comprehensive framework grounded in operations research to model, simulate, and analyze the spread of influence, information, disease, and failures across complex networks. By examining the Linear Threshold and Independent Cascade models, we highlighted both deterministic and stochastic mechanisms of diffusion. The influence maximization problem was formalized as a submodular optimization task, and we explored efficient algorithmic strategies supported by Monte Carlo simulations and structural analysis. Real-world applications—including marketing campaigns, epidemic control, infrastructure resilience, and cybersecurity—demonstrate the versatility and impact of these models. Looking ahead, integration with continuous-time processes, strategic multi-agent interactions, and machine learning promises to enhance the realism and responsiveness of diffusion models. As networks evolve and grow in complexity, diffusion modeling will play an increasingly central role in data-driven decision-making, risk mitigation, and policy design. This work provides a solid foundation for future research and application, ensuring that diffusion insights remain relevant, actionable, and ethically aligned in both digital and physical systems.

References

1. Kempe, D., Kleinberg, J., & Tardos, É. (2003). Maximizing the spread of influence through a social network. *KDD '03*.
2. Rogers, E. M. (2003). *Diffusion of Innovations*. Free Press.
3. Easley, D., & Kleinberg, J. (2010). *Networks, Crowds, and Markets*. Cambridge University Press.
4. Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford University Press.
5. Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.
6. Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87(3), 925–979.
7. Liu, Y., Slotine, J. J., & Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, 473(7346), 167–173.
8. Goyal, A., Lu, W., & Lakshmanan, L. V. S. (2011). CELF++: Optimizing the Greedy Algorithm for Influence Maximization in Social Networks. *WWW '11*.
9. Chen, W., Wang, Y., & Yang, S. (2009). Efficient influence maximization in social networks. *KDD '09*.
10. Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6), 1420–1443.
11. Domingos, P., & Richardson, M. (2001). Mining the network value of customers. *KDD '01*.
12. Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144).
13. Valente, T. W. (1996). Social network thresholds in the diffusion of innovations. *Social Networks*, 18(1), 69–89.
14. Centola, D. (2010). The Spread of Behavior in an Online Social Network Experiment. *Science*, 329(5996), 1194–1197.
15. Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55–71.
16. Kermack, W. O., & McKendrick, A. G. (1927). A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A*.
17. Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
18. Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9), 5766–5771.
19. Liu, B., & Zhang, L. (2020). Real-time influence maximization on dynamic social streams. *Information Sciences*, 512, 902–917.
20. Wang, Y., Cong, G., Song, G., & Xie, K. (2010). Community-based greedy algorithm for mining top-k influential nodes in mobile social networks. *CIKM '10*.