Agent-based Influence Propagation in Social Networks

Weihua Li
Auckland University of Technology
Email: weihua.li@aut.ac.nz

Quan Bai
Auckland University of Technology
Email: quan.bai@aut.ac.nz

Minjie Zhang
University of Wollongong
Email: minjie@uow.edu.au

Abstract—Many studies have been conducted to analyse the influence diffusion in social networks, where the problem is normally modelled from a centralised perspective. In this paper, we propose an agent-based approach to model the bidirectional influence propagation in directed weighted networks in a distributed manner. In the proposed model, each user's personalised traits and individual's social context are taken into consideration on the basis of social theories. Moreover, the model can capture the dynamics in the environment as well. We claim that the proposed approach is more suitable for simulating the real-world complex influence propagation and predicting the evolution of entire network comparing with traditional approaches.

Keywords-Agent-based influence; agent-based model; influence propagation; influence maximization; social network

I. Introduction

Nowadays, social networks have become an essential medium for people to communicate, interact and deliver messages. Basically, these interactions among the users in social networks form the primary influence-diffusion pathway [1]. Influence propagation modelling has been extensively studied and widely applied in many research fields, such as maximization of product adoption, contagion of computer viruses, spread of infectious diseases.

Most contemporary research work have modelled influence propagation as a centralised diffusion process [2] [3], where influence spread is represented as hops from one node to another, and eventually stops when a certain criteria is met. However, the state-of-the-art influence diffusion models, such as Independent Cascade (IC) model, Linear Threshold (LT) model, are not practical when the network topology is undiscovered or the global view is unavailable. Furthermore, centralised influence propagation models cannot handle opinion revisions of each individual, specifically, a vast of research works focus on predicting the status of network at t_{i+1} based on the given snapshot of time t_i . The result is most likely an interim stage of the network, whereas, the final stabilised status, i.e., the convergent state of the network, is not analysed. To explain this further, the principle of social influence suggests that the more that people interact with one another, the more similar they become [4]. The interactions can be regarded as a process of influence and opinions adaptation, while, the homophily will be achieved eventually. This also implies the process of influence will inevitably lead global network convergence [4]. On the other side, there can be thousands of individualised behavioural parameters when analysing influence in a complex social network, moreover, these parameters are globally independent, thus, it is nearly impossible to build any appropriate centralised model for influence simulations. By considering the aforementioned factors, the Agent-Based Modelling (ABM) has demonstrated many advantages in modelling complex systems [5] [6]. ABM is a specific individual-based computational model, which models entities in a system as a number of interactive autonomous agents. Compared with centralised models, agent-based model is more suitable for exploring the macro world through defining micro level of a social system [5] [6].

In this paper, we propose an Agent-based Influence Diffusion Model (AIDM), where the influence propagation demonstrates an evolutionary process. The model is applicable in a dynamic environment and functions even without the network topology. Moreover, AIDM is capable of capturing the prospective state of the network and analysing trends in the long run. In the proposed model, the influence propagation network is modelled as a multiagent system containing a collection of autonomous agents. In addition, users' behaviours, such as estimating level of prior commitment and social pressure, revising preference state, making decisions, are also taken into consideration. Furthermore, we evaluate the proposed model by using a typical application, i.e., influence maximization problem [3]. A novel seed selection approach based on the AIDM has been proposed, namely, the Evolution Based Backward (EBB) selection. The EBB approach utilises advantages offered from AIDM, and can produce better performance than state-of-the-art approaches in terms of efficiency and effectiveness.

The reminder of this paper is organised as follows. Section 2 elaborates agent-based influence diffusion model. Experiments and analysis are presented in Section 3. In Section 4, some related works have been reviewed. Finally, the paper is concluded in Section 5.

II. AGENT-BASED INFLUENCE DIFFUSION MODEL

In ABM, individuals or collective entities are represented as autonomous and self-directed homogeneous agents which can function independently in its local environment by



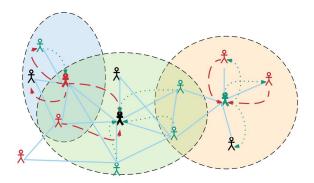


Figure 1. Agent-based Influence Diffusion Model

interacting with the direct neighbours [6]. In AIDM, users are modelled as agents, living in a dynamic local social environment, i.e., a directed weighted influence network. Each agent has a single preference state, i.e., positively activated (PA), negatively activated (NA) or inactivated (IA). Meanwhile, it possesses two fundamental features in terms of influence, i.e., individual's valuation of the product [7] and social conformity [8]. Specifically, an agent not only needs to evaluate its intrinsic appreciation towards a specific item, but is also required to handle the social pressure coming from the neighbours with diverse opinions. Therefore, in the proposed model, agents are capable of adapting their preference states and behaviours based on the internal factor: own opinion and attitude towards an item or event, as well as the external factor: social context. Activated agents diffuse influence by interacting with the neighbours. In the meanwhile, they are getting influenced and attempting to reach a "comfortable" state.

Figure 1 illustrates the general idea of agent-based influence propagation. In this figure, an agent is of different colours: red, blue and black, which denote NA, PA and IA respectively. Every agent locates in the centre of its local social context. An agent's local view covers the immediate neighbours only, thus, the interactions among its peers are not in the scope. We see that three individuals' social contexts presented. PA and NA agents receive influences from the neighbours with adverse opinion to suggest them revise the state. Meanwhile, the neighbours with same state also convey the supportive influence to retain the current state. IA agents are affected by both NA and PA, but it does not exert any influence on their peers.

From macroscopic perspective, ABM demonstrates a decentralised evolutionary pattern driven by the individual's actions. The entire network evolves through a number of discrete time steps according to a set of transitional or behavioural rules of each agent. To be more specific, the network in time t_i is regarded as generation g_i , agents keep adapting and evolving, the network becomes generation g_{i+1} in time step t_{i+1} .

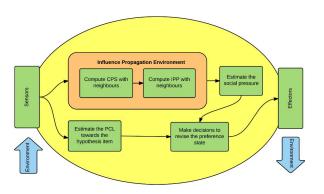


Figure 2. Individual Agent Architecture in AIDM

Inspired by the agent-based model proposed in [9], the individual agent architecture has been described in Figure 2. The agent attempts to construct its local influence environment by considering common preference similarity with the neighbours. Next, the agent analyses prior commitment level towards the hypothesis item and estimate the undertaken social pressure. Subsequently, it decides how to update the preference state. Concrete design and formal definitions are introduced in the following sub-sections.

A. Agent and Environment

An agent is defined as a vertex $v_i, (v_i \in V)$ in a directed weighed social network G = (V, E), where $V = \{v_1, ..., v_n\}$ denotes a set of agents and E represents a set of edges, $E = \{e_{ij} | 1 \leq i, j \leq n\}, i, j \in \mathbb{N}, \{v_i, v_j\} \subseteq V$. The weight (strength) of edge e_{ij} denotes the influence propagation probability from v_i to v_j . Agent v_i has a set of neighbours $\Gamma(v_i)$. If v_j is a neighbour of v_i , then $\{e_{ij}\} \cup \{e_{ji}\} \subseteq E, v_j \in \Gamma(v_i)$. |V| and |E| denote the amount of agents and edges respectively.

Another type of vertex I also exists in same context, where I denotes the item set, $I = \{i_1, ..., i_m\}$. The agent v_j maintains its own rating list R_j . The preference of v_j towards item i_x derives from its ratings to items $\{r_{jx}|1 \leq j, x \leq n\}$, $i, j \in \mathbb{N}, r_{jx} \in R_j$, where r_{jx} signifies the rating score of item i_x given by agent v_j .

Meanwhile, agent v_j has a unique preference state towards item i_x , i.e., $s_{jx}, s_{jx} \in \{PA, NA, IA\}$. $s_{jx} = PA$ implies that v_j shows a favour towards item i_x and tends to diffuse positive influence to its neighbours $\Gamma(v_j)$, enhancing agents with the same opinion and attempting to change those with opposite opinions. On the contrary, $s_{jx} = NA$ indicates that v_j expresses disfavour towards item i_x . While, IA indicates neutral opinion.

Environment refers to the local influence propagation context where a particular agent resides in, incorporating the neighbourhood and corresponding relationships. In our previous work [10], we considered both Common Prefer-

ence Similarity (CPS) and friendship affiliation as essential factors of constructing influence propagation channels, where cps_{ij} denotes the degree of CPS between v_i and v_j . Whereas, the Influence Propagation Probability (IPP) ipp_{ij} refers to the chances that agent v_i propagates influence to v_j effectively, where v_i is the source of influence and v_j is the target, $ipp_{ij} \neq ipp_{ji}$. In the meanwhile, it also represents the weight (strength) of an edge. In this context, the more ratings an agent gives, the more powerful influence it propagates [11]. Hence, IPP is formulated as Equation 1, where $|I_i|$ denotes the number of items rated by agent v_i .

$$ipp_{ij} = cps_{ij} \cdot \frac{|I_i|}{|I_i \cup I_j|} \tag{1}$$

B. Agent Social Behaviours

Prior Commitment Level (PCL) pcl_{jx} is defined as agent v_j 's estimated prior preference state or opinion towards a hypothesis or rated item i_x on the basis of the past ratings or experience. To be more specific, if an estimated or actual rating on item i_x approaches to the highest rating score, v_j has higher tendency of being PA, vice versa. The pcl_{jx} of turning positive could be formulated as Equation 2.

$$pcl_{jx} = \begin{cases} \frac{r_{jx} - min(R_j)}{max(R_j) - min(R_j)}, & max(R_j) \neq min(R_j) \\ 0.5, & max(R_j) = min(R_j) \end{cases}$$
(2)

In Equation 2, $max(R_j) - min(R_j)$ denotes the gap of highest and lowest rating values given by agent v_j . While, the $v_j's$ PCL of turning negative is represented as $1 - pcl_{jx}$; The PCL of retaining neutral opinion on i_x is depicted as $1 - |pcl_{jx} - 0.5|$.

Social Pressure $sp_{jx}|S$ is defined as the influence agent v_j received from its immediate neighbours $\Gamma(v_j)$, to change or stick on its opinion towards item i_x to one particular preference state S ($S \in \{PA, NA, IA\}$).

The value of social pressure is usually measured by examining the numbers of immediate neighbours with different preference states [12]. The more adverse neighbours with strong IPP agent v_j has, the higher social pressure v_j undertakes, so that v_j has higher possibility to revise the opinion. Whereas, neighbours with the same preference states contribute supportive influence, thus, the social pressure becomes the strength of enhancing the current preference state, which reduces the probability of state revision. The social pressure of agent v_j can be calculated by using Equation 3.

$$sp_{jx}|S = 1 - \prod_{v_i \in \Gamma(v_i)} \sum_{s_{i,x} = S} (1 - ipp_{ij})$$
 (3)

In Equation 3, S represents one particular preference state value of agent v_j towards item i_x . If $s_{jx} = S$, then $sp_{jx}|S$ refers to the supportive strength to encourage v_j retain

the current preference state. Otherwise, $sp_{jx}|S$ denotes the social pressure from the neighbours to force v_j to change its state from s_{jx} to S.

Probability of Revising Preference State $prs_{jx}(s'_{jx}|s_{jx})$ denotes agent v_j 's probability of tendency of changing its current preference state s_j towards an item i_x to another state s'_j , which could be opposite or the same as the current preference state on the basis of two factors, i.e., PCL pcl_{jx} and social pressure sp_{jx} . prs_{jx} is formulated as Equations 4, 5 and 6.

$$prs_{jx}(PA|s_{jx}) = \lambda_j \cdot pcl_{jx} + (1 - \lambda_j) \cdot \frac{sp_{jx}|PA}{\sum_{S \in \{PA, NA, IA\}} sp_{jx}|S}$$
(4)

$$prs_{jx}(NA|s_{jx}) = \lambda_j \cdot (1 - pcl_{jx})$$

$$+ (1 - \lambda_j) \cdot \frac{sp_{jx}|NA}{\sum_{S \in \{PA, NA, IA\}} sp_{jx}|S}$$
(5)

$$prs_{jx}(IA|s_{jx}) = \lambda_j \cdot (1 - |pcl_{jx} - 0.5|)$$

$$+ (1 - \lambda_j) \cdot \frac{sp_{jx}|IA}{\sum_{S \in \{PA, NA, IA\}} sp_{jx}|S}$$
(6)

In Equation 4, $prs_{jx}(PA|s_{jx})$ represents the probability of agent v_j to revise the preference state towards i_x from any state to PA. λ_j stands for the personalised parameter of v_j , which is also a trade-off between the PCL and social pressure. Similarly, Equations 5 and 6 formulate the probability of revising or retaining the current preference state s_{jx} as NA and IA respectively.

The opinion revision behaviour of an agent is triggered by the update of the neighbour's opinion or the changes of the ratings. Once notified, the agents start the actions.

C. Evolution Based Backward Algorithm

In this research, we propose a novel influence maximization algorithm by utilising the advantages offered from AIDM, namely, the EBB algorithm. It aims to increase the positive influence by reducing the maximum prospective negative impact. EBB does not focus on the current state of network, but attempts to select seeds by examining the final status of network. Specifically, when the network starts to converge, the negative node with the maximum influence capability is selected as seed, so that the most negative influential node is converted and starts to convey positive influence to the neighbours. EBB is described in Algorithm 1.

The input of Algorithm 1 incorporates the seed set size k and the global preference graph, where all the agents and their connections are included. While, the output is the selected seed set with k seeds: P_k . Lines 2 - 7 consist the main body of the algorithm. In each iteration, v_x is elected when the PA and NA nodes start to converge in the network.

Algorithm 1 Evolution Based Backward Algorithm

```
Input: G_p, k
Output: P_k

1: Initialize P := \varnothing
2: for i = 1 to k do
3: Network evolves
4: Find the NA user with maximum influence, v_x
5: v_x := \max_{v_x \in V \setminus P, s_{ix} = NA} \sum_{v_j \in \Gamma(v_x)} ipp_{x_j}
6: P_k := P_k \cup \{v_x\}
7: end for
```

III. EXPERIMENTS AND ANALYSIS

Two experiments have been conducted. The first experiment aims to track, forecast and analyse the evolution pattern of a social network by using AIDM, where different scenarios are demonstrated. The second experiment tends to compare the performance of AIDM under influence maximization problem. Specifically, influential users (seed set) identified using AIDM and EBB algorithm are compared with three classic seed selection approaches, i.e., greedy selection, degree-based selection and random selection.

Movielens¹ dataset is used for the experiments, which is a stable benchmark dataset containing 1,000,209 anonymous ratings of approximately 3,900 movies made by 6,040 MovieLens users who joined in 2000. We select 500 users randomly for the experiments in order to reduce the computing time. Individual's personalised parameter λ_j is randomized in our experiments.

In the current setting, two constrains have been declared for agents' actions. First, seed set can be selected from the agents with any initial preference state: PA, NA or IA. Once a user is selected as a member of the seed set, its preference state is locked as PA. Second, there are very low probabilities to revise the preference state if a user undertakes low social pressure generated from the adjacent neighbours with adverse attitudes.

Performance metrics are defined by considering the amount of both positive and negative nodes, which has been formulated as Equation 7, where |PA|, |NA| and |V| denote the count of PA, NA and all the agents in the network respectively. The award from positive nodes and penalization from the negative ones present in an asymmetric way by using a trade-off factor $\alpha, \alpha \in [0, 1]$.

$$f(PA, NA) = \alpha \cdot \frac{|PA|}{|V|} + (1 - \alpha) \cdot (1 - \frac{|NA|}{|V|})$$
 (7)

A. Experiment 1: Network Evolution Simulation

Experiment 1 aims to track, forecast and analyse network evolution by adopting AIDM. We define a fixed time step:

60 for the network evolution, so that the full evolutionary picture could be presented clearly.

In Figure 3, one of the most popular items (receive over 250 ratings from the 500 users) is selected, the initial state of the network shows that more than 150 users turn out to be PA, whereas, less than 25 users are NA. After five time steps, we could observe the global stability of social homogeneity, where both PA and NA nodes start to converge. Moreover, both lines walk steadily and only slight fluctuation is observed. The count of the positive rises up to approximate 425, while the negative ones' stays only around 75. Hence, we could conclude that the positive influence dominates through the network. Figure 4 demonstrates an evolution pattern of an unknown or unfamiliar item. The network evolves towards a undesirable direction, where the negative influence appears more dominant. Both types of nodes compete for the resources, IA nodes has been consumed thoroughly within 5 time steps; convergence tendency of the PA and NA emerges after 20 time steps. In this case, proper measures should be taken to remedy the inferior situation. Assume the company selects some users from the network to promote this item, and Figure 5 illustrates the trend after implementing the viral marketing strategy. The PA is far beyond the NA this time, and the convergence tends show after 15 time steps.

In summary, AIDM can be adopted to forecast and analyse the overall trend of network evolution, so that proper business strategy can be made. In the aforementioned two scenarios, no further actions are required for the first case since the network is approaching to a beneficial direction. However, more efforts or investment should be put for the second scenario.

B. Experiment 2: Influence Maximization by Using EBB

In this experiment, we compare the EBB algorithm with the other three seed selection algorithms, including greedy algorithm [3], degree-based selection and random selection.

Figure 6 demonstrates the comparison result of influence effectiveness with different seed numbers. X-axis denotes the seed set size, while y-axis represents the influence effectiveness (using Equation 7, $\alpha=0.8$). We could observe from this figure that greedy selection outperforms the other algorithms when the seed set size is small. While, EBB performs much better than the rest when the size of seed set reaches 30. Three ranking-based selections are not bad. As expected, random selection plays the worst. In Figure 7, efficiency among different seed selection approaches is measured. It is obvious that EBB performs as efficiently as the ranking based algorithms. By contrast, greedy selection is the worst due to its high computational complexity.

According to the experimental results, we can conclude that the EBB algorithm in general is more effective and efficient by comparing with other seed selection algorithms. Furthermore, the EBB algorithm is based on the AIDM,

¹https://grouplens.org/datasets/movielens

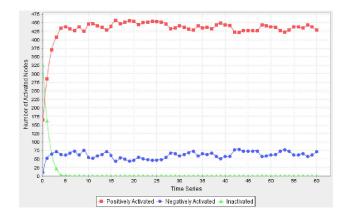


Figure 3. Network Evolution Trend of a Popular Item

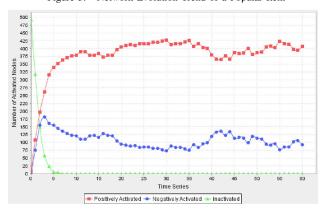


Figure 5. Network Evolution Trend after Implementing the Viral Marketing Strategy

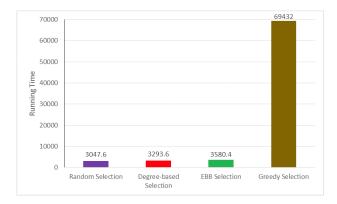


Figure 7. Seeds Selection Efficiency Comparison: EBB vs Ranking Based Methods

namely, many features are enabled by agent-based modelling. Hence, the result from Experiment 2 also proves the rationality of AIDM.

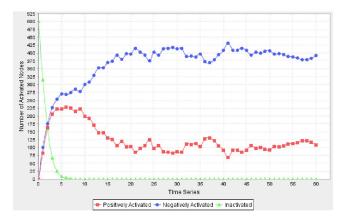


Figure 4. Network Evolution Trend of an Unknown Item

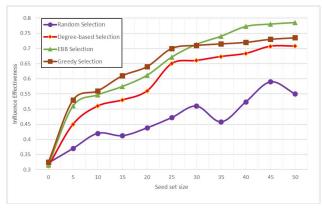


Figure 6. Seeds Selection Approaches Performance Comparison

IV. RELATED WORK

Kempe et al. research influence maximization on the basis of two fundamental stochastic influence diffusion models, i.e., the IC and LT Model [3]. Jiang et al. propose the preference-based trust independent cascade model, considering both user preference and trust connectivity in constructing influence propagation networks [10]. Liu et al. propose an ising model to predict positive and negative opinion formation in social network by considering the neighbourhood-based interactions [13]. Bhagat et al. study an interesting variation of the classical influence maximization problem by considering different states of customers [7]. Tang et al. formally defined the conformity influence and conducted contribution analysis of various social factors; their studies showed conformity plays a significant role in predicting influence acceptance [8].

In these existing research works, one of the critical assumptions is the user's preference state is static, it does not change once assigned. However, this may not hold in general since individuals may revise the state due to the changing environment. Influence propagation demonstrates

an evolutionary pattern of a network, rather than simply hopping. However, to the best of our knowledge, this feature is seldom systematically articulated when modelling influence propagation. Nearly all of the existing approaches model the diffusion process from a centralized point of view, where the network is static and the topological structure is available. Furthermore, the existing approaches cannot handle the dynamics of social networks or capture the individuals' complex behaviours.

ABM has a substantial and active literatures including both methodologies and applications [14] [15]. A simple and generic ABM is cellular automata (CA) [16], where each agent keeps evolving by looking at the neighbours' states. Sophisticated ABM sometimes utilises artificial intelligence approaches, such as neural networks, evolutionary algorithms, etc., which enable the agents to learn and adapt their behaviours [5] [9]. Li and Tang analyse the group polarization based on ABM and empirical data analysis, where a novel distributed voter model has been proposed. The social network demonstrates a convergent trend when considering three types of valence of the social identity tie [17].

V. CONCLUSION AND FUTURE WORK

In this research work, we have modelled the real-world influence propagation network as a complex system, where weighted bidirectional influence exists. We have proposed a distributed influence propagation model, i.e., AIDM, to simulate and forecast the trend of network evolutions driven by influence. In the proposed model, individual agent architecture is further elaborated. Individuals' personalised characteristics and social context have been modelled. Furthermore, behaviours including PCL estimation, opinion revision, decision making have been modelled. Experiments have been conducted to reveal the network evolution in different scenarios. We make use of influence maximization as a typical application of AIDM and explore a novel seed selection algorithm called EBB. The influence efficiency and effectiveness of EBB prove to be higher than stateof-the-art approaches. In the future, we will compare the computational complexity of AIDM with other traditional influence diffusion models, such as IC, TL, using larger datasets. We will further model the personalised parameter λ for each user by utilising machine learning algorithms.

REFERENCES

- [1] S. Aral, L. Muchnik, and A. Sundararajan, "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 51, pp. 21544–21549, 2009.
- [2] W. Chen, Y. Wang, and S. Yang, "Efficient influence maximization in social networks," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2009, pp. 199–208.

- [3] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proceed*ings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003, pp. 137–146.
- [4] D. Centola, J. C. Gonzalez-Avella, V. M. Eguiluz, and M. San Miguel, "Homophily, cultural drift, and the coevolution of cultural groups," *Journal of Conflict Resolution*, vol. 51, no. 6, pp. 905–929, 2007.
- [5] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the National Academy of Sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.
- [6] C. M. Macal and M. J. North, "Agent-based modeling and simulation," in *Proceedings of the Winter Simulation Conference*. IEEE, 2009, pp. 86–98.
- [7] S. Bhagat, A. Goyal, and L. V. Lakshmanan, "Maximizing product adoption in social networks," in *Proceedings of the* fifth ACM international conference on Web search and data mining. ACM, 2012, pp. 603–612.
- [8] J. Tang, S. Wu, and J. Sun, "Confluence: Conformity influence in large social networks," in *Proceedings of the* 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013, pp. 347–355.
- [9] F. Chantemargue, O. Krone, M. Schumacher, T. Dagaeff, and B. Hirsbrunner, "Autonomous agents: from concepts to implementation," in *Proceedings of the Fourteenth European Meeting on Cybernetics and Systems Research (EMCSR'98)*, vol. 2, 1998, pp. 731–736.
- [10] C. Jiang, W. Li, Q. Bai, and M. Zhang, "Preference aware influence maximization," in *Proceedings of International Joint Agents Workshop and Symposium (IJAWS)*. Springer, 2015.
- [11] W. Zhang, S. Wang, G. Han, Y. Yang, and Q. Wang, "Hiprank: Ranking nodes by influence propagation based on authority and hub," in *Smart Modeling and Simulation for Complex Systems*. Springer, 2015, pp. 1–14.
- [12] Z. Li and X. Tang, "Group polarization: connecting, influence and balance, a simulation study based on hopfield modeling," in *PRICAI 2012: Trends in Artificial Intelligence*. Springer, 2012, pp. 710–721.
- [13] S. Liu, L. Ying, and S. Shakkottai, "Influence maximization in social networks: An ising-model-based approach," in *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2010, pp. 570–576.
- [14] X. Su, M. Zhang, and Q. Bai, "Decentralised task allocation under space, time and communication constraints in disaster domains," in *Smart Modeling and Simulation for Complex Systems*. Springer, 2015, pp. 41–57.
- [15] Q. Bai, M. Zhang, and H. Zhang, "A colored petri net based strategy for multi-agent scheduling," in *Rational, Robust,* and Secure Negotiation Mechanisms in Multi-Agent Systems, 2005. IEEE, 2005, pp. 3–10.
- [16] D. Shiffman, S. Fry, and Z. Marsh, *The nature of code*. D. Shiffman, 2012, ch. 7 Cellular Automata, pp. 323–330.
- [17] Z. Li and X. Tang, "From global polarization to local social mechanisms: A study based on abm and empirical data analysis," in *Smart Modeling and Simulation for Complex Systems*. Springer, 2015, pp. 29–40.