

Exploring individuals' effective preventive measures against epidemics through reinforcement learning

Ya-Peng Cui(崔亚鹏), Shun-Jiang Ni (倪顺江), and Shi-Fei Shen(申世飞)
Citation: Chin. Phys. B, 2021, 30 (**4**): 048901. DOI: 10.1088/1674-1056/abcfa5
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Modeling and analysis of the ocean dynamic with Gaussian complex network

Xin Sun(孙鑫), Yongbo Yu(于勇波), Yuting Yang(杨玉婷), Junyu Dong(董军宇)†, Christian B\"ohm, and Xueen Chen(陈学恩)
Chin. Phys. B, 2020, 29 (**10**): 108901. DOI: 10.1088/1674-1056/aba27d

Finite density scaling laws of condensation phase transition in zero-range processes on scale-free networks

Guifeng Su(苏桂锋), Xiaowen Li(李晓温), Xiaobing Zhang(张小兵), Yi Zhang(张一)
Chin. Phys. B, 2020, 29 (**8**): 088904. DOI: 10.1088/1674-1056/ab8a41

Influential nodes identification in complex networks based on global and local information

Yuan-Zhi Yang(杨远志), Min Hu(胡敏), Tai-Yu Huang(黄泰愚)
Chin. Phys. B, 2020, 29 (**8**): 088903. DOI: 10.1088/1674-1056/ab969f

Identifying influential spreaders in complex networks based on entropy weight method and gravity law

Xiao-Li Yan(闫小丽), Ya-Peng Cui(崔亚鹏), Shun-Jiang Ni(倪顺江)
Chin. Phys. B, 2020, 29 (**4**): 048902. DOI: 10.1088/1674-1056/ab77fe

Pyramid scheme model for consumption rebate frauds

Yong Shi(石勇), Bo Li(李博), Wen Long(龙文)
Chin. Phys. B, 2019, 28 (**7**): 078901. DOI: 10.1088/1674-1056/28/7/078901

Exploring individuals' effective preventive measures against epidemics through reinforcement learning*

Ya-Peng Cui(崔亚鹏)[1,2,3], Shun-Jiang Ni (倪顺江)[1,2,3,†], and Shi-Fei Shen(申世飞)[1,2,3]

[1] *Institute of Public Safety Research, Tsinghua University, Beijing* 100084, *China*
[2] *Department of Engineering Physics, Tsinghua University, Beijing* 100084, *China*
[3] *Beijing Key Laboratory of City Integrated Emergency Response Science, Beijing* 100084, *China*

(Received 22 September 2020; revised manuscript received 21 October 2020; accepted manuscript online 2 December 2020)

Individuals' preventive measures, as an effective way to suppress epidemic transmission and to protect themselves from infection, have attracted much academic concern, especially during the COVID-19 pandemic. In this paper, a reinforcement learning-based model is proposed to explore individuals' effective preventive measures against epidemics. Through extensive simulations, we find that the cost of preventive measures influences the epidemic transmission process significantly. The infection scale increases as the cost of preventive measures grows, which means that the government needs to provide preventive measures with low cost to suppress the epidemic transmission. In addition, the effective preventive measures vary from individual to individual according to the social contacts. Individuals who contact with others frequently in daily life are highly recommended to take strict preventive measures to protect themselves from infection, while those who have little social contacts do not need to take any measures considering the inevitable cost. Our research contributes to exploring the effective measures for individuals, which can provide the government and individuals useful suggestions in response to epidemics.

Keywords: epidemic simulation, complex networks, reinforcement learning, preventive measures

PACS: 89.75.Hc **DOI:** 10.1088/1674-1056/abcfa5

1. Introduction

Epidemic disease has become one of the most concerned threats to human society. According to statistics of Johns Hopkins University, by 19 June 2020, over 200 countries or territories have reported COVID-19 cases and over 8.46 million confirmed cases and 453 thousand deaths had been reported globally.[1] In addition, only in the United States, it is estimated that there are 31.4 million outpatient visits due to seasonal influenza. Apart from large-scale infections, epidemics can also cause huge economic losses and social panic.[2–4] According to the International Monetary Fund (IMF), the global economy is projected to contract sharply by −3 percent in 2020 due to the COVID-19 pandemic, much worse than that during the 2008–2009 financial crisis.[3] The economic burden caused by seasonal influenza in the United States is estimated to be $87.1 billion.[4]

Aiming to relieve the severe consequences caused by epidemics, scholars and the government have increasingly paid attention to preventive measures against epidemics. It is shown that preventive measures taken by individuals play a crucial role in suppressing the epidemic transmission.[5,6] On the one hand, some scholars concerned medical measures, such as vaccination[7–11] and antiviral drugs.[12–15] For example, Keeling *et al.*[7] reported that mass prophylactic vaccination could reduce the potential for a major epidemic significantly. By comparing the direct and indirect losses, Sartore *et al.*[8] found that vaccination could reduce the overall costs associated with avian influenza epidemics. Ferguson *et al.*[13] modeled the effects of antiviral drugs on influenza and found that antiviral drugs would be useful for epidemic elimination. On the other hand, researchers also paid attention to nonmedical measures.[15–20] Scott *et al.*[17] revealed that people voluntarily can quarantine themselves to avoid infection when the plague occurs. During the influenza pandemic, individuals tend to stay away from congregated places.[16] Lau *et al.*[18] discovered that wearing face masks were widespread in severe acute respiratory syndrome (SARS) affected areas. Ma *et al.*[21] reported that limiting individuals' movement to high-layer nodes which mean large cities can suppress the epidemic transmission.

Given different preventive measures, scholars also concern individuals' selection process of these measures. Bauch *et al.*[22] proposed a game theory-based model and found that if individuals decide whether to vaccinate according to self-interest, the epidemic transmission cannot be suppressed effectively. They further indicated that there exists a conflict between self-interest and group interest.[23] Researchers also found that individuals are not always rational and they may make decisions by imitating others.[24–27] For example, Shi *et al.*[24] explored individuals' voluntary vaccination behav-

*Project supported by the National Key Technology Research and Development Program of China (Grant No. 2018YFF0301000) and the National Natural Science Foundation of China (Grant Nos. 71673161 and 71790613).

†Corresponding author. E-mail: sjni@tsinghua.edu.cn

© 2021 Chinese Physical Society and IOP Publishing Ltd

ior considering bounded rationality. Other factors influencing individuals' vaccination behavior, such as individuals' adaptability[28] and human contact structure,[29–31] were also investigated previously. Besides the selection process of vaccination, Reluga[32] studied the game process of social distancing in response to epidemics and found that social distancing is most beneficial to individuals for basic reproduction numbers around 2. Boven *et al.*[33] analyzed individuals' selection process when people took antiviral drugs to response to epidemics.

In those previous studies, only one preventive measure was incorporated at one time and individuals decided whether or not to take this measure. However, in real world, when epidemics happen, an individual may choose to get vaccinated, as well as wear face masks or take antiviral drugs. Therefore, new models need to be proposed to bridge the gap. In this paper, we introduce effective contact rate (ECR) to represent these different preventive measures. Specifically, the smaller ECR refers to the less effective contact process between individuals, which corresponds to stronger preventive measures. For example, when an individual choose to get vaccinated as his/her epidemic preventive measure, his/her social contacts with others will be ineffective because he/she will not get infected through the contacts. However, if wearing face masks is adopted by an individual, the contact process will still work while the efficiency will be reduced. As a result, we can set ECR as 0 and 0.8 (or other values between 0 and 1) to represent getting vaccinated and wearing face masks respectively.

In epidemic scenarios, especially for epidemics that happen repeatedly such as seasonal influenza, individuals' learning behavior plays a key role in the selection process of preventive measures. Individuals will learn from their historical experience and choose epidemic preventive measures adaptively, which is seldom studied by previous research. Therefore, in this paper, a reinforcement learning-based (RL-based) model is proposed to explore the effective preventive measures for different people. Through reinforcement learning process, individuals try to obtain the most effective preventive measure for themselves. In Section 2, we present the RL-based model in detail. Extensive simulations are conducted and the results are described in Section 3. Discussion and conclusions are in Sections 4 and 5.

2. Method

In this section, we propose a reinforcement learning-based (RL-based) model to investigate individuals' selection process of preventive measures against flu-like seasonal epidemics. Through reinforcement learning process, each individual eventually obtains the most effective preventive measure in response to epidemics. Specifically, the model includes two stages. In the first stage is the reinforcement learning process where individuals decide which preventive measure to take before an epidemic season begins. The second stage is the epidemic transmission process on complex networks. In the following subsection, the two stages are presented in detail.

2.1. Reinforcement learning process

In this paper, different values are assigned to ECR to represent different epidemic preventive measures. We set $S = [0, 0.2, 0.4, 0.6, 0.8, 1.0]$ to describe individuals' decision space. The smaller value refers to stronger preventive measures. For example, $S = 0$ describes that the contact process between individuals is absolutely ineffective, which means that strict prevention measures like vaccination or isolation are taken; while $S = 1.0$ indicates that the contact process is effective, representing no preventive measures are taken. Each individual will select a prevention measure from decision space S before an epidemic season begins.

Considering the impact of historical experience, reinforcement learning mechanism is introduced to explore individuals' effective preventive measures against epidemics. Specifically, reinforcement learning with ε-greedy algorithm[34] is adopted where individuals choose a prevention measure randomly with probability ε and choose the prevention measure with highest playoff with probability $1-\varepsilon$. The process can be expressed as follows:

$$S_i(n) = \begin{cases} s | s \in S, & \text{with probability } \varepsilon, \\ \arg\max_s Q_i^s(n-1), & \text{with probability } 1-\varepsilon, \end{cases} \tag{1}$$

where $S_i(n)$ is preventive measure adopted by individual i in season n; and s refers to a random measure in decision space S, representing a random prevention measure taken by individuals. $Q_i^s(n-1)$ indicates the average playoff of measure s taken by individual i until season $n-1$, and ε is the reinforcement learning parameter.

Once an individual selects his/her preventive measures, he/she must afford cost for the preventive measure, as well as obtains reward. The cost includes not only the money to pay but also side effects that the preventive measure caused, and the reward refers to the reduction of infection risk. For example, if a person decide to vaccinate in advance, he/she must pay for the vaccination and accept the possible side effects caused by vaccination. At the same time, he/she will hardly get infected when the epidemic season ends. Moreover, the stronger the preventive measures are, the more cost an individual has to afford. One has to balance the cost and the reward to make proper decisions. As soon as the epidemic season ends, individuals will evaluate the efficiency of their preventive measures and obtain the playoff in this season. The playoff of an

individual can be expressed as

$$U_i(n) = \begin{cases} 1 - c(1 - S_i(n)), & \text{if individual } i \text{ is not infected,} \\ 0 - c(1 - S_i(n)), & \text{if individual } i \text{ is infected .} \end{cases} \quad (2)$$

Here $U_i(n)$ is the playoff of individual i in season n; c represents the relative cost of strict preventive measures over infection. If strict preventive measures are adopted by individuals, they will not get infected. Considering the strong preventive measures indicate high cost, $c(1 - S_i(n))$ is used to describe the relative cost of preventive measure taken by individual i in season n. Individuals will obtain reward 1 if they do not get infected. Instead, there is no reward when they are infected.

At the end of the epidemic season, individuals will update the playoffs of different preventive measures,

$$Q_i^s(n) = \begin{cases} \dfrac{Q_i^s(n-1)\,(k_i^s - 1) + U_i(n)}{k_i^s}, & \\ \qquad \text{if } s \text{ is taken in season } n, & \\ Q_i^s(n-1), \quad \text{if } s \text{ is not taken in season } n, & \end{cases} \quad (3)$$

where $Q_i^s(\mathrm{n})$ is the average playoff of measure s taken by individual i until season n; and k_i^s refers to the times of measure s chosen by individual i until season n and $U_i(n)$ means the playoff of individual i in season n. Algorithm 1 demonstrates individuals' selection process based on reinforcement learning mechanism.

Algorithm 1 The reinforcement learning algorithm ( )

Input: $U_i(n)$, k_i^s, $Q_i^s(n-1)$, relative cost c, ε
Output: individuals' preventive measures $S_i(n)$

```
for i = 1 → N do  // for N individuals
    if a random number <ε then
        S_i(n)  ← randomly choose a prevention measure from S;
    else // choose the prevention measure with the highest playoff;
        S_i(n)  ← argmax_s Q_i^s(n − 1);
    end if
end for
Simulate epidemic spread process; // presented in 2.2 in detail
Update Q_i^s(n) based on Eqs. (2) and (3)
```

2.2. Epidemic spread model

In this paper, the susceptible-infected-recovered (SIR) model.[35] is adopted to simulate the epidemic spread process. Individuals can be one of the three states: susceptible (S), infected (I) and recovered (R). Individuals in state S are not infected while those in state I are infected and will cause further infections through contacts with others. The individuals in state R will not infect others and cannot be infected in the current epidemic season. In addition, complex networks are generated to represent the human contact structure where nodes represent the individuals and edges refer to social contacts among individuals.

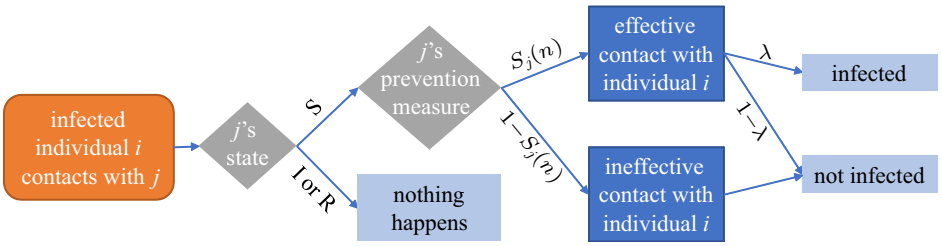


Fig. 1. Diagrammatic description of the epidemic spread process. When susceptible individual j contacts with an infected individual i, the contact process can be effective or ineffective depending on the prevention measure taken by individual j. If the contact process is effective, individual j will get infected with probability λ.

Initially, an individual is randomly selected as the infection source (i.e., set it as state I). At each time step, nodes in state I contact with their neighbors, which represents the individuals' social contacts with others in daily life. In the epidemic season, individuals will take different preventive measures in response to the epidemic, which impacts the efficiency of the contact process. Specifically, when an infected individual i contacts with individual j, nothing happens if individual j is not in state S. Otherwise, according to the preventive measure adopted by individual j, the contact process works with a probability $S_j(n)$, that is, the stronger preventive measure indicates that infection will spread through the contact with a lower probability. It will lead to further infections only when the contact process is effective. Then, individual j may get infected with a probability λ. All infected individuals also recover with a probability μ at each time step. The whole process evolves until there are no infected individuals. The epidemic spread process is shown in Fig. 1.

3. Results and discussion

In this section, a series of epidemic simulations on different networks are conducted to explore individuals' effective preventive measures against epidemics. Studies have revealed that social networks have the characteristics of power-law degree distribution and high clustering.[36–39] Therefore, in this paper, scale-free (SF) networks[40] and small-world (WS) networks[41] are used to depict the population contact structure, on which the epidemic spreads. The formers show the power-law degree distribution, while the latter ones characterize high clustering. At the same time, we also generated the Holme–Kim (HK) network,[42] which has both the characteristics of power-law degree distribution and high clustering. Specifically, the concerned parameters are set as follows.

Network parameter Each network has 10000 nodes,

representing 10000 individuals. The scale-free (SF) networks are generated based on the uncorrelated configuration model,[40] where the degree distribution follows power-law $p(k) = \varphi k^{-\gamma}$ and we set $\gamma = 3.0$ in our simulations. We also generate small-world (WS) networks with $p_w = 0.2$ according to the model proposed by Watts and Strogatz.[41] HK networks are generated by the algorithm proposed by Holme and Kim,[42] and we set $m = 4$ and $p = 0.5$.

Reinforcement learning parameters Similar to Ref. [24], the parameter of ε-greedy algorithm is set as $\varepsilon = 0.05$. Initially, individuals' playoffs are set as $Q_i^s(0) = 0$, which means that at the beginning, individuals have no historical experience about the epidemic and will choose an epidemic preventive measure randomly. In the following seasons, the playoffs of different measures are updated according to Eqs. (2) and (3).

Epidemic spread parameters We set epidemic infection rate $\lambda = 0.6$ and recover rate $\mu = 0.2$ to guarantee the epidemics can spread all over the networks.

Simulation parameters For each scenario, we carry out simulations for 5000 seasons and the epidemic infection scale and individuals' selection data are averaged over the last 200 seasons when epidemic transmission process and individuals' selection process are stable enough. Moreover, we randomly generate 50 SF, WS and HK networks respectively and all the results are averaged over the 50 networks.

3.1. The final infection scale

We first investigate the effect of relative cost c on the final infection scale. Relative cost c refers to the cost that individuals must afford when taking epidemic preventive measures. It not only includes money to pay but also physical, mental or social effects. For example, when a person decides to vaccinate in advance, he/she must pay for the vaccination and accept the possible side effect of the vaccination, as well as the time he/she spends to go to hospitals. As shown in Fig. 2, our simulation results indicate that epidemic infection scale increases as the relative cost c becomes larger. The difference between simulations on different networks is small. When c is small ($c < 0.5$), the results obtained on SF, WS and HK networks are almost the same. The difference gets larger as c continues to increase. With a fixed cost c (except $c = 1.0$), the final infection scale on SF network is the largest and that on HK network is the smallest.

The phenomena stated above can be explained as follows. The clustering coefficient of HK network is the largest and that of SF network is the smallest. Higher clustering coefficient means that there are more local communities on the networks, where the epidemic transmission process will be limited considering that after people in these communities are infected, social contacts are ineffective and cannot cause more infections. Therefore, the infection scale is the smallest on HK networks and the largest on SF networks. According to the results, we can see that there are two ways to suppress the epidemic transmission: increasing clustering coefficient of network or reducing relative cost c. In real world, individuals' contact structure (network structure) can hardly change, which means that it is not realistic to control epidemics by changing human contact structure. Fortunately, the spread of epidemics can also be suppressed by reducing the cost of preventive measures, regardless of the human contact structure. The results suggest that the government can prevent epidemic infections by providing preventive measures with low lost or reducing the cost of preventive measures, such as making vaccines safer and lowering the price.

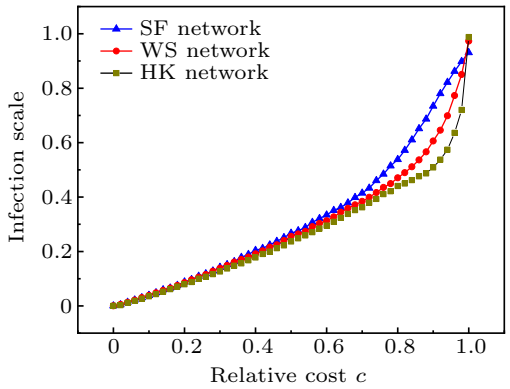


Fig. 2. The infection scale as a function of relative cost c. The infection scale increases as the relative cost c grows. With the same relative cost c (except $c = 1.0$), the infection scale is the largest on SF networks and the smallest on HK networks.

3.2. The fractions of different preventive measures

In this subsection, we try to figure out why the infection scale gets larger when relative cost c increases. We calculate the fractions of individuals adopting different preventive measures with different relative cost c, which is presented in Fig. 3. From Fig. 3(b), on the WS networks, the number of individuals adopting strict preventive measure (ECR = 0) decreases sharply and then remains stable when c increases from 0. However, if c continues to increase ($c > 0.8$), the number quickly dwindles to 0. On the contrary, the number of individuals taking no measures (ECR = 1) always increases as c becomes larger. However, on SF and HK networks [see Figs. 3(a) and 3(c)], the fractions fluctuate when the relative cost$c$ is large. Even though the exact values of the fractions are different in SF, WS and HK networks, the trends of the fractions are the same. We also find that after reinforcement learning, individuals tend to take either strict preventive measures (ECR = 0) or no measures (ECR = 1). Intermediate measures (ECR = 0.2–0.8) are seldom adopted especially when relative cost are large.

When an individual chooses a preventive measure, he/she must afford the cost of the measure. Therefore, once the cost of preventive measure exceeds the threshold that an individual can afford, he/she tends to take no measures. As a result, more

individuals are exposed to the infections, leading to a larger infection scale as shown in Fig. 2. The reason for individuals' preference to strict preventive measures (ECR = 0) and no measures (ECR = 1) can be explained as follows. If a person adopts an intermediate preventive measure, such as ECR = 0.6, he/she still has a possibility to get infected through social contacts. The stronger the preventive measures, the smaller the possibility and the higher the cost. Consequently, he/she has to suffer the risk of infection and must afford the cost of the measure. However, if individuals take strict preventive measures (ECR = 0), they just afford the cost of the measure. Similarly, there is only risk of infection when nothing is done (ECR = 1). During the COVID-19 pandemic, people's preventive measures against the infectious disease also tend to be polarized. For example, in some East Asian countries such as China, people tend to adopt very strict measures to avoid getting infected, such as wearing facial masks and self-isolation at home. However, in some countries, although the number of confirmed cases is already very large, people still refuse to wear masks and even hold large parties as usual, which is consistent with the results obtained by our simulation.

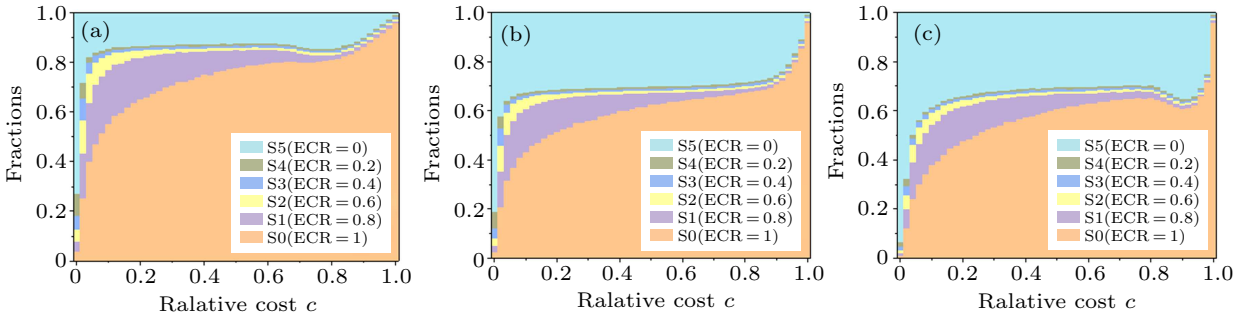


Fig. 3. The fractions of individuals adopting different preventive measures with different relative cost c. Here (a), (b) and (c) show the results on SF, WS and HK networks respectively. As relative cost c increases, more individuals choose not to take measures.

Interestingly, on SF and HK networks, the number of individuals taking strict preventive measures fluctuates slightly when c is larger than 0.7, which does not exist on WS networks. Due to the heterogeneity of networks, SF and HK networks have more hub nodes which can facilitate the spread process but also can be infected easily. When c is large, many individuals tend to take no measures, which makes the hub nodes get infected more easily. Therefore, more hub nodes take strict preventive measures again, leading to a slight increase when c continues to grow. However, if the cost continues to increase, these hub nodes will also choose not to take preventive measures. Consequently, the number of individuals adopting strict preventive measures decreases. Specifically, if the relative cost increases to 1, that is, the consequences of infection are comparable to the cost of preventive measures, all individuals choose to take no measures regardless of the network structure as shown in Fig. 3. The fluctuations in the fraction of taking measure S5 in Fig. 3(c) are meaningful, which indicates that in some cases, although the cost of preventive measures increases, the fraction of taking preventive measures in the population will rise. For example, although the price of facial masks rose rapidly in China, many people still chose to wear facial masks in response to COVID-19.

3.3. The performance of different individuals

We further explore the effective preventive measures of different individuals against the epidemic. Figure 4 demonstrates the selection behaviors of different individuals on SF, WS and HK networks. From Fig. 4, with the increase of relative cost c, more and more individuals do not take measures (S0) [see Fig. 4(a)], while fewer and fewer individuals take strict preventive measures (S5) [see Fig. 4(b)], no matter whether the degree of individuals is high or low. However, the behaviors of individuals with high or low degree are quite different. The fraction of individuals taking strict preventive measures among high degree individuals is much greater than that among low degree individuals as shown in Fig. 4(b). Similarly, compared to low-degree individuals, fewer individuals among high-degree individuals choose not to take measures [see Fig. 4(a)]. Moreover, when c is smaller than 0.1, the fractions of individuals taking no measures and strict preventive measures among low degree individuals change faster than that among high degree individuals, which means that low-degree individuals are more sensitive to the cost of preventive measures compared to high-degree individuals.

The results on SF, WS and HK networks are similar. However, the changing process of the fractions is quite different on the HK network. For example, on HK networks, the fraction of taking measure S0 among high degree individuals first increases slowly and then rapidly as c grows [see the square curve in Fig. 4(a)], but it first increases rapidly and then slowly on SF and WS networks [see the triangle and circle curves in Fig. 4(a)]. At the same time, compared with the SF and WS networks, on the HK networks, more individuals take strict measures no matter whether the degree is high or low. This can be explained as follows. On the HK networks, there

are more communities. When infectious disease spreads to a community, individuals in this community have large probabilities to get infected due to the frequent social contacts. Thus, individuals in HK networks are not only affected by its neighbors but also individuals in their communities. In order to protect themselves from getting infected, more individuals choose to take strict preventive measures. However, on SF and WS networks, individuals are mainly affected by their neighbors and the infection probability is relatively low, so more people do not take measures.

To further understand the selection process of different individuals, we calculate the fractions of individuals with the same degree who take different preventive measures as shown in Figs. 5 and 6. With the degree increases, more and more individuals adopt strict preventive measures (S5) and fewer and fewer individuals take no measures (S0). Specifically, among the individuals with degree larger than 16, almost all of them choose to take strict preventive measure. We also discover that measures S0 and S5 are the most frequently adopted by individuals. As shown by the dashed lines in Fig. 5, the fraction of individuals taking measures S0 or S5 is up to 80% or more as the degree of individuals increases. Moreover, when the cost of preventive measures changes, only the specific fraction values change, but the principle of individuals' selection behavior remains the same [see Fig. 6].

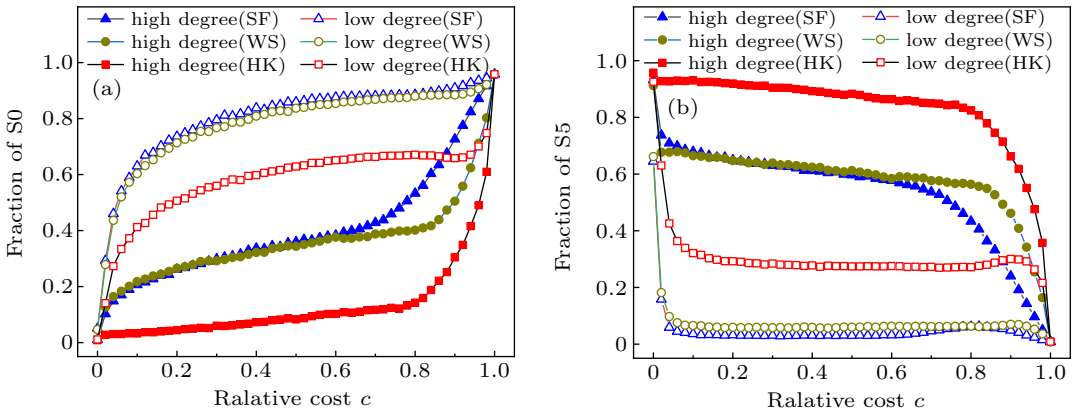


Fig. 4. The performance of individuals with high and low degree on SF, WS and HK networks. Here (a) and (b) show the fraction of individuals taking no measures (S0) and strict preventive measures (S5), respectively.

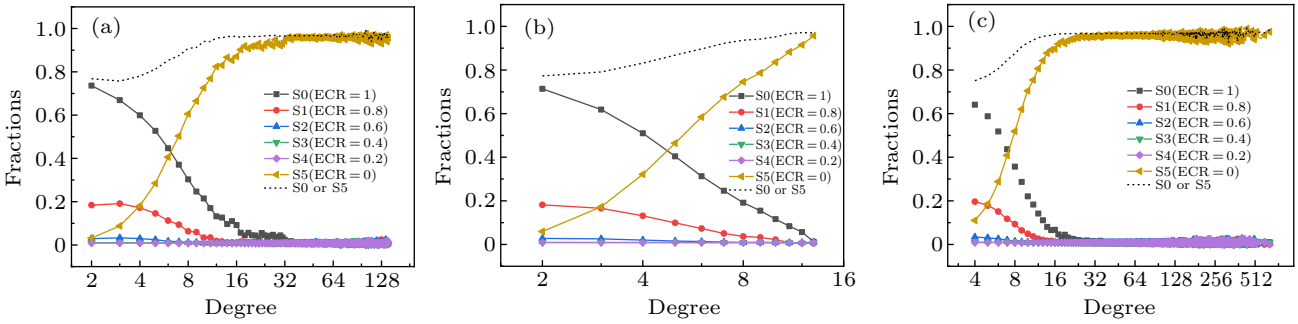


Fig. 5. The individuals' selection behavior changes with individuals' degree. Here (a), (b) and (c) present the simulations on SF, WS and HK networks respectively. The larger the degree of the individuals, the more likely to take strict preventive measures. We set relative cost $c = 0.2$.

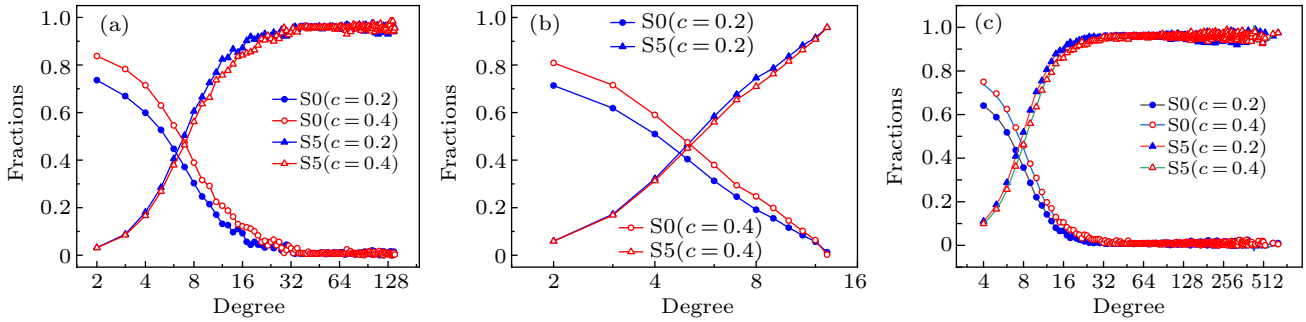


Fig. 6. The individuals' selection behavior under different relative cost c. Here (a), (b) and (c) present the simulations on SF, WS and HK networks respectively. The relative cost c has slight effects on the whole trend.

We explain the aforementioned phenomena as follows. Individuals with high degree refer to people who contact with others frequently in daily life. As a result, these individuals are more likely to be infected when epidemics spread. Learning from the historical experience and balancing the risk of infection and the cost of preventive measures, individuals with high degree tend to take strict preventive measures while individuals with low degree prefer not to take measures because they are seldom infected. Based on reinforcement learning mechanism, in our simulations, each individual learns from his experience and tries to obtain the best preventive measure in response to epidemics. The results are useful for individuals to deal with epidemics in real world. Individuals with higher degree who contact with others more frequently in daily life are highly recommended to take stronger preventive measures to protect themselves from infection. However, individuals who have little social contact with others can take certain measures or no measures considering the cost of the preventive measures cannot be ignored.

4. Discussion

In this paper, reinforcement learning is introduced to obtain the most effective preventive measures for each individual. The effective contact rate (ECR) is proposed to represent different epidemic preventive measures. The epidemic transmission has also been simulated on different complex networks. This research contributes to understanding individuals' selection process of preventive measures, which can help the government and individuals better cope with the epidemics.

In previous research, individuals' historical experience was seldom studied, which is very important for dealing with epidemics, especially the seasonal influenza-like infectious diseases. Some researchers used game theory to study whether a preventive measure would be adopted by individuals,[22,23,32,33] but the situation in which people may take various measures in real life is insufficiently considered. This paper proposes a reinforcement learning-based (RL-based) epidemic transmission model to simulate individuals' selection process of different preventive measures based on their historical experience. Through the reinforcement learning process, the most effective preventive measures for each individual can be obtained, which is very useful for the government and individuals in response to epidemics. Moreover, most of the previous studies only focused on one kind of epidemic preventive measures, such as vaccination, antiviral drugs, isolation, and so on.[8,13,17] However, in real life, the measures that individuals can take are varied and even combined measures can be adopted at the same time. Considering these preventive measures are essentially affecting the contact process between individuals, in this paper, the effective contact rate (ECR) is proposed to represent the different measures. The stricter the measures, the lower the effective contact rate.

In this paper, seasonal influenza-like epidemics are the main concern considering this kind of epidemics occur every year and take away countless lives. However, for a new type of epidemics, such as COVID-19, reinforcement learning mechanism may not be appropriate due to the lack of coping experience. Under this circumstance, how to predict individuals' effective preventive measures against epidemics needs further study. In this paper, it is assumed that whether an individual chooses a preventive measure depends on its benefits and costs. If the playoffs of the measure is large, individuals prefer to take that measure. Therefore, the results we obtained are smooth and monotonic. However, in real life, there may be a threshold for individuals' perception of the cost of the preventive measures. For example, when the cost is lower than a certain value, it may be considered negligible, and when the cost is greater than a certain value, it cannot be tolerated. In this scenario, the simulation results may not be linear and smooth. This threshold is too difficult to determine, so we consider introducing it into the further model. In addition, there are many other factors that affect people's choice of preventive measures. For example, in some East Asian countries like China, Japan and South Korea, almost everyone wears face masks, but in some developed countries, people generally resist wearing facial masks. This phenomenon is similar to the findings in this article, that is, individuals either take strict preventive measures or no measures in response to seasonal influenza-like epidemics. Nevertheless, its determinants may be more complicated, which remains to be studied further. In short, how to predict individuals' selection process of preventive measures is of great significance to predict the epidemic trend.

5. Conclusions

Individuals' historical experience plays an important role in coping with epidemics, which has rarely been paid attention to in previous studies. In real life, individuals may choose different preventive measures in response to epidemics, but most of the previous studies only focused on one preventive measure at one time. Therefore, aiming to bridge the gaps above, this paper proposes effective contact rate (ECR) to represent different preventive measures. Reinforcement learning mechanism is also introduced to study the effects of individuals' historical experience and explore individuals' effective preventive measures against epidemics. Through extensive numerical simulations, we find that:

(i) In addition to the effect of network structure, the cost of preventive measures impacts the epidemic transmission process significantly. When the cost of preventive measures is lower than the threshold that an individual can afford, the in-

dividual's initiative to take preventive measures is enhanced, thus effectively inhibiting the spread of epidemics.

(ii) After reinforcement learning, individuals tend to take either strict preventive measures (S5) or no measures (S0). Intermediate measures are seldom adopted especially when the cost preventive measures are large. Regardless of the degree of individuals, the fraction of individuals taking S0 or S5 is up to 80% or more.

(iii) Individuals with high or low degree perform quite differently in the epidemic spreading process. Individuals with high degree prefer strict epidemic preventive measures while those with low degree tend to take no measures. The willingness of an individual to take strict preventive measures gets stronger as the degree of the individual grows because individuals with higher degree are more likely to be infected during the epidemic transmission. As a result, they are more inclined to take strict measures to protect themselves from infection.

The preventive measures adopted by individuals play a crucial role in suppressing the epidemic transmission. In this paper, it is assumed that people choose different preventive measures based on the costs and benefits of the measures. Some other cultural and social factors are not considered. Although our model is a simplification of the real world, the results of the model still have important significance, considering that most people still balance the benefits and costs to choose different measures. It is the premier goal for the government to get individuals to take preventive measures when epidemics occur. In this research, the government is recommended to provide low-cost preventive measures because more individuals will choose to take preventive measures after balancing the risk of infection and the cost of the measures, leading to a small infection scale. In addition, in this paper, each individual tries to learn from his/her historical experience to obtain the best preventive measure by reinforcement learning mechanism, which is useful for individuals in response to epidemics. Individuals who have frequent contacts with others in daily life are highly recommended to take stronger preventive measures to protect themselves from infection. In summary, our research contributes to exploring the effective measures for individuals, which can provide the government and individuals useful suggestions in response to epidemics.

References

[1] Dong E, Du H and Gardner L 2020 *Lancet Infect. Dis.* **20** 533
[2] Helbing D, Ammoser H and Kühnert C 2006 *Extreme Events in Nature and Society* (Berlin: Springer)
[3] IMF 2020 *World Economic Outlook April 2020* The Great Lockdown, Report
[4] Molinari N A M, Ortega-Sanchez I R, Messonnier M L, Thompson W W, Wortley P M, Weintraub E and Bridges C B 2007 *Vaccine* **25** 5086
[5] Sahneh F D, Chowdhury F N and Scoglio C M 2012 *Sci. Rep.* **2** 632
[6] Wang Z, Andrews M A, Wu Z X, Wang L and Bauch C T 2015 *Phys. Life Rev.* **15** 1
[7] Keeling M J, Woolhouse M E J, May R M, Davies G and Grenfell B T 2003 *Nature* **421** 136
[8] Sartore S, Bonfanti L, Lorenzetto M, Cecchinato M and Marangon S 2010 *Poult. Sci.* **89** 1115
[9] Tildesley M J, Savill N J, Shaw D J, Deardon R, Brooks S P, Woolhouse M E, Grenfell B T and Keeling M J 2006 *Nature* **440** 83
[10] Galvani A P, Reluga T C and Chapman G B 2007 *Proc. Natl. Acad. Sci. USA* **104** 5692
[11] Hu Z L, Liu J G and Ren Z M 2013 *Acta Phys. Sin.* **62** 218901 (in Chinese)
[12] Ferguson N M, Mallett S, Jackson H, Roberts N and Ward P 2003 *J. Antimicrob. Chemother.* **51** 977
[13] Ferguson N M, Cummings D A T, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S and Burke D S 2005 *Nature* **437** 209
[14] Ferguson N M, Cummings D A, Fraser C, Cajka J C, Cooley P C and Burke D S 2006 *Nature* **442** 448
[15] Longini I M, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings D A T and Halloran M E 2005 *Science* **309** 1083
[16] Crosby A 1990 *America's forgotten pandemic: The influenza of 1918* (Cambridge: Cambridge University Press)
[17] Scott S and Duncan C J 2001 *Biology of Plagues: Evidence from Historical Populations* (Cambridge: Cambridge University Press)
[18] Lau J T F, Yang X, Pang E, Tsui H Y, Wong E and Wing Y K 2005 *Emerg. Infect. Dis.* **11** 417
[19] Jiang C 2007 *Int. J. Syst. Sci.* **38** 451
[20] Funk S, Salathe M and Jansen V A A 2010 *J. R. Soc. Interface* **7** 1247
[21] Ma X, Cui Y P, Yan X L, Ni S J and Shen S F 2019 *Chin. Phys. B* **28** 128901
[22] Bauch C T, Galvani A P and Earn D J D 2003 *Proc. Natl. Acad. Sci. USA* **100** 10564
[23] Bauch C T and Earn D J D 2004 *Proc. Natl. Acad. Sci. USA* **101** 13391
[24] Shi B Y, Liu G L, Qiu H J, Wang Z, Ren Y Z and Chen D 2019 *Physica A* **515** 171
[25] Bauch C T 2005 *Proc. R. Soc. B* **272** 1669
[26] Fu F, Rosenbloom D I, Wang L and Nowak M A 2011 *Proc. Royal Soc. B* **278** 42
[27] Zhang H F, Shu P P, Wang Z, Tang M and Small M 2017 *Appl. Math. Comput.* **294** 332
[28] Vardavas R, Breban R and Blower S 2007 *PLoS Comput. Biol.* **3** e85
[29] Perisic A and Bauch C T 2009 *BMC Infect. Dis.* **77** 15
[30] Perisic A and Bauch C T 2009 *PLoS Comput. Biol.* **5** e1000280
[31] Cornforth D M, Reluga T C, Shim E, Bauch C T, Galvani A P and Meyers L A 2011 *PLoS Comput. Biol.* **7** e1001062
[32] Reluga T C 2010 *PLoS Comput. Biol.* **6** e1000973
[33] van Boven M, Klinkenberg D, Pen I, Weissing F J and Heesterbeek H 2008 *PLoS One* **3** e1558
[34] Sutton R and Barto A 1998 *Reinforcement Learning: An Introduction* (Cambridge: MIT Press)
[35] Kermack W O and McKendrick A G 1991 *Bull. Math. Biol.* **53** 33
[36] Watts D J 2004 *Small Worlds: the Dynamics of Networks between Order and Randomness* (Princeton: Princeton University Press)
[37] Newman M E 2000 *J. Stat. Phys.* **101** 819
[38] Strogatz S H 2001 *Nature* **410** 268
[39] Xu X L, Liu C P and He D R 2016 *Chin. Phys. Lett.* **33** 048901
[40] Catanzaro M, Boguna M and Pastor-Satorras R 2005 *Phys. Rev. E* **71** 027103
[41] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[42] Holme P and Kim B J 2002 *Phys. Rev. E* **65** 026107