Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/69/20241525

Integrating threshold models with Markov chains for information diffusion in social networks

Zhichao Liang

Centre in Computational Technologies for Finance, Nanyang Technological University, 50 Nanyang Ave, Singapore

zliang016@e.ntu.edu.sg

Abstract. In this paper, I introduce an agent-based approach that leverages Markov models to formalize and understand information propagation within social networks. This methodology aims to explain the internal information states of agents and employs probabilistic model checking to analyze these models. Initially, I provide a comprehensive overview of the current research on information diffusion in social networks. Following this, I delve into the fundamentals of model checking and illustrate how this technique can be used to assess model accuracy, particularly in managing large networks with high precision. To demonstrate the effectiveness of this approach, I conduct a series of experiments. The results show that this paper provides a robust and effective method for analyzing complex information diffusion and network behavior.

Keywords: Agent-based approach, Information diffusion, Threshold model, Monte Carlo simulation, Model checking

1. Introduction

With the widespread growth of social networks, the landscape of information dissemination has undergone a profound shift, affecting fields such as politics and business [1]. Social network analysis, a potent technique, investigates the connections and interactions between network nodes, enhancing our understanding of social dynamics [2]. These connections can manifest as approvals, conflicts, or exchanges of services.

Social networks are composed of nodes interconnected by various relationships [3]. Social network analysis explores these connections to understand their impact on information flow. Information dissemination within social networks can be represented as a graph, where nodes signify agents and edges represent communication links [4]. This study focuses on how an agent's internal state influences its decision to share information, using agent-based models for this purpose.

Markov chains are particularly suitable for modeling real-world social networks [5]. These stochastic models predict future states based on the current state alone, disregarding past states, and effectively simulate the spread of information among agents [6].

Contribution. I introduce a framework for formalizing information diffusion that accounts for an agent's internal state using Markov chains, with a particular emphasis on the broadcasting of opinions. Model-checkers and Monte Carlo simulation serve not only as simulation tools but also as powerful instruments capable of exhaustively exploring all potential system states. This exhaustive exploration allows for highly precise results.

2. Background

There are several models for disseminating information through influence, but finding the best one is still challenging for us [7]. Social influence models that can help us define the spreading process can be roughly divided into two categories: threshold models and infection models. The Simmelian model may not be considered [8]. In this paper, we focus on the threshold model, where an agent is affected when the number of its affected neighbors surpasses a specific threshold [9].

2.1. Threshold models

Threshold influence models determine if a node will become influenced based on the degree or set of neighbors of that node. For a node x, let n(x) denote the group of agents directly linked to x. These models establish a threshold q. The node x will be influenced if the size of n(x) is greater than or equal to q.

2.2. Discrete-time Markov chains

Markov chains provide a sophisticated formalization of probabilistic processes that move between states. A discrete-time Markov chain (DTMC) adopts a discrete-time model, where state transitions occur at defined time points. We express it mathematically:

A DTMC is defined as a tuple $D = (S, s_0, P, L)$. S: a finite set of states, s_0: distinguished initial state, $P: S \times S \rightarrow [0, 1]$ is a transition probability matrix and $L(s) \subseteq AP$ is a labeling function [10].

A DTMC represents a collection of execution paths through S, with P providing the probability of transitioning from one state to another, and L specifying the propositions that hold true in each state.

Another important concept is the reward function. We can define a reward function, $\rho: S \rightarrow R$, which assigns a reward value to each state in S. Using a reward function can help us to analyze the expected reward at a specific time step t. This is especially useful in studying information diffusion, where the reward can be modeled as the number of agents who have embraced a specific idea in a given state.

2.3. PRISM probabilistic model checker

PRISM is a probabilistic symbolic model checker. It verifies system models, provided as probabilistic automata, against properties written in its own specification language, incorporating several probabilistic logics [11]. Our models employ Discrete-Time Markov Chains (DTMC) with a reward function, functioning as probabilistic automata.

2.4. Monte Carlo simulation

Monte Carlo simulation utilizes random sampling and statistical modeling to to simulate intricate systems. When applied to a Discrete-Time Markov Chain (DTMC), each run starts from the initial state and progresses based on transition probabilities. Model checkers like PRISM can also be used for such simulations.

3. A threshold model based on Markov chains for information diffusion in social networks

I propose a threshold model for information diffusion in social networks, represented by the parallel combination of derived DTMCs. These DTMCs are produced from a social network graph, the agents' initial states, and two probabilities. Each agent can be in one of three states: agree, disagree, or indifferent. With n agents, the global model has 3^n states.

Each agent can either support or not support φ. Supporting agents broadcast to neighbors, but this isn't a network transition. Instead, a joint "think" transition updates support. The likelihood of adopting φ depends on the proportion of an agent's connections that support φ, scaled by a factor λ.

Let n_{ci} represent the number of connections. Additionally, let n_{ci}^{φ} represent the number of connections of agent a_i who are broadcasting messages in endorsing of the idea φ.

4. Model-Checking Threshold Models

In Figure 1, we examined a fully connected network (FCN) comprising 10 agents. In this scenario, instead of expecting all agents to eventually express φ, we anticipated reaching a state where half of the agents express φ. Then, I turn to randomly generated model in Figure 2.

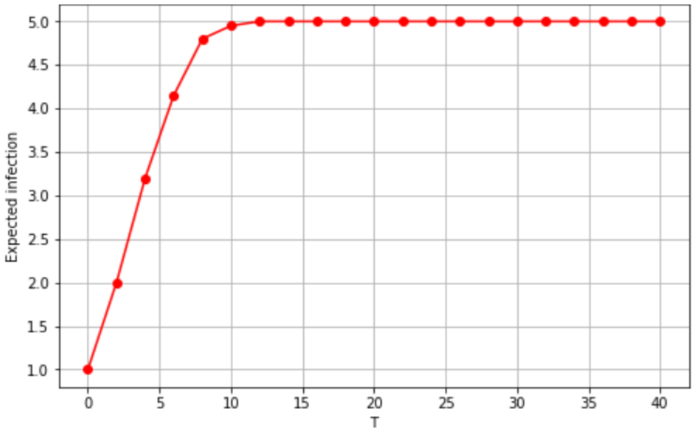

Figure 1. Anticipated number of agents adopting φ after T transitions in a FCN

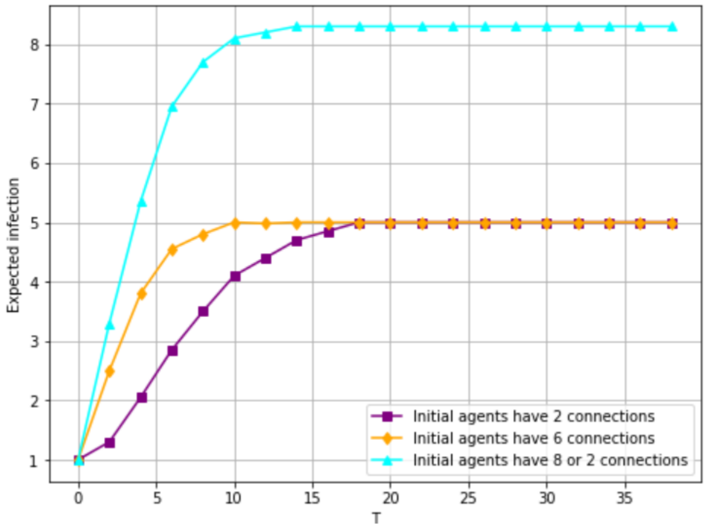

Figure 2. Expected number of agents adopting φ after T transitions in a randomly generated network.

For randomly generated models, we expect the network to converge to states where more than 8 agents express φ.

5. A Tool for Simulating Threshold Models Using Monte Carlo Methods

I designed a simulation tool for modeling influence using Markov chains. It takes a graph, two seed nodes for φ and ¬φ, the probability of changing minds, and the number of message exchanges. Here is the algorithm. It runs the simulation and calculates how many agents agree with φ. It is just a basic threshold model. I add probabilistic models and network change models to this basis.

```
Function REWARD (idea, anti-idea, graph, t)

    edges = edges(graph)
    threshold = t
    idea.state = agree
```

```
    anti-idea.state = disagree
    for node in nodes do
        agree = 0; disagree = 0
        for x in neighbor of node do
            if x.state == agree then
                agree += 1
            else
                disagree += 1
            end if
        end for
           if agree > disagree and agree > threshold then
               node.state = agree
           else if disagree > agree and disagree > threshold then
               node.state = disagree
           end if
    end for
    reward = 0
    for n in vertices (graph) do
        if n.state == agree then
            reward = reward + 1
        end if
    end for
    return reward

end funtion
```

I study information diffusion in social networks ranging from 100 to 1000 agents. The final experimental results are in line with expectations. Due to the addition of the probability of network changes, some agents will be ignored, so the percentage of agreement will be higher than the probability model. And because this experiment also has the parameter of probability, the result is still lower than the basic threshold model. Result shows in Figure 3.

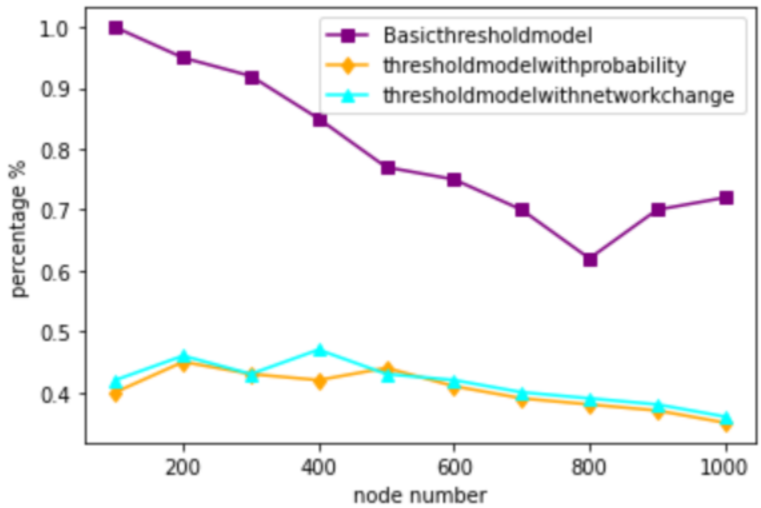

Figure 3. Experimental results of changing the number of agents

6. Conclusion

I integrated a threshold model with a Markov chain framework to study information diffusion in social networks. This combined model effectively captures the complexities of information spread and network

DOI: 10.54254/2755-2721/69/20241525

behavior evolution. Through mathematical simulations and scenario analyses, I demonstrated that Markov chain models provide robust and accurate predictions of information diffusion. My findings highlight critical factors such as thresholds and network structures that influence the diffusion process. The model's high predictive accuracy in real-world scenarios underscores its practical applications, offering a more precise and insightful understanding of information diffusion compared to previous models.

References

[1] Wihbey, J. (2015). How does social media use influence political participation and civic engagement? A meta-analysis. Journalist's Resource.
[2] Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications.
[3] Scott, J. (2024). Social Network Analysis: Research Methods. Bloomsbury Publishing.
[4] Newman, M. E. (2003). The structure and function of complex networks. SIAM review, 45(2), 167-256.
[5] Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. Social networks, 32(1), 44-60.
[6] Norris, J. R. (1998). Markov chains (No. 2). Cambridge university press.
[7] Christoff, Z., Hansen, J. U., & Proietti, C. (2016). Reflecting on social influence in networks. Journal of Logic, Language and Information, 25, 299-333.
[8] Pedersen, T., & Slavkovik, M. (2017). Formal models of conflicting social influence. In PRIMA 2017: Principles and Practice of Multi-Agent Systems: 20th International Conference, Nice, France, October 30–November 3, 2017, Proceedings 20 (pp. 349-365). Springer International Publishing.
[9] Dodds, P., & Watts, D. J. (2011). Threshold models of social influence.
[10] Kwiatkowska, M., Norman, G., & Parker, D. (2007). Stochastic model checking. Formal Methods for Performance Evaluation: 7th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2007, Bertinoro, Italy, May 28-June 2, 2007, Advanced Lectures 7, 220-270.
[11] Kwiatkowska, M., Norman, G., & Parker, D. (2011). PRISM 4.0: Verification of probabilistic real-time systems. In Computer Aided Verification: 23rd International Conference, CAV 2011, Snowbird, UT, USA, July 14-20, 2011. Proceedings 23 (pp. 585-591). Springer Berlin Heidelberg.