

A Novel Discrete-Time Model of Information Diffusion on Social Networks Considering Users Behavior

1st Tran Van Khanh
*Faculty of Information Technology
Posts and Telecommunications
Institute of Technology
Hanoi, Vietnam
khanh080605@gmail.com*

2nd Do Xuan Cho
*Faculty of Information Security
Posts and Telecommunications
Institute of Technology
Hanoi, Vietnam
chodx@ptit.edu.vn*

3rd Hoang Phi Dung
*Faculty of Fundamental Sciences
Posts and Telecommunications
Institute of Technology
Hanoi, Vietnam
dunghp@ptit.edu.vn*

Abstract—In this paper, we introduce the SDIR (Susceptible–Delayable–Infected–Recovered) model, an extension of the classical SIR epidemic framework, to provide a more explicit characterization of user behavior in online social networks. The newly merged state D (delayable) represents users who have received the information but delayed its spreading and may eventually choose not to share it at all. Based on the mean-field approximation method, we derive the dynamical equations of the model and investigate its convergence and stability conditions. Under these conditions, we further propose a greedy algorithm and a sandwich approximation algorithm for the edge-deletion problem, aiming to minimize the influence of information diffusion by identifying approximate solutions.

Index Terms—SIR epidemic, social networks, complex networks, Markov chains, discrete optimization, edge deletion, mean-field approximation

I. INTRODUCTION

Online social networks have become one of the most crucial and essential information platforms for communication and commerce on a global scale. Because of their highly complex data structures, information spreading in social networks has emerged as an ideal environment for propagation—deep, wide, and significantly faster than any previous medium. Studying information diffusion on platforms such as Facebook, Twitter, and TikTok plays an essential role in communication media, information security, and the social sciences [8], [12], [26], [29]. In these platforms, the influence maximization (IM) problem [8], [12], [22], [29] has been recognized as a fundamental problem in viral marketing, while the influence minimization (IMIN) problem [13], [23], [29], [31] is central to controlling the spread of harmful or false information in online social networks.

Epidemic models have long been studied using mathematical formulations, for instance, the Susceptible-Infected-Recovered (SIR) model is used for epidemic forecasting in epidemiology [28]. Initially, epidemic models such as SIR and SIS (Susceptible-Infected-Susceptible) were applied to epidemiology for disease forecasting [4], [28]. Nowadays, however, these models are also widely applied to other fields,

including viral advertising [24], [26], cybersecurity [1], [3], and information diffusion [9], [16]. Models of information propagation are generally categorized into stochastic and deterministic approaches, with several comprehensive surveys available [19], [21]. For example, Wang et al. [32] introduced a discrete-time virus spread model and evaluated the epidemic threshold as a function of network structure. Miegheem et al. [16], [18] applied Markov chain formulations to analyze biological diffusion models such as SIR, and subsequent works explored SIR variants, including SIS and SIRS [5], [27]. Through mean-field approximation, stochastic SIR models can be transformed into deterministic and discrete forms, enabling linearization and tractable analysis [10], [16], [20], [34], [35]. More recently, Yi et al. [34] employed mean-field approximations to study discrete SIR models, reformulated them into matrix-based dynamical systems, and studied influence minimization via edge deletion, along with convergence properties of the resulting models.

At present, personalization of user experience has become a dominant trend and a primary objective of online social networks, especially with the increasing use of artificial intelligence algorithms [30] to enhance user engagement on platforms such as Facebook, Twitter, YouTube and TikTok. Consequently, the study of users' behavior has become particularly relevant in this context [35].

In this paper, we present a novel discrete-time SDIR model, a new extension of the classical discrete SIR framework. The SDIR model incorporates an intermediate state, D (Delayable), between S and I. This new state captures scenarios where a node that receives information does not immediately become “infected,” but may instead exhibit a delay before spreading the information—or may ultimately choose not to spread it at all. This behavioral shade reflects realistic interactions in online social networks, where users exercise discretion in processing, accepting, and sharing information. Each seed node corresponds to a social media account (e.g., a Facebook/Twitter/YouTube/TikTok user account or fan page) that initiates diffusion to its followers at the beginning. In reality,

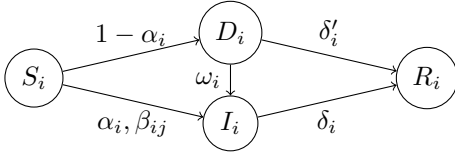


Fig. 1. The SDIR model

many users either ignore their received information or share it only after a delay in judgment. User procrastination on online social networks has been examined as an inherent behavioral factor in the use of such platforms [6], [25]. By introducing the Delayable state, the SDIR model offers a more accurate representation of individual-level behavior in information diffusion.

Our paper focus on discrete-time SDIR model by using the mean-field approximation method to transform the model SDIR to deterministic SDIR model. Using some techniques in linear algebra and spectral matrix theory, we study optimization problems that minimize the number of infections in SDIR Markov chain model on a network. By using some technique assumptions, we give the sufficient condition for the convergence and the stability of our model. Moreover, we investigate the change of the number of infections after deleting edges. We propose the an efficient modified greedy algorithm based on [34] and Sandwich algorithm based on Sandwich framework [15], [33] for minimize infections.

II. PROPOSED MODEL

The proposed SDIR model is extended from the SIR model to better suit the application of simulating information diffusion on social networks. A key characteristic of the SDIR model is the introduction of state D (Delayable) which means that a node can be delayed in its infection process. More specifically, consider a directed graph $G(V, E)$ with $|V| = n$, where each node represents a user and can be in one of four states: *Susceptible* (S), *Delayable* (D), *Infected* (I), or *Recovered* (R). Notably, Delayable means that a delayed state after being exposed to information but not immediately spreading it, and more importantly, it is still considered as an infected state. It is evident that $S_i(t) + I_i(t) + D_i(t) + R_i(t) = 1$. At time t , a node i in state S can be infected by an adjacent node j in state I that has an edge pointing to i with probability $\beta_{ij}(t)$. If successful, node i will transition to one of two states: state I with probability $\alpha_i(t)$ or state D with probability $1 - \alpha_i(t)$. Furthermore, if a node i is in state D at time t , it may transition to another state or remain in its current state according to the following rule: A real number $p \in [0, 1]$ is randomly chosen following a uniform distribution. Then, if $p < \omega_i(t)$, it transitions to state I; if $\omega_i(t) \leq p < \omega_i(t) + \delta'_i(t)$, it transitions to state R; otherwise, it remains in its current state D. When node i is infected at time t , it also heals with rate $\delta_i(t)$. Before establishing the equations for the model, we introduce the following assumptions.

Assumption 1. $\delta_i(t) \leq \delta'_i(t)$, $\forall i \in \overline{1, n}$, $t \geq 0$.

Assumption 2. $\alpha_i(t) \in (0, 1]$, $\forall i = \overline{1, n}$, $t \geq 0$.

The set of inequalities in Assumption 1 is derived from the observation that during the process of information diffusion and reception, users tend to ignore or quickly forget information (transition to state R) if they do not share or interact immediately.

Next, for every node $i \in 1, 2, \dots, n$, we have:

$$I_i(t+1) = S_i(t)\alpha_i(t) \left(1 - \prod_{j=1}^n (1 - \beta_{ij}(t)I_j(t)) \right) + \omega_i(t)D_i(t) + (1 - \delta_i(t))I_i(t) \quad (1)$$

$$D_i(t+1) = S_i(t)(1 - \alpha_i(t)) \left(1 - \prod_{j=1}^n (1 - \beta_{ij}(t)I_j(t)) \right) + (1 - \omega_i(t) - \delta'_i(t))D_i(t) \quad (2)$$

$$R_i(t+1) = \delta_i(t)I_i(t) + \delta'_i(t)D_i(t) + R_i(t) \quad (3)$$

By evaluating $S_i(t) \leq S_i(0)$, taking the expectation on both sides of each equation, and then linearizing the model using the mean-field approximation and the approximation formula $e^x \approx x + 1$ when $x \rightarrow 0$ for the quantity $(1 - \prod_{j=1}^n (1 - \beta_{ij}(t)I_j(t)))$ under the assumption that $\beta_{ij}(t)$ are independent for every pair i, j , and the coefficients $\beta_{ij}(t)$, $\delta_i(t)$, $\alpha_i(t)$, $\omega_i(t)$, $\delta'_i(t)$ are independent and identically distributed, we obtain the following system:

$$\mathbf{x}(t+1) = (\mathbf{I} - \mathbf{D} + \mathbf{A}\mathbf{S}(0)\mathbf{B})\mathbf{x}(t) + \mathbf{W}\mathbf{y}(t) \quad (4)$$

$$\mathbf{y}(t+1) = (\mathbf{I} - \mathbf{A})\mathbf{S}(0)\mathbf{B}\mathbf{x}(t) + (\mathbf{I} - \mathbf{W} - \mathbf{D}')\mathbf{y}(t) \quad (5)$$

$$\mathbf{r}(t+1) = \mathbf{D}\mathbf{x}(t) + \mathbf{D}'\mathbf{y}(t) + \mathbf{r}(t) \quad (6)$$

Here, $\mathbf{x}(t)$, $\mathbf{y}(t)$, $\mathbf{r}(t)$ are vectors whose i -th element takes the value $\mathbb{E}[I_i(t)]$, $\mathbb{E}[D_i(t)]$, $\mathbb{E}[R_i(t)]$, respectively. \mathbf{A} , \mathbf{W} , \mathbf{D} , \mathbf{D}' are diagonal matrices whose i -th diagonal element takes the value $\mathbb{E}[\alpha_i(t)]$, $\mathbb{E}[\omega_i(t)]$, $\mathbb{E}[\delta_i(t)]$, $\mathbb{E}[\delta'_i(t)]$, respectively. $\mathbf{S}(0)$ is a diagonal matrix whose i -th diagonal element is $1 - x_i(0) - y_i(0) - r_i(0)$. The matrix \mathbf{B} consists of elements $B_{ij} = \mathbb{E}[\beta_{ij}(t)]$. We also assume that $\sum_{j=1}^n B_{ij} < 1$, $\forall i = 1, 2, \dots, n$.

III. SUFFICIENT CONDITION FOR GLOBAL CONVERGENCE

One difference between the SDIR model and the SEIR model in [21] is that the delayed state in the SDIR model is still considered an infected state, while the SEIR model only considers E as an exposed state. Therefore, when considering global stability in the SDIR model, we need to find a sufficient condition for $\mathbf{x}(t)$ and $\mathbf{y}(t)$ to both approach the zero state. Consider a vector $\mathbf{q} \in (0, 1]^n$ which has n elements q_1, q_2, \dots, q_n , then define

$$\mathbf{C}(\mathbf{q}) := \text{diag}(\min(D_i, D'_i + (1 - 1/q_i)W_i)),$$

$$\mathbf{M}(\mathbf{q}) := \mathbf{I} - \mathbf{C}(\mathbf{q}) + (\mathbf{A} + \mathbf{Q}(\mathbf{I} - \mathbf{A}))\mathbf{S}(0)\mathbf{B},$$

where $Q = \text{diag}(q)$. Moreover, if there is no confusion, we convention to write $M(q)$ as M . Recall that the sufficient condition for the convergence of the SIR model is $\rho(M_{\text{SIR}}) < 1$, where $M_{\text{SIR}} = I - D + S(0)B$. Under Assumption 1, we can give a convergence condition that is better than the SIR model and is stated in the following theorem.

Theorem III.1. *Under Assumption 1, there always exists a choice of vector $q \in (0, 1]^n$ such that $\rho(M) \leq \rho(M_{\text{SIR}})$. Moreover, if $\rho(M) < 1$ then $x(t)$ and $y(t)$ converge to the $\mathbf{0}_{n \times 1}$.*

Corollary III.2. *If there exist a vector $q \in (0, 1]^n$ such that $\rho(M) < 1$ then $x(t)$ and $y(t)$ converge to the $\mathbf{0}_{n \times 1}$.*

Theorem III.1 provides a "loose" condition for the convergence of the model due to the presence of n parameters $q_i, i = \overline{1, n}$. The variation of these n parameters may change the convergence rate of the model, although it is not certain whether this quantity is positively correlated with the spectral radius of M or not. It should also be noted that introducing the delayed state only helps the model to simulate in more detail and better capture the information dissemination behavior of social network users, but it does not assert that the convergence condition of the SDIR model is always better than existing related models. We observe that when Assumption 1 is removed, the result of Theorem III.1 may not always be achieved.

IV. THE PROBLEMS AND BOUNDING FUNCTIONS

A. The main problem

Let the vector $m(t) := x(t) + y(t) + r(t)$ and $m^* \in \mathbb{R}^n$ consisting of elements $m_i^* = \sup_t m_i(t)$.

Definition IV.1. The quantity $\|m^* - m(0)\|_1$ is called the number of increased infections on G after the diffusion process ends.

In the following, we consider the problem of minimizing the diffusion quantity from the nodes in the given seed set by deleting some appropriate edges. Suppose we choose a set of edges $P \subseteq Q$ to delete from the graph. Denote B_{-P} as the matrix obtained from B , $G_{-P} = G(V, E \setminus P)$, and let $\sigma(P) := \|m^* - m(0)\|_1$ be the function of the number of increased infections in the network when the diffusion ends after deleting the set of edges P .

Problem IV.2. Given a directed graph $G(V, E)$ with $|V| = n$, representing SDIR diffusion model, an initial state vector $x(0) \in [0, 1]^n$ and $y(0), r(0)$ such that $x(0) + y(0) + r(0) \in [0, 1]^n$. Let Q be a candidate set of edges such that $Q \subseteq E$ and a positive integer k satisfying $k \leq |Q| = q$. Find a set of edges $P \subseteq Q, |P| \leq k$ to delete from the graph such that the infection amount on G_{-P} is minimized i.e., find

$$P^* \in \underset{P \subseteq Q, |P| \leq k}{\operatorname{argmin}} \sigma(P)$$

An easily noticeable point is that the objective function for Problem IV.2 lacks submodularity or supermodularity. This motivates us to find a good heuristic algorithm that provides a

solution with good optimality approximation. Similar to [34], if the condition of Theorem III.1 is satisfied, combined with the constraints of the above assumptions, we can find a monotonic upper bounding function that possesses supermodularity.

B. Supermodular Upper Bound

Theorem IV.3. *Under Assumption 1, when $\rho(M_{-P}) < 1$, the infection amount of the SDIR model with the removed edge set P does not exceed*

$$\sigma^U(P) = \mathbf{1}^\top (A + Q(I - A))^{-1} (D(I - M_{-P})^{-1} - I)(x(0) + Qy(0)).$$

If the constraints of Assumption 1 are satisfied, it is clear that the obtained upper bound function is not greater than the upper bound function for the D-SIR model of [34] and is significantly smaller if there exists a node i satisfying $D'_i > D_i$ (the reader is recommended to see the proof of Theorem III.1 in the appendix for more details). We have the following lemma to prove the monotonicity and supermodularity of the function $\sigma^U(\cdot)$.

Lemma IV.4. $\sigma^U(\cdot)$ is a non-increasing and supermodular function.

C. Supermodular Lower Bound

A noteworthy point in the SDIR model is that one can find a lower bound function that is also supermodular. Hence, instead of running the greedy algorithm only for the upper bound function $\sigma^U(\cdot)$, one can run the greedy algorithm once more for the following lower bound function $\sigma^L(\cdot)$. This is precisely the idea of the Sandwich approximation and will be presented in more detail in the next section. Let $N_{-P} := I - D + AS(0)B_{-P}$.

Theorem IV.5. *Under Assumption 2, when $\rho(N_{-P}) < 1$ the number of infections of the SDIR model with the set of removed edges P is at least*

$$\sigma^L(P) = \mathbf{1}^\top A^{-1} (D(I - N_{-P})^{-1} - I) (x(0) + W(W + D')^{-1}y(0)).$$

Lemma IV.6. $\sigma^L(\cdot)$ is a non-increasing and supermodular function.

The proofs of Lemmas IV.4 and IV.6 can be argued in a manner similar to that in [34]. Now, we have $\sigma^L(P) \leq \sigma(P) \leq \sigma^U(P)$, with $\sigma^L(P), \sigma^U(P)$ being two non-increasing and supermodular functions. Accordingly, the Sandwich approximation principle [15], [33] implies that the solution returned by the greedy algorithm always guarantees an approximation ratio relative to the optimal solution; specifically,

$$\sigma^U(\emptyset) - \sigma(P_{\text{Sand}}) \geq \max \left\{ \frac{\sigma^U(\emptyset) - \sigma(P_L)}{\sigma^U(\emptyset) - \sigma^L(P_L)}, \frac{\sigma^U(\emptyset) - \sigma^U(P^*)}{\sigma^U(\emptyset) - \sigma(P^*)} \right\} \left(1 - \frac{1}{e} - \epsilon \right) (\sigma^U(\emptyset) - \sigma(P^*)),$$

where $P_{\text{Sand}} = \operatorname{argmin}_{P \in \{P_L, P_0, P_U\}} \sigma(P)$ with P_L, P_U being the solutions returned by the greedy algorithm applied to $\sigma^L(\cdot)$

and $\sigma^U(\cdot)$, respectively, P_0 being a solution returned by some algorithm (possibly greedy) applied to $\sigma(\cdot)$, and P^* denoting the optimal solution for objective function.

V. ALGORITHMS

First, we have the following proposition showing that solving Problem IV.2 still remains computationally hard due to the use of a more general model than SIR model.

Proposition V.1. *The problem of finding the optimal edge set P^* to remove in order to minimize the spread in the SDIR network model, as formulated in IV.2, is NP-hard.*

Since the problem is NP-hard, it is very difficult to design an algorithm that provides an exact solution in all cases within polynomial time. Fortunately, with the two bound functions derived in the previous section, and when the condition $\rho(M_{-P}) < 1$ is satisfied, we propose the Sandwich approximation algorithm, which is described in Algorithm 2.

Algorithm 1 Greedy Algorithm (GA)

Input: A function $f \in \{\sigma^L, \sigma^U\}$, a graph G , initial states, a candidate edge set Q , and an integer k .

Output: An edge set $P \subseteq Q$ of size k .

Initialize $P \leftarrow \emptyset$

for $i = 1$ **to** n **do**

Compute $f(P \cup \{e\})$ for each $e \in Q \setminus P$
 $e^* \leftarrow \operatorname{argmax}_{e \in Q \setminus P} (f(P) - f(P \cup \{e\}))$
 $P \leftarrow P \cup \{e^*\}$

end

return P

Specifically, Algorithm 1 is applied to the two functions $\{\sigma^L, \sigma^U\}$ in order to obtain three corresponding edge sets by greedily selecting an edge in each iteration. Based on Lemmas IV.4 and IV.6, the solution returned by this greedy algorithm guarantees an approximation ratio of $(1 - 1/e - \epsilon)$ for the two functions $\sigma^U(\emptyset) - \sigma^L(\cdot)$ and $\sigma^U(\emptyset) - \sigma^U(\cdot)$ with respect to the optimal solution. Next, among the resulting edge sets, we compute the estimated infection for each set and select the one that yields the smallest spreading influence.

Algorithm 2 Sandwich Approximation Algorithm (SAA)

Input: A graph G , initial states, a candidate edge set Q , and an integer k .

Output: An edge set $P \subseteq Q$ of size k .

$P_L \leftarrow$ Result of GA for $\sigma^L(\cdot)$

$P_0 \leftarrow$ Result of a heuristic algorithm for $\sigma(\cdot)$

$P_U \leftarrow$ Result of GA for $\sigma^U(\cdot)$

return $P_{\text{Sand}} \leftarrow \operatorname{argmin}_{P \in \{P_L, P_0, P_U\}} \sigma(P)$

Remark V.2. The solution P_0 of $\sigma(\cdot)$ does not play a crucial role in proving the approximation guarantee of SAA. Moreover, the greedy method for $\sigma(\cdot)$ does not provide any guaranteed approximation ratio compared to the optimal solution, while incurring higher computational complexity. Therefore, P_0 can be obtained using other simpler methods, such as

random selection. We propose two algorithms that are GA and SAA.

VI. EXPERIMENTS

In this section, we present numerical examples to demonstrate the convergence and effectiveness of the proposed algorithms. We evaluate the performance of our method on a real-world dataset and compare it against some known strategies.

A. Comparison with Heuristic Algorithms

First, we compare the proposed SAA with two heuristic algorithms, namely *Max-degree* [2] and *Random* [7], applied to the SDIR model. The algorithms are defined as follows:

- 1) **Max-degree:** This algorithm iteratively removes the edge incident to the node with the highest weighted degree at each step.
- 2) **Random:** This algorithm removes an edge selected uniformly at random at each iteration. Note that the procedure for determining P_0 is integrated into this selection process.

We conducted experiments using the real-world contact network dataset collected in Haslemere, England [14]. The parameter settings and preprocessing steps follow the methodology described in Yi *et al* [34]. The new simulation parameters are initialized as follows: The initial state $\mathbf{y}(0)$ is set to zero for all nodes, except for five seed nodes which are assigned random values in the interval $[0, 0.05]$. Also, the initial state $\mathbf{x}(0)$ is set to zero for all nodes, except for five seed nodes which are assigned random values in the interval $[0.8, 0.85]$. Next, the infection rates A_i are distributed based on node categories: the five seed nodes have $A_i \in [0.9, 1]$; forty-five randomly selected nodes have $A_i \in [0.65, 0.87]$; and the remaining nodes have $A_i \in [0.15, 0.35]$. For each node $i \in V$, the weight W_i is chosen uniformly at random from $[0.15, 0.35]$. To ensure Assumption 1 is satisfied, D'_i is selected from the interval $[\max(0, D_i - W_i), 0.95 - W_i]$.

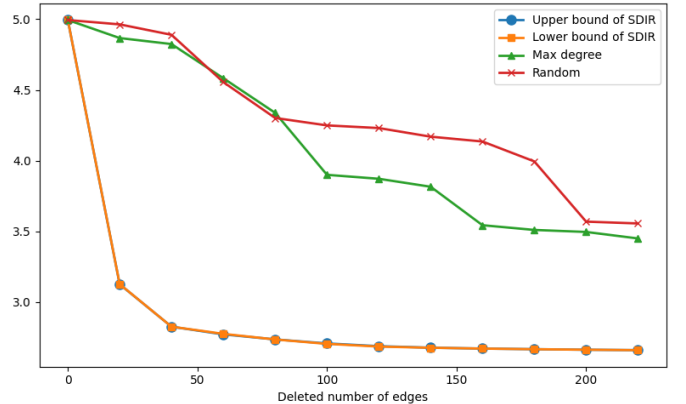


Fig. 2. Haslemere Network with preprocessing

The simulation results are presented in Fig. 2. It is evident that the solutions obtained from the proposed sandwich bounding functions significantly outperform the two heuristic

algorithms. The number of increased infections decreases rapidly within the removal of the first 50 edges. Notably, the upper and lower bounds derived from the sandwich algorithms are very close to each other, indicating the tightness of the approximation with respect to the objective function. It is similar to the results or the complexity of the Greedy algorithm in [34], in our simulations, the running time of our Algorithm 1 for the SDIR model is $O(n^3 + k(n^2 + qn))$.

B. Convergence Analysis

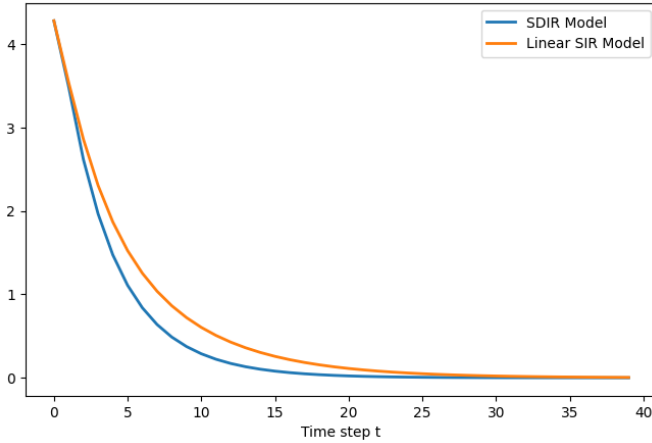


Fig. 3. Convergence of infection states over time

Next, utilizing the same parameter setup, we examine the convergence of the infection states in both the proposed SDIR model and the linear SIR model. The total amount of infection state of the SDIR model is defined as $\|x(t) + y(t)\|_1$. For comparison, the total amount of infection state of the linear SIR model is calculated as $\|\hat{x}(t)\|_1$, where the initial condition is set to $\hat{x}(0) = x(0) + y(0)$.

The results illustrated in Fig. 3 demonstrate that the SDIR model exhibits faster convergence compared to the linear SIR model. Specifically, the infection states in the SDIR model rapidly decays to $0_{n \times 1}$ within the first 10 time steps. In contrast, the linear SIR model exhibits a slower decay, requiring approximately 5 additional time steps to reach a similar extinction state. This empirical observation reinforces the theoretical validity of Theorem III.1 under Assumption 1.

VII. CONCLUSION

We studied the discrete-time SDIR model, a new modified SIR model. We showed that the sufficient condition for the convergence and the stability of the SDIR model. Furthermore, we investigate the problem that minimize the number of infections where deleting some edges. Two algorithms, greedy and Sandwich algorithms were proposed to solve the minimizing infections problem. Some experimental results were given.

REFERENCES

[1] D. Acemoglu, A. Malekian, and A. Ozdaglar, "Network security and contagion," *J. of Economic Theory*, 166, 2013, pp. 536–585.

[2] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, 406.6794, 2000, pp. 378–382.

[3] T. Alpcan and T. Basar, *Network security: A decision and game-theoretic approach*, Cambridge University Press, 2010.

[4] N. T. Bailey, *The mathematical theory of infectious diseases and its applications*, Griffin, London, 1975.

[5] H. J. Ahn, B. Hassibi, "Global dynamics of epidemic spread over complex networks," *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)*, IEEE, 2013, pp. 4579–4585.

[6] A. L. Barabasi, "The origin of bursts and heavy tails in human dynamics," *Nature*, 435, 2005, pp. 207–211.

[7] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Physical Review letters*, 85, 2000, pp. 5468–5471.

[8] W. Chen, C. Castillo, L. V. Lakshmanan, *Information and influence propagation in social networks*, Morgan & Claypool Publishers, 2014.

[9] P. Jacquet, B. Mans, and G. Rodolakis, "Information propagation speed in mobile and delay tolerant networks," *IEEE Transactions on Information Theory*, 56 (10), 2010, pp. 5001–5015.

[10] S. Gracy, P. E. Pare, H. Sandberg, K. H. Johansson, "Analysis and distributed control of periodic epidemic processes," *IEEE Trans. Control Netw. Syst.*, 8, 2021, pp. 123–134.

[11] D. Kempe, J. M. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2003, pp. 137–146.

[12] D. Kempe, J. Kleinberg, E. Tardos, "Maximizing the Spread of Influence through a Social Network," *Theory of Computing*, 11 (4), 2015, pp. 105–147.

[13] M. Kimura, K. Saito, and H. Motoda, "Solving the Contamination Minimization Problem on Networks for the Linear Threshold Model," *PRICAI (Lecture Notes in Computer Science)*, 5351, 2008, pp. 977–984.

[14] P. Klepac, S. Kissler, and J. Gog, "Contagion! the bbc four pandemic—the model behind the documentary," *Epidemics*, 24, 2018, pp. 49–59.

[15] W. Lu, W. Chen, L. V. Lakshmanan, "From Competition to Complementarity: Comparative Influence Diffusion and Maximization," *Proc. VLDB Endow.*, 9 (2), 2015, pp. 60–71.

[16] P. Van Mieghem, "Virus Spread in Networks," *IEEE/ACM Transactions on Networking*, 17 (1), 2009, pp. 1–14.

[17] P. Van Mieghem, *Graph spectra for complex networks*, Cambridge University Press, Cambridge, 2011.

[18] P. Van Mieghem, F. D. Sahneh, C. Scoglio, "An upper bound for the epidemic threshold in exact Markovian SIR and SIS epidemics on networks," *53rd IEEE Conference on Decision and Control*, Los Angeles, CA, USA, 2014, pp. 6228–6233.

[19] C. Nowzari, V. M. Preciado, G. J. Pappas, "Analysis and control of epidemics: A survey of spreading processes on complex networks," *IEEE Control Syst.*, 36, 2016, pp. 26–46.

[20] P. E. Pare, J. Liu, C. Beck, B. Kirwan, T. Basar, "Analysis, Estimation, and Validation of Discrete-Time Epidemic Processes," *IEEE Transactions on Control Syst. Tech.*, 28 (1), 2020, pp. 79–93.

[21] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, "Epidemic processes in complex networks," *Rev. Mod. Phys.* 87, 2015, pp. 925–979.

[22] Pham, C.V., Thai, M.T., Duong, H.V., Bao, Q.B., Hoang, X.H., "Maximizing misinformation restriction within time and budget constraints," *J. Comb. Optim.* 35, 2018, pp. 1202–1240.

[23] Pham, C.V., Phu, Q.V., Hoang, H.X., J. Pey, My. T. Thai, "Minimum budget for misinformation blocking in online social networks," *J. Comb. Optim.* 38, 2019, pp. 1101–1127.

[24] J. E. Phelps, R. Lewis, L. Mobilio, D. Perry, and N. Raman, "Viral marketing or electronic word-of-mouth advertising: Examining consumer responses and motivations to pass along email," *Journal of Advertising research*, 44 (04), 2004, pp. 333–384.

[25] J. Qi, X. Liang, Y. Wang, H. Cheng, "Discrete time information diffusion in online social networks: micro and macro perspectives," *Sci. Rep.* 8, 2018, pp. 11872.

[26] M. Richardson and P. Domingos, "Mining knowledge-sharing sites for viral marketing," *Proceedings of the 8th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2002, pp. 61–70.

[27] A. Ruhi, B. Hassibi, "SIRS epidemics on complex networks: Concurrence of exact Markov chain and approximated models," *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, IEEE, 2015, pp. 2919–2926.

- [28] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 115 (772), 1927, pp. 700–721.
- [29] P. Shakarian, A. Bhatnagar, A. Aleali, E. Shaabani, R. Guo, *Diffusion in Social Networks*, Springer, 2015, xi+101 pp.
- [30] T. C. Silva; L. Zhao, *Machine learning in complex networks*, Springer, Cham, 2016. xviii+331 pp.
- [31] J. Xie, F. Zhang, K. Wang, X. Lin, and W. Zhang, "Minimizing the Influence of Misinformation via Vertex Blocking," *ICDE. IEEE*, 2023, pp. 789–801.
- [32] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," *Proc. 22nd Int. Symp. Reliable Distributed Systems (SRDS'03)*, 2003, pp. 25–34.
- [33] Z. Wang, Y. Yang, J. Pei, L. Chu, E. Chen, "Activity maximization by effective information diffusion in social networks," *IEEE Transactions on Knowledge and Data Engineering*, 29.11 2017, pp. 2374–2387.
- [34] Y. Yi, L. Shan, P. Pare, K. H. Johansson, "Edge deletion algorithms for minimizing spread in SIR epidemic models," *SIAM Journal on Control and Optimization*, 60.2, 2022, pp. 246–273.
- [35] M. Youssef, C. Scoglio, "An individual-based approach to SIR epidemics in contact networks," *J. Theoret. Biol.*, 283, 2011, pp. 136–144.

APPENDIX

A. Proof of Theorem III.1

Under Assumption 1, it is clear that $D_i \leq D'_i, \forall i = 1, 2, \dots, n$. Hence we can choose $q_i \in \left[\frac{W_i}{W_i + D'_i - D_i}, 1 \right]$. We have: $\mathbf{x}(t+1) + \mathbf{Qy}(t+1)$

$$\begin{aligned} &= (\mathbf{I} - \mathbf{D} + (\mathbf{A} + \mathbf{Q}(\mathbf{I} - \mathbf{A}))\mathbf{S}(0)\mathbf{B})\mathbf{x}(t) \\ &\quad + (\mathbf{Q}\mathbf{I} - \mathbf{Q}\mathbf{D}' - (\mathbf{I} - \mathbf{Q})\mathbf{W})\mathbf{y}(t) \\ &= (\mathbf{I} - \mathbf{D} + (\mathbf{A} + \mathbf{Q}(\mathbf{I} - \mathbf{A}))\mathbf{S}(0)\mathbf{B})\mathbf{x}(t) \\ &\quad + (\mathbf{I} - \mathbf{D}' - (\mathbf{Q}^{-1} - \mathbf{I})\mathbf{W})\mathbf{Qy}(t) \\ &\leq (\mathbf{I} - \mathbf{C} + (\mathbf{A} + \mathbf{Q}(\mathbf{I} - \mathbf{A}))\mathbf{S}(0)\mathbf{B})\mathbf{x}(t) \\ &\quad + (\mathbf{I} - \mathbf{C} + (\mathbf{A} + \mathbf{Q}(\mathbf{I} - \mathbf{A}))\mathbf{S}(0)\mathbf{B})\mathbf{Qy}(t) \\ &= \mathbf{M}(\mathbf{x}(t) + \mathbf{Qy}(t)), \forall t \geq 0. \end{aligned}$$

Therefore, $\mathbf{x}(t) + \mathbf{Qy}(t) \leq \mathbf{M}^t(\mathbf{x}(0) + \mathbf{Qy}(0))$. If $\rho(\mathbf{M}) \leq 1 - \epsilon$, then $\|\mathbf{x}(t) + \mathbf{Qy}(t)\| \leq (1 - \epsilon)^t \|\mathbf{x}(0) + \mathbf{Qy}(0)\|$, which forces $\mathbf{x}(t)$ and $\mathbf{y}(t)$ to converge to 0 as $t \rightarrow \infty$. This completes the proof.

B. Proof of Theorem IV.3

From the proof of Theorem III.1, we have $\mathbf{x}(t) \leq \mathbf{M}_{-P}^t(\mathbf{x}(0) + \mathbf{Qy}(0)), \forall t \geq 0$. Next, observe that

$$\begin{aligned} \mathbf{m}(t) &= \mathbf{x}(t) + \mathbf{y}(t) + \mathbf{r}(t) \\ &= \mathbf{x}(0) + \mathbf{y}(0) + \mathbf{r}(0) + \mathbf{S}(0)\mathbf{B}_{-P} \sum_{l=0}^{t-1} \mathbf{x}(l) \\ &\leq \mathbf{m}(0) + \mathbf{S}(0)\mathbf{B}_{-P} \sum_{l=0}^{t-1} \mathbf{M}_{-P}^l(\mathbf{x}(0) + \mathbf{Qy}(0)). \\ &\Rightarrow \|\mathbf{m}^* - \mathbf{m}(0)\|_1 = \lim_{t \rightarrow \infty} \|\mathbf{m}(t) - \mathbf{m}(0)\|_1 \\ &\leq \mathbf{1}^\top \mathbf{S}(0)\mathbf{B}_{-P}(\mathbf{I} - \mathbf{M}_{-P})^{-1}(\mathbf{x}(0) + \mathbf{Qy}(0)). \end{aligned}$$

This completes the proof.

C. Proof of Theorem IV.5

First, we have the following lemma:

Lemma VII.1. *Let \mathbf{A} and \mathbf{B} be two square nonnegative real matrices of the same dimension with spectral radius $\rho(\mathbf{A}) < 1$ and $\rho(\mathbf{B}) < 1$. Then*

$$L = \lim_{t \rightarrow \infty} \sum_{s=0}^t \sum_{l=0}^s \mathbf{A}^l \mathbf{B}^{s-l} = (\mathbf{I} - \mathbf{A})^{-1}(\mathbf{I} - \mathbf{B})^{-1}.$$

Proof. Denote $\mathbf{C}_l = \mathbf{I} + \mathbf{B} + \dots + \mathbf{B}^l$ for $l = 0, 1, 2, \dots$. For $t > t_0$ with fixed $t_0 \in \mathbb{N}^*$ we have $\sum_{s=0}^t \sum_{l=0}^s \mathbf{A}^l \mathbf{B}^{s-l} \geq \sum_{l=0}^{t_0} \mathbf{A}^l \mathbf{C}_{t-t_0} + \mathbf{A}^{t_0+1} \sum_{l=0}^{t-t_0-1} \mathbf{A}^l$. Fixing t_0 and letting $t \rightarrow \infty$ yields $L \geq (\mathbf{I} + \dots + \mathbf{A}^{t_0})(\mathbf{I} - \mathbf{B})^{-1} + \mathbf{A}^{t_0+1}(\mathbf{I} - \mathbf{A})^{-1}$. Now letting $t_0 \rightarrow \infty$ gives $L \geq (\mathbf{I} - \mathbf{A})^{-1}(\mathbf{I} - \mathbf{B})^{-1}$. Similarly, one may show the reverse bound by observing $\sum_{s=0}^t \sum_{l=0}^s \mathbf{A}^l \mathbf{B}^{s-l} \leq \sum_{l=0}^{t_0} \mathbf{A}^l \mathbf{C}_t + \mathbf{A}^{t_0+1} \sum_{l=0}^{t-t_0-1} \mathbf{A}^l \mathbf{C}_{t-t_0-1}$, and repeating the same limiting argument to obtain $L \leq (\mathbf{I} - \mathbf{A})^{-1}(\mathbf{I} - \mathbf{B})^{-1}$. Thus the lemma is proved. \square

Returning to Theorem IV.5. We have $\mathbf{x}(t+1) = \mathbf{N}_{-P}\mathbf{x}(t) + \mathbf{W}\mathbf{y}(t) = \mathbf{N}_{-P}^{t+1}\mathbf{x}(0) + \sum_{l=0}^t \mathbf{N}_{-P}^{t-l}\mathbf{W}\mathbf{y}(l), \forall t \geq 0$. It is easy to see that $\mathbf{y}(t) \geq \mathbf{F}\mathbf{y}(t-1)$ with $\mathbf{F} = \mathbf{I} - \mathbf{W} - \mathbf{D}'$, it follows that $\mathbf{y}(t) \geq \mathbf{F}^t\mathbf{y}(0)$ for all $t \geq 0$. So

$$\begin{aligned} \mathbf{m}(t+1) - \mathbf{m}(0) &= \mathbf{S}(0)\mathbf{B}_{-P} \left(\sum_{s=1}^t (\mathbf{N}_{-P}^s \mathbf{x}(0) \right. \\ &\quad \left. + \sum_{l=0}^{s-1} \mathbf{N}_{-P}^{s-1-l} \mathbf{W}\mathbf{y}(l)) + \mathbf{x}(0) \right) \geq \\ &\mathbf{S}(0)\mathbf{B}_{-P} \sum_{l=0}^t \mathbf{N}_{-P}^l \mathbf{x}(0) + \mathbf{S}(0)\mathbf{B}_{-P} \sum_{s=0}^{t-1} \sum_{l=0}^s \mathbf{N}_{-P}^{s-l} \mathbf{F}^l \mathbf{W}\mathbf{y}(0). \end{aligned}$$

Setting $t \rightarrow \infty$ and applying Lemma VII.1 yields $\mathbf{m}^* - \mathbf{m}(0) \geq \mathbf{S}(0)\mathbf{B}_{-P} \left((\mathbf{I} - \mathbf{N}_{-P})^{-1} \mathbf{x}(0) + (\mathbf{I} - \mathbf{N}_{-P})^{-1} (\mathbf{I} - \mathbf{F})^{-1} \mathbf{W}\mathbf{y}(0) \right)$. Then, equivalently, $\|\mathbf{m}^* - \mathbf{m}(0)\|_1 \geq \mathbf{1}^\top \mathbf{A}^{-1}(\mathbf{D}(\mathbf{I} - \mathbf{N}_{-P})^{-1} - \mathbf{I})(\mathbf{x}(0) + \mathbf{W}(\mathbf{W} + \mathbf{D}')^{-1}\mathbf{y}(0))$.

The theorem is proved.