Article
Minimizing the Spread of Negative Influence in SNIR Model by Contact Blocking

Caiyan Dai 1, Ling Chen 2,*, Kongfa Hu 1,3 and Youwei Ding 1

1 College of Artificial Intelligence and Information Technology, Nanjing University of Chinese Medicine, Nanjing 210023, China
2 College of Information Engineering, Yangzhou University, Yangzhou 225012, China
3 Jiangsu Collaborative Innovation Center of Traditional Chinese Medicine in Prevention and Treatment of Tumor, Nanjing 210023, China
* Correspondence: yzulchen@163.com

Abstract: This paper presents a method to minimize the spread of negative influence on social networks by contact blocking. First, based on the infection-spreading process of COVID-19, the traditional susceptible, infectious, and recovered (SIR) propagation model is extended to the susceptible, non-symptomatic, infectious, and recovered (SNIR) model. Based on this model, we present a method to estimate the number of individuals infected by a virus at any given time. By calculating the reduction in the number of infected individuals after blocking contacts, the method selects the set of contacts to be blocked that can maximally reduce the affected range. The selection of contacts to be blocked is repeated until the number of isolated contacts that need to be blocked is reached or all infection sources are blocked. The experimental results on three real datasets and three synthetic datasets show that the algorithm obtains contact blockings that can achieve a larger reduction in the range of infection than other similar algorithms. This shows that the presented SNIR propagation model can more precisely reflect the diffusion and infection process of viruses in social networks, and can efficiently block virus infections.

Keywords: precise isolation; minimize virus infection; SNIR model

1. Introduction

In recent years, there have been many infectious diseases in the international community which generally have the characteristics of rapid transmission and a wide range of harms. Therefore, scholars have studied the spread and control of epidemics [1] and proposed the SIR model. In this model, every individual is in one of three states: susceptibility, infection, or recovery. According to the characteristics of disease transmission, there are different improved SIR models. The specific information of the SIR model was described in [2]. Subsequently, a series of infectious diseases were studied based on the SIR model [3-6]. In this context, how to effectively control the spread of an epidemic is a research hotspot. One method to control the spread is to use the spectral norm of the minimized transfer matrix [7-12]. Another method is to use resource allocation [8] and improve the network model [11,12]. Some researchers have proposed studying the optimal control problem based on a given boundary [13,14] to minimize the cost of controlling the spread of an epidemic.

In the process of fighting COVID-19, isolating infected persons is an effective way to prevent infection. However, isolating every infected person (symptomatic or asymptomatic) comes with a certain price. In the case of large-scale infection, simply isolating all contacts of all infected persons would come with a great price. Therefore, we should adopt a strategy of precise isolation to block some essential contacts (i.e., edges in the network) among the people who have contact with the infected person, to minimize the final scope of virus infection at a reasonable cost.
Due to the openness and high speed of social networks, all types of false information and rumors can spread quickly and widely. Such false messages are typically accompanied by network hotspots or eye-catching information and tend to attract the attention of network users. False news will cause harm to individuals to some extent and can lead to social panic, a decline in business credibility, damage to personal reputation, and a serious loss of network security. Therefore, such misinformation must be controlled to make social networks more reliable and secure for information exchange. We must find efficient approaches to limit the destructive impact of negative influences.

Based on the infection process of COVID-19, we extended the traditional SIR model. The traditional model has the following limitations:

1. The SIR model using traditional methods does not consider the state of asymptomatic infection. In the actual process of a COVID-19 infection, it has been proven that there is an asymptomatic phase; although if infected people are asymptomatic, they will infect others.
2. The SIR model using traditional methods assumes that recovered people have antibodies, without considering the possibility of them becoming infected again. In the actual process of a COVID-19 infection, it has been proven that when antibodies are no longer present, recovered people will become susceptible and could be reinfected.
3. In most traditional methods, only one transmission source, namely “patient zero”, is considered for a certain area. In the actual process of a COVID-19 infection, it has been proven that there may be multiple sources of infection in one region from various other regions.

Therefore, we propose the SNIR infection model. We studied the method of minimizing virus infection using the SNIR model based on the accurate isolation of multiple infection sources. At the same time, the model was also applied to describe the process of spreading rumors, in order to more accurately suppress them.

In recent years, methods have been proposed for the influence-blocking maximization problem using different propagation models. However, there are still many challenges in maximally blocking negative influences in complex social networks. The main challenges and difficulties are as follows:

1. Most of the influence-blocking maximization algorithms assume that there is only a single source of negativity in the network. However, in real-world social networks, negativity is probably sent out from different channels that are linked with multiple sources. In order to block the influence from multiple propagation sources, we need to know the relationship between each node and each source. In addition, we also must consider the interactions between influences from different propagation sources, which is difficult to analyze due to the randomness of influence spreading.
2. Most of the influence-blocking maximization algorithms try to limit the spread of negative influence by deleting some nodes, specifically by isolating some individuals in the social network. However, to prevent the spread of an epidemic, it is impossible to isolate individuals. A feasible way is to stop contact between some people.
3. Some existing methods of influence-blocking leverage node centrality to select the nodes or edges to be blocked and ignore the propagation probability between nodes. Other methods use a BFS tree instead of the original network to simulate the influence propagation process. However, the BFS tree cannot precisely reflect the real process of influence propagation. In addition, these methods treat all nodes and edges in the network equally and ignore the latent topological features of the nodes. All of these factors reduce the quality of the results of influence-blocking.
4. In most existing influence-blocking algorithms, a greedy strategy is used to select the nodes or edges to be blocked. In each step of greedy source node selection, a large number of simulations are needed to estimate the propagation range of the candidate source node set. Such a simulation of influence spreading is #P-complete, which requires large amounts of computation time; therefore, it is not applicable to
large-scale networks. It is challenging to quickly and effectively find the contacts to be blocked in large-scale networks.

To tackle these difficulties, we propose the SNIR propagation model for influence spreading based on the spread of COVID-19 infections. Based on this model, we propose an algorithm named MaxExpectedH (maximum expected $H$ values), in which the value of $H$ changes with the spreading of the virus.

The main innovation and contributions of this paper are as follows:

1. We propose the SNIR propagation model, which adds the asymptomatic infected state. The new model is more accurate than the traditional SIR model and can reflect the real propagation process of viral infection spreading.

2. We propose a method for estimating the influence propagation range of each node at different times. Since we define a set of functions to estimate the probability for each node at different states at each time step, our method does not need to perform time-consuming simulations, thus it requires much less computation time than other methods.

3. We present a method for selecting sets of contacts that need to be isolated in ascending order according to the value of the infection range of the virus.

4. The experimental results show that the algorithm proposed in this paper can block negative influences more effectively than other methods. This shows that the presented SNIR propagation model can more precisely reflect the diffusion and infection process of viruses in social networks.

The remainder of this paper is structured as follows: A review of related works is presented in Section 2. In Section 3, the SNIR propagation model is proposed and the problem is defined. In Section 4, we propose a method to calculate the probability of nodes in different states. In Section 5, we propose an algorithm to calculate the range of influence spreading. In Section 6, an algorithm for selecting contacts to be blocked is proposed. Section 7 shows and analyzes the experimental results. Section 8 presents conclusions and further research.

2. Related Work

In recent years, by extending two traditional propagation models, the linear threshold (LT) model and the independent cascade (IC) model, many improved models have been proposed for analyzing maximum propagation in social networks [15].

Sahar et al. [16] proposed a path-based method for analyzing influence maximization in social networks. They noted that a small set of nodes, if activated, would spread information all over the network from two complementary perspectives, adapting the proposed algorithm to large-scale networks. Su et al. [17] proposed an algorithm for minimizing the seed set cost of influence spreading with a probabilistic guarantee. They define the problem as the minimum cost seed selection with probabilistic influence spreading guarantee in the linear threshold (LT) model. To avoid simulating the propagation of influence, an algorithm is adopted for estimating propagation by path counting in sample graphs. Masmoud et al. [18] considered a new hybrid greedy approach based on a community detection algorithm and propose a MADM technique to cope with the optimization of influence when analyzing complex networks. They referenced community detection and the TOPSIS technique. Li et al. [19] proposed an alternative solution for the IM problem that attempts to select ordinary grassroots as seeds. First, they empirically proved that grassroots are a better choice than elites in the IM problem from the aspects of relationship strength and polarities, based on statistics and the analysis of real datasets. Then they developed a grassroots-oriented seed-users-seeking algorithm that fully explores the community information of the network structure. Caliò et al. [20] proposed integrating a categorical-based notion of seed diversity into the objective function of a targeted influence maximization problem. They assumed that the users of a social network are associated with a categorical dataset, with each tuple expressing the profile of a user according to a predefined schema of categorical attributes.
In real social networks, there can be both positive and negative influences. Some researchers have focused on negative influences. The goal, through searching and setting negative influence sources, is to maximize the positive influence [21,22]. Kuhnle et al. [23] identified a new property, a generalized deterministic submodule, that ensures that propagation on the multiplex overall is submodular. In this case, they formulated an influential seed finder, a greedy algorithm with an approximation ratio (1-1/e). They also formulated an algorithm for knapsack seeding of the network that runs on each layer of the multiplex in parallel. Wang et al. [24] proposed the time-sensitive positive influence maximization problem by considering two factors simultaneously, to select the seed node set that would achieve the maximum spreading of positive influence within a specified time limit. Furthermore, they constructed a heat diffusion-based polarity influence diffusion model and an improved \( k \)-step greedy seed node selection algorithm to solve the TP-IM problem.

Some factors, such as vaccines and medical help, are also important in blocking the process of viral spreading. Khubchandani et al. [25] investigated the impact of COVID-19 morbidity and mortality among family and friends on vaccination preferences. They suggested that the dangers of not receiving the vaccine should be emphasized, as many people who do not know someone who was affected by COVID-19 are hesitant about vaccination. Long et al. [26] examined the spread of the COVID-19 pandemic in terms of social relationships. They specifically focused on the relational mechanisms of medical aid for people infected by COVID-19 and made recommendations for future public health policy and recovery.

At present, blocking negative influences is also a hotspot in the study of influence transmission. Song et al. [27] considered a more realistic situation, with the goal of reducing the number of rumor-infected users before a deadline, which they called the temporal influence blocking (TIB) problem. They proposed a two-phase solution called TIB-Solver to select \( k \) nodes to start a truth campaign such that the number of people reached by a rumor is minimized. Ghoshal et al. [28] leveraged the community structure of online social networks to select seed nodes statistically, independent of the distribution of misinformed nodes, for faster containment of misinformation with a simple one-time computation. They extended the work to include OSNs with the overlapped community as well. A competitive diffusion model was proposed for modeling the propagation of two types of competitive information in the same network [29]. The problem of minimizing the spread of rumors in social networks was explored and a novel heuristic based on diffusion dynamics was proposed to solve the rumor propagation problem under the LT1DT. The rumor propagation model can also be applied to the spread of infectious diseases.

Although the abovementioned methods are effective for the problem of blocking negative influences, their efficiency needs to be improved. Some of the algorithms select nodes or edges to be blocked based on their centrality and ignore the probability of propagation between nodes. Other methods simulate the process of influence spreading on trees instead of a general structured network. It is difficult for these methods to precisely detect the real diffusion sources without utilizing the latent structural information of the network. Furthermore, most of these approaches block the negative influence by isolating the nodes instead of the edges. This strategy is not practicable for preventing epidemics.

To overcome these defects, we propose a representation learning-based method for locating multiple influence sources. Compared with the other source-detecting methods, our algorithm has the following advantages:

1. It is based on the SNIR propagation model, which enables us to learn more precisely about viral spreading. Therefore, the method can obtain more accurate results than the other methods.
2. It establishes a set of functions to estimate the probability of each node in different states at each time step of influence propagation. Due to the nonparametric functions of the SNIR model, it is much easier to calculate probability with our method than with other methods.
(3). Based on the nodes’ probabilities in different states at a given time, our method precisely calculates the probability that each link will be blocked to prevent the spreading of negative influence. Since the model reflects influence propagation in the network, the influence probability obtained is more accurate than that estimated by Monte Carlo sampling.

3. Spreading Model and Problem Definition

In this section, we first propose the SNIR infection model, then give the problem definition.

3.1. SNIR Spreading Model

According to the actual process of COVID-19 infection, we propose the following SNIR infection model. In this model, every individual is in one of four states:

Susceptible (S): People in this state have not been infected, thus they will not infect others. They may become infected (turning to state I) with probability $\beta$, and they may also become infected but asymptomatic with probability $\alpha$ (turning to state N). Here, $\alpha + \beta < 1$.

Non-symptomatic (N): People in this state have been infected and will infect others, but they are asymptomatic. They will become infected with symptoms (turning to state I) with probability $\delta$, and will recover with probability $\eta$ and become convalescent (turning to state R). Here, $\delta + \eta < 1$.

Infectious (I): People in this state have been infected and have symptoms and will infect others. They will recover with probability $\gamma$ and become convalescent (turning to state R). Here, $\gamma < 1$.

Recovered (R): People in this state were infected, and now they have recovered. They have antibodies and will not infect others. However, without antibodies present, they will become susceptible (turning to state S) with probability $\xi$. Here, $\xi < 1$.

The transformation relationship of the above four states is shown in Figure 1.

![Figure 1. State transition process of the SNIR model.](image)

This model can also be used to describe the process of rumor propagation. Everyone’s state in the process of rumor propagation is as follows:

S: People in this state have not been exposed to rumors, thus they will not spread rumors to others. However, they will hear rumors with probability $\beta$ and believe them, and then become believers (turning to state I). They will also hear rumors with probability $\alpha$ but have a neutral attitude about them (turning to state N).

N: People in this state have heard rumors and will spread them to others, but they have a neutral attitude about rumors themselves and do not fully believe them. They will become believers (turning to state I) with probability $\delta$, and will reject rumors (turning to state R) with probability $\eta$. 

I: People in this state have heard and believed rumors, and will spread them to others. They will also be influenced by positive information that changes their response to disbelief (turning to state R) with probability $\gamma$.

R: People in this state have accepted rumors in the past, but later, due to the influence of positive information, changed their ideas, became nonbelievers, and will not spread rumors to others. However, with changes in public opinion and the environment, over time they will also change to state S with probability $\xi$.

3.2. Setting the Transition Probabilities

In order to control the spread of an epidemic, people can be vaccinated to prevent infection, and after infected people recover, they may have antibodies to fight off the virus. All of these factors must be considered in the propagation model.

Let the protection rate of the vaccine be $\omega$ and the protection rate of antibodies be $\psi$. For people who have not been vaccinated and do not have antibodies, let the value of their probabilities $\alpha, \beta, \delta, \gamma, \eta,$ and $\xi$ be $\alpha', \beta', \delta', \gamma', \eta'$, and $\xi'$, respectively. Considering the effect of vaccines and antibodies, the transition probabilities can be set as follows:

$$
\alpha = \begin{cases} 
1 - (1 - \alpha') \cdot (1 - \omega) & \text{if vaccinated} \\
1 - (1 - \alpha') \cdot (1 - \psi) & \text{if antibodies present} \\
\alpha' & \text{otherwise}
\end{cases}
$$

$$
\beta = \begin{cases} 
\beta' \cdot (1 - \omega) & \text{if vaccinated} \\
\beta' \cdot (1 - \psi) & \text{if antibodies present} \\
\beta' & \text{otherwise}
\end{cases}
$$

$$
\gamma = \begin{cases} 
1 - (1 - \gamma') \cdot (1 - \omega) & \text{if vaccinated} \\
1 - (1 - \gamma') \cdot (1 - \psi) & \text{if antibodies present} \\
\gamma' & \text{otherwise}
\end{cases}
$$

$$
\delta = \begin{cases} 
\delta' \cdot (1 - \omega) & \text{if vaccinated} \\
\delta' \cdot (1 - \psi) & \text{if antibodies present} \\
\delta' & \text{otherwise}
\end{cases}
$$

$$
\eta = \begin{cases} 
1 - (1 - \eta') \cdot (1 - \omega) & \text{if vaccinated} \\
1 - (1 - \eta') \cdot (1 - \psi) & \text{if antibodies present} \\
\eta' & \text{otherwise}
\end{cases}
$$

$$
\xi = \begin{cases} 
\xi' \cdot (1 - \omega) & \text{if vaccinated} \\
\xi' \cdot (1 - \psi) & \text{if antibodies present} \\
\xi' & \text{otherwise}
\end{cases}
$$

3.3. Problem Definition

Given a social network $G = (V, E, P)$, where $V$ is the set of individuals, i.e., the network nodes, and $E$ is the contact relationship between individuals, i.e., the network edges, probability $p_{u,v}$ on edge $(u, v)$ represents the probability that the virus will be transmitted from $u$ to $v$. Given the set of infected people observed $O = \{o_1, o_2, \ldots, o_m\}$, $O \subset V$, the $i$th observer $o_i \in O$ can be represented by a binary tuple $(o_i, Q_i)$. Here, $Q_i \in \{N, I\}$, indicating $o_i$ in an infectious or non-symptomatic state. We assume that the cost of isolating a contact is 1, and a positive integer $k$, which is our predetermined cost, is given. Given the positive integer $T$, which is the maximum transmission time of the virus, we need to find out the contact edges $X = \{e_1, e_2, \ldots, e_k\}$ whose number is no more than $k$, so that in the graph $G = (V, E \setminus X, P)$, the range of virus infection is the smallest after time $T$. This supports the set of individuals in four states in the network at a certain time is $S, N, I, R$, and the range of virus infection at that time is $|N| + |I|$. 
4. Calculation of the Probability of Nodes in Various States at a Given Time

Suppose the negative influence was instigated at time \( t = 0 \), and the probabilities of node \( u \) in states \( S, N, I, \) and \( R \) at time \( t \) are \( p_S(u, t), p_N(u, t), p_I(u, t), \) and \( p_R(u, t) \), respectively. Here, \( p_S(u, t) + p_N(u, t) + p_I(u, t) + p_R(u, t) = 1 \).

Let \( r(u, t) \) be the probability of virus transmission to \( u \) at time \( t \) and let \( a(u, t) \) be the probability that \( u \) is infected by the virus at time \( t \).

Set the initial value of the above variables (i.e., when \( t = 0 \)) as:

\[
p_S(u, 0) = \begin{cases} 1 & u \in S \\ 0 & \text{else} \end{cases} \quad a(u, 0) = \begin{cases} 1 & u \in I \\ 0 & \text{else} \end{cases} \quad p_R(u, 0) = \begin{cases} 1 & u \in R \\ 0 & \text{else} \end{cases} \quad p_N(u, 0) = \begin{cases} 1 & u \in N \\ 0 & \text{else} \end{cases}
\]

(1)

Then, at each time \( t \):

\[
r(u, t + 1) = 1 - \prod_{v \in \Gamma(u)} [1 - p_{v,u}a(u, t)]
\]

(4)

Here, \( \Gamma(u) = \{v| (v, u) \in E\} \) is the set of incoming neighbors of \( u \), then

\[
a(u, t + 1) = r(u, t + 1)\{[\alpha + \beta]p_S(u, t) + (1 - \gamma \cdot p_I(u, t + 1) - \eta \cdot p_N(u, t))\}
\]

(5)

\[
p_I(u, t + 1) = (1 - \gamma) \cdot p_I(u, t) + r(u, t + 1) \cdot \beta \cdot p_S(u, t) + \delta \cdot p_N(u, t)
\]

(6)

\[
p_R(u, t + 1) = (1 - \zeta) \cdot p_R(u, t) + \gamma \cdot p_I(u, t) + \eta \cdot p_N(u, t)
\]

(7)

\[
p_N(u, t + 1) = (1 - \delta - \eta) \cdot p_N(u, t) + \alpha \cdot p_S(u, t)
\]

(8)

\[
p_S(u, t + 1) = 1 - p_N(u, t + 1) - p_I(u, t + 1) - p_R(u, t + 1)
\]

(9)

At each time, the individuals in the network \( W = N \cup I \) are the source of infection. Let the time when the virus begins to spread be \( t = 0 \), and the probability that vertex \( u \) is in states \( S, N, I, \) and \( R \) at time \( t \) is \( p_S(u, t), p_N(u, t), p_I(u, t), \) and \( p_R(u, t) \). Then, the initial value can be set for each vertex \( u \) by Equations (1)–(3) as \( p_S(u, 0), p_N(u, 0), p_I(u, 0), \) and \( p_R(u, 0) \). Then, \( p_S(u, t), p_N(u, t), p_I(u, t), \) and \( p_R(u, t) \) can be obtained for each vertex \( u \) by iterative Formulas (4)–(9).

5. Calculating the Range of Influence Spreading

We first give an algorithm to calculate the range \( I(G, \Omega) \) of virus infection at time \( T \) when the initial state is \( \Omega = (S,N,I,R) \) on network \( G = (V, E, P) \). The algorithm first computes \( p_Q(u, t) \) for all \( u \in V \). When calculating the transmission process with \( W = N \cup I \) as the infection source, because the virus only infects the neighbors related to vertex \( v \) in \( W \), and not other individuals, it is not necessary to calculate the probability for all individuals, only the neighbors directly or indirectly related to \( v \). After obtaining the probability \( p_Q(u, t) \) \((u \in V, Q = S,N,I,R, t = 1, 2, \ldots, T)\), we know each vertex \( u \) with \( W = N \cup I \) as the infectious source in state \( Q \) at time \( t \). Later, the statistics of the following value are calculated:

\[
I(G, \Omega) = \sum_{t=0}^{T} \left[ \sum_{u \in V} p_Q(u, t) + \sum_{u \in V} p_Q(u, t) \right]
\]

(10)

We present an algorithm to estimate the expected value of influence spreading at each time \( t \). The framework of the algorithm is given in Algorithm 1. The value of \( H \) is used to indicate the infection range of virus infection that changes with time \( T \) and it is related to the different state probabilities of each node. (V, E, P):(S, N, I, R):
Entropy 2022, 24, 1623
8 of 17

Algorithm 1: Calculating the range of influence spreading. $I(G, \Omega)$

**Input:** $G = (V, E, P)$: Social network;  
$p_{u,v}$: Probability that the virus on edge $(u,v)$ is transmitted from $u$ to $v$;  
$\Omega = (S, N, I, R)$: Set of individuals in four states in a network;  

**Output:** $I(G, \Omega)$: Expected value of the range of virus infection up to time $T$;

**Begin**

$H = 0$;

**For each** $u \in V$ **do**

Initialize $p^W_N(u, 0)$ and $p^W_I(u, 0)$ based on (1);

Initialize $p^W_N(u, 0)$ and $p^W_I(u, 0)$ based on (2);

Initialize $p^W_N(u, 0)$ and $a^W(u, 0)$ based on (3);

$H = H + p^W_I(u, 0) + p^W_N(u, 0)$;

**Endfor** $u$;

$U = N \cup I$;

**For** $t = 1$ **to** $T$ **do**

$U \leftarrow U \cup \{$ neighbor nodes of vertices in $U$ $\}$;

**For each** $u \in U$ **do**

Compute $p^W_N(u, t)$ and $a^W(u, t)$ based on (4) and (5);

Compute $p^W_N(u, t)$, $p^W_I(u, t)$, $p^W_N(u, t)$ and $p^W_I(u, t)$ based on (6) to (9);

$H = H + p^W_I(u, t) + p^W_N(u, t)$;

**Endfor** $u$;

**Endfor** $t$;

$I(G, \Omega) = H$;

**Output**($I(G, \Omega)$);

**End**

Complexity analysis of the algorithm: Set the number of individuals in the network as $n$; the main calculation of the algorithm is the double cycle of $t$ and $u$, and its complexity is $O(Tn)$. Taking $T$ as a constant, the complexity of the algorithm is $O(n)$.

Compared with other methods based on epidemic spreading models, such as the SIR model, our method can block negative influence in a larger range in the network. Let $\Omega_t = (S, N, I, R)$ be the set of states in the SNIR model at time $t$, and let $\Omega_f = (S, I, R)$ be the corresponding set of states in the SIR model. From (10) we can see that:

$$I(G, \Omega_t) = \sum_{t=0}^{T} \left[ \sum_{u \in V} p^W_N(u, t) + \sum_{u \in V} p^W_I(u, t) \right]$$

$$\geq \sum_{t=0}^{T} \sum_{u \in V} p^W_I(u, t) = I(G, \Omega_f)$$

Therefore, the influence range estimated in our model is larger than that in SIR; consequently, the range of blocked negative influence is always larger.

6. Select Contacts to Be Blocked

The first infected individual should be the contact in $W = N \cup I$. We determine the set of edges in contact with individuals in $W$ as $\Gamma(W) = \{e = (u, v) \mid u \in W, v \in V \setminus W, (u, v) \in E\}$. For each edge $e$ in $\Gamma(W)$, we constructed a graph $G_e$ after blocking $e$, then calculated the expected value $I(G_e, \Omega)$ of the range of virus infections in $G_e$. We used the greedy method, taking edge $e$ with the smallest expected value $I(G_e, \Omega)$ step by step, and added set $X$ of contacts that need to be isolated one by one until $|X| = k$ or all contact edges in $W = N \cup I$ were blocked. The framework of the algorithm is given in Algorithm 2.

Complexity analysis of the algorithm: Suppose there are $m$ infected persons, that is, $m = |N \cup I|$, and the maximum number of contacts of the infected person is $d_{max}$, so we determine that the edges of contact with the individual in $W$ do not exceed $m d_{max}$. The number of “while” loops in the algorithm is $O(k \cdot m d_{max})$. Algorithm 1 is called in each loop, so the complexity of Algorithm 2 is $O(k \cdot m d_{max} n)$. Taking $k$, $m$, and $d_{max}$ as constants, the complexity of the algorithm is $O(n)$. 
Algorithm 2: Identify the contacts to be blocked

**Input:** $G = (V, E, P)$: Interactive network;  
$W = N \cup I$: Vertex set of infectious viruses in the network;  
$k$: Threshold of isolation cost;  
**Output:** $X$: Set of contacts requiring isolation;

**Begin**

Construct the set of edges $\Gamma(W) = \{e = (u, v)|u \in W, (u, v) \in E\}$;  
$j = 0; F = E; X = \Phi$;  
**While** $j < k$ and $\Gamma(W) \neq \Phi$ **do**

For each edge $e \in \Gamma(W)$ do  
Construct graph $G_e = (V, F \setminus \{e\})$;  
$H(e) = I(G_e, \Omega)$ /* Call algorithm 1 to calculate the expected value of the range of virus infection $I(G_e, \Omega)$*/

Endfor $e$;  
Select edge $e$ with the smallest $H(e)$ in $\Gamma(W)$;  
$F = F \setminus \{e\}; X = X \cup \{e\}$

Endfor $j$;  
output($X$);

**End**

7. Experiments

7.1. Experimental Environment

The algorithm experiments were based on Windows 10, coded with Python 3.8, and run on an Intel® Core™ i7 CPU, 1.10 GHz. In order to verify the effectiveness of our proposed max expected $H$ values (MaxExpectedH) algorithm, we tested it on multiple groups of real and synthetic networks, and compared its performance with the Random algorithm and MaxDegree algorithm [30].

7.2. Dataset and Parameter Setting

To verify the effectiveness of our MaxExpectedH algorithm, we tested it on four real networks and three synthetic networks with the other two algorithms. The three groups of real network data were dolphins [31], football [32], power [33], and Facebook [34]. Dolphins is a network that describes the family relationship of dolphins, in which each node represents a dolphin, and each edge indicates an association between two dolphins. Football is a network of American football, in which each node represents a college team participating in the 2000 football season, and edges connecting two nodes represent different matches between two teams. Power is a network of the topology of the power grid in the western United States, in which each node represents a power supply facility, and edges represent connections between power supply facilities.

Table 1 shows the main topological characteristics of the four actual networks, where $N$ represents the number of nodes in the network, $|E|$ represents the number of edges among the networks, and $<d>$ represents the average degree of nodes in the network. The value of $H$ is used to indicate the viral infection range, which changes with time $T$, and it is related to the different state probabilities of each node.

| Name     | $N$  | $|E|$  | $<d>$ |
|----------|------|-------|-------|
| Dolphins | 62   | 159   | 2.6   |
| Football | 115  | 613   | 5.3   |
| Power    | 4941 | 6594  | 1.4   |
| Facebook | 4037 | 88,234| 21.9  |
The three synthetic networks are ER, WS small-world, and BA scale-free networks [34]. These are very close to real data, and the data of the ER and WS small-world networks are shown in Figures 2 and 3.

Figure 2. Small-world network.

Figure 3. BA scale-free network.

The probability of $S \rightarrow N$ randomly takes $\alpha \in [0.022, 0.044]$, the probability of $S \rightarrow I$ randomly takes $\beta \in [0.011, 0.034]$, the probability of $I \rightarrow R$ randomly takes $\gamma \in [0.28, 0.35]$, the probability of $N \rightarrow I$ randomly takes $\delta \in [0.012, 0.031]$, and the probability of $R \rightarrow S$ randomly takes $\xi \in [0.2, 0.4]$. The experiments in the real and synthetic datasets are described below.

Table 2 shows the main topological characteristics of three synthetic networks, where $N$ represents the number of nodes in the network, $|E|$ represents the number of edges in different networks, and $<d>$ represents the average degree of nodes in the network.

| Name                        | $N$  | $|E|$  | $<d>$ |
|-----------------------------|------|-------|-------|
| ER network                  | 500  | 3153  | 6.3   |
| WS small-world network      | 500  | 2500  | 5     |
| BA scale-free network       | 300  | 300   | 1     |

7.3. Experiments on Real Datasets and Analysis of Results

7.3.1. Setting $p_{v,u}$ as a Fixed Value

Five nodes were randomly selected as the initially infected nodes, and their initial state values were set. Based on the average output of nodes, the probability $p_{v,u}$ of transmitting the virus from $v$ to $u$ was set as a value fixed to the opposite of the average degree of nodes in different networks.

In the four real networks, when the probability of $N \rightarrow R$ is randomly selected as $\eta \in [0.1, 0.18]$ and $\eta \in [0.18, 0.25]$, the operation of the three algorithms is as follows:

We tested the MaxExpectedH algorithm on four real networks. It can be seen from Figure 4 that when $p_{v,u}$ is set to a fixed value, $H$ changes when $k$ changes, and the change in $H$ was compared with the Random algorithm and the MaxDegree algorithm. As can be seen from Figure 4, in the Football network, when fewer than five edges are removed, the
superiority of the MaxExpectedH algorithm is not obvious. When more than five edges are removed, the $H$ value of the MaxExpectedH algorithm is significantly less than the other two algorithms. In general, in these three real datasets, with an increase in the number of removed edges $k$, the performance of the MaxExpectedH algorithm is better than that of the other two algorithms.

Figure 4. Change in $H$ with change in $k$ when $p_{v,u}$ is set to a fixed value in three real datasets.

7.3.2. Setting $p_{v,u}$ as a Variable Value

According to the different out-degrees of nodes, the probability $p_{v,u}$ of transmitting the virus from $v$ to $u$ was set to the reciprocal of the output of $v$. In the four real networks,
when the probability of $N \rightarrow R$ is randomly selected as $\eta \epsilon [0.1, 0.18]$ and $\eta \epsilon [0.18, 0.25]$, the operation of the three algorithms is as follows:

We tested the accuracy of the MaxExpectedH algorithm on four real networks with $p_{v,u}$ set as a variable value. The change in $H$ with the change in $k$ can be seen in Figure 5. It can be seen in the figure that the MaxExpectedH is overall better than the other two algorithms in the Dolphins and Power networks. In the Football network, when fewer than five edges are removed, the $H$ value of the MaxExpectedH algorithm is close to that of the MaxDegree algorithm. When more than five edges are removed, there is a wider gap between the two algorithms.

Figure 5. Change in $H$ with change in $k$ when $p_{v,u}$ is set to a variable value in three real datasets.
In the Facebook dataset, it can be seen that when fewer than 20 links are removed, the effect of our algorithm is not as good as that of the MaxDegree algorithm. This network link is tight. After removing an edge with the largest $H$ value, the propagation of other edges is frequent and complex. However, when more than 20 edges are removed, the MaxExpectedH algorithm is superior to the other two models.

7.4. Experiments on Synthetic Datasets and Analysis of Results

The efficiency of the proposed algorithm was tested by generating three synthetic datasets: ER network, WS small-world network, and BA scale-free network.

The generated ER network contained 500 nodes, and the link probability was set to 0.025. The generated network $G(V,E)$ contained 3153 edges. Five nodes were randomly selected as the initially infected nodes, and the state value of each node was set. The generated WS small-world network contained 500 nodes, and the generated network $G(V,E)$ contained 2500 edges. Five nodes were randomly selected as the initially infected nodes, and the state value of each node was set. In the small-world network, the value of $T$ was set to three. The generated BA scale-free network contained 300 nodes, and the generated network $G(V,E)$ contained 300 edges. Three nodes were randomly selected as the initially infected nodes, and the state value of each node was set.

7.4.1. Setting $p_{v,u}$ as a Fixed Value

According to the average out-degree of the nodes, the probability $p_{v,u}$ of transmitting the virus from $v$ to $u$ was set as a value fixed to the opposite number of average degree of nodes in different networks. When the probability of $N \to R$ randomly choosing $\eta$ in different ranges, the operation of the three algorithms is as follows:

We compared the three algorithms in the three synthetic networks. It can be seen in Figure 6 that, when $p_{v,u}$ is set to a fixed value, $H$ changes when $k$ changes, and the change in $H$ was compared with the other two algorithms. The figure shows that the MaxExpectedH algorithm performed better than the other two algorithms in the synthetic ER and BA networks. Especially in the ER network, the MaxExpectedH algorithm had a lower $H$ value than the other two algorithms when fewer edges were deleted. In the WS small-world network, the superiority of the MaxExpectedH algorithm was not obvious, because the average degree of nodes in this network was high. In the process of propagation, with an increase in deleted edges, there was still more propagation in the whole network. However, overall, the MaxExpectedH algorithm always achieved a lower $H$ value than the Random and MaxDegree algorithms.

It can be seen in Figure 7 that, when there are more initial nodes, the performance of our algorithm is better than that of the other two algorithms when the number of edges removed reaches about five. This shows that the performance of our algorithm is also ideal when the number of initially infected nodes in the ER network increases.

![Figure 6. Cont.](image-url)
Figure 6. Change in $H$ with change in $k$ when $p_{v,u}$ is set as a fixed value in three synthetic datasets.

Figure 7. Cont.
7.4.2. Setting $p_{v,u}$ as a Variable Value

According to the different out-degrees of nodes, the probability $p_{v,u}$ of transmitting the virus from $v$ to $u$ was set to the reciprocal of the output of $v$. When the probability of $N \rightarrow R$ is randomly taken as $\eta$ with different ranges, the operation of the three algorithms is as follows:

We compared three algorithms in three synthetic networks. It can be seen in Figure 6 that when $p_{v,u}$ is set as a variable value, $H$ changes when $k$ changes, and the change in $H$ was compared with the other two algorithms. Figure 7 shows that the MaxExpectedH algorithm performs better than the other two algorithms in the synthetic ER and BA networks. Especially in the ER network, the MaxExpectedH algorithm had a lower $H$ value than the other two algorithms when fewer edges were deleted. In the process of propagation, with an increase in deleted edges, there was still more propagation in the whole network. However, overall, the MaxExpectedH algorithm always achieved lower $H$ values than the Random and MaxDegree algorithms.

In the WS small-world network, the $H$ value of the MaxExpectedH algorithm was less than the other methods, but the difference was not significant. This is because the average degree of nodes in the WS network was high, and the superiority of MaxExpectedH is not obvious in a dense network.

8. Conclusions and Further Work

Based on the proposed SNIR model, in this paper, we mainly propose using the precise isolation method to minimize the spread of an epidemic. First, by combining the infectious characteristics of COVID-19 and the traditional SIR model, the SNIR model was proposed. Next, based on the established model, the infection range of the virus at different times was calculated according to its different state and diffusion probability. Then, the greedy method was used to select the set of contacts that need to be isolated until the preset value of the contact that needs to be isolated is reached or the infection source is completely blocked. In experiments on three real datasets and three simulated datasets, the results confirm the effectiveness of this algorithm in blocking virus infection.

One weak point of our method is that in a network with dense connections, the precision of our algorithm is not obviously higher than the others. Therefore, in further work, we will consider infected individuals and their adjacent edges as a whole, and the overall cost of isolating infected individuals and their adjacent edges together. The higher the isolation cost of infected individuals and their adjacent edges, the earlier they should be isolated so as to determine the sequence of isolated contacts. We will also consider the known propagation situation and subsequently locate the source of influence. We will look into finding the solution to the problem of maximizing influence as an optimization problem from the perspective of artificial intelligence and reduce the impact of randomness through the approach of linear regression in machine learning.
Author Contributions: Conceptualization, L.C.; methodology, L.C. and K.H.; software, C.D.; validation, Y.D.; formal analysis, L.C.; writing—original draft preparation, C.D.; writing—review and editing, L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [Youth Natural Science Foundation of China] grant number [61906100 and 82004499] and [National Natural Science Foundation of China] grant number [82074580] and [National Key Research and Development Program of China] grant number [2022YFC3502302].

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found by the following links: https://download.csdn.net/download/u012311410/7670521, https://www.csdn.net/tags/NtzaIg5sMTczNzAtYmxvZwO0O0OO00000.html, accessed on 31 August 2022.

Acknowledgments: Thanks to Qi’an Wang for technical guidance on the configuration of the experimental environment.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

- SNIR: Susceptible, non-symptomatic, infectious, and recovered
- SIR: Susceptible, infectious, and recovered
- MaxExpectedH: maximum expected $H$ value

References

1. Chen, B.L.; Jiang, W.X.; Chen, Y.X.; Chen, L.; Wang, R.J.; Han, S.; Lin, J.H.; Zhang, Y.C. Influence blocking maximization on networks: Models, methods and applications. *Phys. Rep.* 2022, 976, 1–54. [CrossRef]

2. Nowzari, C.; Preciado, V.M.; Pappas, G.J. Analysis and Control of Epidemics: A survey of spreading processes on complex networks. *IEEE Control Syst. 2016*, 36, 26–46.

3. Goutsias, J.; Jenkinson, G. Markovian Dynamics on Complex Reaction Networks. *Phys. Rep.* 2013, 529, 199–264.

4. Oupechoux, E.C.; Lelarge, M. How clustering affects epidemics in random networks. *Adv. Appl. Probab.* 2014, 46, 985–1008.

5. Maliyoni, M. Probability of Disease Extinction or Outbreak in a Stochastic Epidemic Model for West Nile Virus Dynamics in Birds. *Acta Biotheor.* 2021, 69, 91–116. [CrossRef]

6. Fried, M.N. Ways of Relating to the Mathematics of the Past. *J. Humanist. Math.* 2018, 8, 3–23. [CrossRef]

7. Saha, S.; Adiga, A.; Prakash, B.A. Approximation Algorithms for Reducing the Spectral Radius to Control Epidemic Spread. In Proceedings of the 2015 SIAM International Conference on Data Mining; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2015; pp. 568–576. [CrossRef]

8. Harling, G.; Onnela, J.P. Impact of degree truncation on the spread of a contagious process on networks. *Netw. Sci.* 2018, 6, 34–53.

9. Matamalas, J.T.; Arenas, A.; Gómez, S. Effective approach to epidemic containment using link equations in complex networks. *Physics* 2018, 4, eaau4212.

10. Mai, V.S.; Battou, A.; Mills, K. Distributed algorithm for suppressing epidemic spread in networks. *IEEE Control Syst. Lett.* 2018, 2, 555–560. [CrossRef]

11. Emoglu, D.A.; Ozdaglar, A.E.; Salehi, A.T. *Networks, Shocks, and Systemic Risk*; Oxford University Press: Oxford, UK, 2015; pp. 569–608. [CrossRef]

12. Rüdiger, S.; Pletzsch, A.; Sagués, F. Epidemics with mutating infectivity on small-world networks. *Sci. Rep.* 2020, 10, 5919. [CrossRef]

13. Wang, Y.; Ma, J.; Cao, J.; Li, L. Edge-based epidemic spreading in degree-correlated complex networks. *J. Theor. Biol.* 2018, 454, 164. [PubMed]

14. Toli, D.; Kleineberg, K. Simulating SIR processes on networks using weighted shortest paths. *Sci. Rep.* 2018, 8, 6562. [CrossRef]

15. Gao, C.; Su, Z.; Liu, J.M.; Kurths, J. Even central users do not always drive information diffusion. *Commun. ACM* 2019, 62, 61–67. [CrossRef]

16. Kianian, S.; Rostamnia, M. An efficient path-based approach for influence maximization in social networks. *Expert Syst. Appl.* 2020, 167, 114168.

17. Su, J.; Chen, L.; Chen, Y.; Li, B.; Liu, W. Minimizing the seed set cost for influence spreading with the probabilistic guarantee. *Knowl.-Based Syst.* 2021, 216, 106797.

18. Jalayer, M.; Azheian, M.; Mehrdad, A. A hybrid algorithm based on community detection and multi attribute decision making for influence maximization. *Comput. Ind. Eng.* 2018, 120, 234–250.

19. Li, D.; Wang, W.; Liu, J.M. Grassroots V5 elites: Which ones are better candidates for influence maximization in social networks? *Neurocomputing* 2019, 358, 321–331. [CrossRef]
20. CaliÓ, A.; Tagarelli, A. Attribute based Diversification of Seeds for Targeted Influence Maximization. *Inform. Sci.* **2020**, *546*, 1273–1305.

21. Sb, A.; Mj, A.; Dkp, B. ComBIM: A community-based solution approach for the Budgeted Influence Maximization Problem. *Expert Syst. Appl.* **2019**, *125*, 1–13.

22. Ju, W.J.; Chen, L.; Li, B.; Liu, W.; Sheng, J.; Wang, Y.W. A new algorithm for positive influence maximization in signed networks. *Inform. Sci.* **2020**, *512*, 1571–1591. [CrossRef]

23. Kuhnle, A.; Alim, M.A.; Li, X.; Zhang, H.; Thai, M.T. Multiplex Influence Maximization in Online Social Networks with Heterogeneous Diffusion Models. *IEEE TCSS* **2018**, *5*, 418–429. [CrossRef]

24. Wang, W.Y.; Zhang, Y.; Yang, F. Time-sensitive Positive Influence Maximization in signed social networks. *Physica A* **2021**, *584*, 126353. [CrossRef]

25. Khubchandani, J.; Sharma, S.; Price, J.H.; Wiblishauser, M.J.; Webb, F.J. COVID-19 Morbidity and Mortality in Social Networks: Does It Influence Vaccine Hesitancy? *Int. J. Environ. Res. Public Health* **2021**, *18*, 9448. [CrossRef] [PubMed]

26. Long, E.; Patterson, S.; Maxwell, K.; Blake, C.; Bosó, R.; Lewis, R.; McCann, M.; Riddell, J.; Skivington, K.; Wilson, R.; et al. COVID-19 pandemic and its impact on social relationships and health. *J. Epidemiol. Community Health* **2022**, *76*, 128–132. [CrossRef]

27. Song, C.; Hsu, W.; Lee, M.L. Temporal Influence Blocking: Minimizing the Effect of Misinformation in Social Networks. In *Proceedings of the 2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, San Diego, CA, USA, 19–22 April 2017. [CrossRef]

28. Ghoshal, A.K.; Das, N.; Das, S. Influence of community structure on misinformation containment in online social networks. *Knowl.-Based Syst.* **2020**, *213*, 106693.

29. Lan, Y.; Zla, B.; Ag, C. Containment of rumor spread in complex social networks. *Inform. Sci.* **2020**, *506*, 113–130.

30. Yu, Y.Y.; Li, S.; Philip, P.E. Edge Deletion Algorithms for Minimizing Spread in SIR Epidemicm Models. *arXiv* 2020, arXiv:2011.11087.

31. Lusseau, D.; Schneider, K.; Boisseau, O.J.; Haase, P.; Slooten, E.; Dawson, S.M. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.* **2003**, *54*, 396–405. [PubMed]

32. Girvan, M.; Newman, M.E.J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826. [CrossRef]

33. Watts, D.J.; Strogatz, S.H. Collective dynamics of ‘small-world networks. *Nature* **1998**, *393*, 440–442. [PubMed]

34. Fu, L.; Shen, Z.; Wang, W.X.; Fan, Y.; Di, Z. Multi-source localization on complex networks with limited observers. *EPL* **2016**, *113*, 18006. [CrossRef]