scientific reports

OPEN

A computational model of epidemic process with three variants on a synthesized human interaction network

Seprianus✉, Nuning Nuraini & Suhadi Wido Saputro

Virus mutations give rise to new variants that cause multiple waves of pandemics and escalate the infected number of individuals. In this paper, we develop both a simple random network that we define as a synthesized human interaction network and an epidemiological model based on the microscopic process of disease spreading to describe the epidemic process with three variants in a population with some features of social structure. The features of social structure we take into account in the model are the average number of degrees and the frequency of contacts. This paper shows many computational results from several scenarios both in varying network structures and epidemiological parameters that cannot be obtained numerically by using the compartmental model.

All viruses change over time through mutations and give rise to new variants. Although most of them have little biological significance, a small number of them, called 'variants of concern', appear to cause multiple waves of pandemic and escalate the infected number of individuals[1,2]. New variants not only result in increased transmissibility, morbidity, and mortality but also cause reinfection in previously infected and recovered individuals, and ineffectiveness of vaccines[3]. For example, SARS-CoV-2 the virus causing COVID-19 emerged originally in Wuhan, China, in December 2019. Several novel variants of SARS-CoV-2 emerged and took over the world since then. The alpha variant was initially detected in the UK in September 2020 and is more transmissible than the original strain. The beta variant was detected in South Africa in November 2020 and is more transmissible than previous strains. It can reinfect those infected with the original strain. The gamma variant was detected in November 2020 in Brazil. It can also reinfect and be more infectious. Delta variant was initially detected in India in October 2020. It is more transmissible, infectious, and lethal[4,5]. Omicron variants were first detected in November 2021 in Africa and since December 2021 confirmed cases have been reported from many countries in Europe[6]. To date, caused by COVID-19, more than 590 million confirmed cases and at least 6.5 million deaths have been reported worldwide[7]. Although the virus is less fatal than it was in 2020, it is likely to continue to undergo mutation and produce future variants. It forces all countries in the world to enter the endemic stage of COVID-19[8].

Virus mutation is natural to occur. Once a virus invades a host (human or animal), it replicates itself to produce more viruses. In this process, some errors occur. We call the error a mutation and new variants are expected to rise[9]. As a consequence, we can not stop the virus from mutating as long as the virus infects[10,11]. Unfortunately, in this current era, globalization has made the world interconnected. It increases the amount, frequency, and speed of population mobility and in turn, causes the range of disease spread to become borderless. Every person in any country becomes vulnerable when a pandemic attacks a certain region. As a result, infectious diseases caused by mutating viruses spread over the world at an unprecedented frequency and speed[12].

Nowadays, although vaccination is the most effective way to control the epidemic spreading[13], no vaccine is 100% effective in the prevention of the disease spread[3]. Moreover, when the vaccination rate is low both in production and administration, this provides enough time for the virus to mutate and become vaccine-resistant[14]. As a consequence, the emergence of new variants can not be prevented[15]. The low rate of vaccination is caused by many factors such as uneven distribution of vaccination, and a lack of consideration of vaccination scenarios[16].

Most epidemiological models use a compartmental model to describe the epidemic process where each individual is supposed to be one of the compartments such as susceptible (S), infected (I), or recovered (R), and

Department of Mathematics, Institut Teknologi Bandung, Bandung, Indonesia. ✉email: 30119003@mahasiswa.itb.ac.id

commonly follow a general susceptible-infected-recovered (SIR) model[17]. However, most of these models assume that the population is "fully mixed' or "mass-action approximation", meaning that every individual has an equal chance, per unit of time, of coming into contact with every other member of the population they belong[18,19]. As a result, these models are not a realistic representation of the disease spread in a human population. Since in practice, each individual has a finite set of connections and frequency of contacts that the infection can pass through. It means that an infected individual does not have an equal probability of infecting a small number of others in the population. Therefore, a network will be used to capture the human interactions and how the disease spreads at the individual level or microscopic scale.

Networks and the epidemiology of directly transmitted infectious diseases are fundamentally linked[20]. As we said above each individual in a human population has a finite set of connections and the intensity of interactions, meaning there are certain connections among individuals in the population that can be a structure of human interaction. The structure underlying human interaction can be captured by a network. In a network, all individuals in the population can be represented by a set of nodes or vertices, and connections among them can be represented by edges. The number of edges incident to a node is called the degree of a node and the weight of an edge can represent the intensity of interaction between two nodes.

The epidemic model with two variants using the compartmental model has been studied analytically[2]. It shows that the cumulative number of infected individuals of two variants can be well approximated by independent logistic functions. It also shows that the model can validate empirically the multiple waves of the COVID-19 pandemic in some countries. The study of two variants of the epidemic by performing a numerical simulation of the microscopic model on the networks has been done[21]. In the study, a new variant is added in the middle of a simulation. As a result, the infectivity of variants determines the epidemic size. When a highly infectious variant is added, the variant spreads quickly. The network structure also stimulates the rapid increase of infection.

Human interactions in a population can be represented as a contact network where each individual is regarded as a node and interaction between two individuals can be considered as an edge. The intensity of interaction between two nodes can be regarded as a weight given to the edge. In this study, the weight we give to each edge in the network is a natural number that represents how many times two connected individuals interact in proximity of each other in one day. This is what we call contact which is defined as a two-way conversation between two individuals in a certain range of distance such that possible for an infection to transmit[22].

In this paper, we first generate a synthesized human interaction network with a given average degree and random frequency of contacts. Then, we develop a mathematical model that can capture the microscopic process of disease spreading in the network. At the initial time, we choose randomly N_0 individuals to get infected by Variant 1. Variant 2 will emerge from infections of Variant 1. It is described by randomly moving an infected individual due to Variant 1 to the infected individual due to Variant 2 with a certain probability. Likewise, Variant 3 will emerge from the infection of Variant 2, and Variant 2 mutates to Variant 3 with similar procedures.

With the varying average degrees and a certain range of contact frequency, we will show that the epidemic size is not only determined by varying the epidemiological parameters but also by a complex combination with the network structures. By varying the network structure, we can accommodate the mitigation strategies to suppress the spread of the disease when no vaccine and medicine are administered.

However, this study has some limitations. The model we will develop relies heavily on computation and strongly depends on the real data of social contacts. The simulation we have performed is limited to 10, 000 individuals and in turn, the network we generate with this number of individuals does not represent the real social network in a big population. In addition, the real data on social contact is not easy to collect and has also not been provided yet. Therefore, we limit our model applications to smaller-scale environments such as campuses, schools, work environments, and public places as theoretical modeling and simulations. We also agree for ethical reasons that the process of the spread of infectious disease can never be experimentally studied in our society[23]. This paper is just proposed to provide decision-makers with the predictive power of the epidemiological models with more than one variant in a smaller population from theoretic simulations.

Results

In this section, we will show some simulation results obtained from the model we developed in the "Method" Section. We perform each simulation with 100 samples for $t_f = 400$ days. We set the total number of nodes $N = 10, 000$ and the initial condition $N_0 = 5$. By varying the average degree, the range of contact frequency, and some epidemiological parameters, we can obtain the various scenarios. We can also obtain abundant results by only varying the epidemiological parameters when the network structure is fixed. However, in this simulation, we want to show how the network structure greatly affects the dynamics of disease spread in the network with fewer variations in the value of epidemiological parameters. We want to show that mitigation strategies like containment measures can be considered as variations of the average degree and contact frequency. To simplify our choice in simulations, we also set the average duration of the infectious period $\tau_v = 15$ days for each variant. Here, we chose only four types of scenarios based on our assumption when some containment measures are implemented in the population during the pandemic. The four types of scenarios are the following:

1. Epidemic model with three variants on the synthesized human interaction network without any containment measure. We assume that $\bar{k} = 7$ and $\omega = 5$.
2. Epidemic model with three variants on the synthesized human interaction network with self-quarantine but no physical distancing. We assume that $\bar{k} = 3$ and $\omega = 5$.
3. Epidemic model with three variants on the synthesized human interaction network with no self-quarantine but with physical distancing. We assume that $\bar{k} = 7$ and $\omega = 1$.

4. Epidemic model with three variants on the synthesized human interaction network with self-quarantine and physical distancing. We assume that $\bar{k} = 3$ and $\omega = 1$.
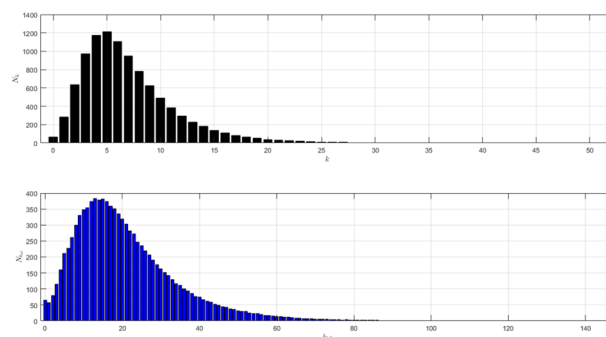
Scenario 1

The network structure of Scenario 1 with $\bar{k} = 7$ and $\omega = 5$ is shown in Fig. 1. The black histogram shows the degree distribution of the network. The blue one shows the contact distribution of the network. We can refer to the nodes with high degrees as hubs[24]. We will call the node a hub if its degree $k \geq 14$. The role of hubs in the epidemic process has been studied in[25].

From the black histogram in Fig. 1, we can infer that there are fewer nodes with no degree. We call them isolated nodes. Almost all nodes in the network are connected. In addition, there are also many hubs in the network with degree 14–27. From the blue histogram in Fig. 1, we can infer that there are a lot of nodes with very high frequency of contact. Even, there are some nodes with contacts in the range 40–80. It makes sense if the disease spreads quickly as shown in Figure 2a. The figure is the time evolution of the number of individuals of all states (S, I, R) with 95% confidence interval. We divide this scenario into four cases.

In the first case we set $\mu_{1\to2} \sim \text{Bernoulli}(0.005)$, $\mu_{2\to3} \sim \text{Bernoulli}(0.005)$ and $\beta_1 = \beta_2 = \beta_3 = 0.01$. The emergence of variants causes multiple waves as shown in Fig. 2a as the red line. We can see the emergence of each variant and their time evolution in detail in Fig. 2b. The figure is the average plot of 100 samples. The infection peak of Variant 1 is 4480 nodes at the time $t = 46$. The infection of Variant 2 is 614 nodes at the time $t = 94$. The infection of Variant 3 is 197 at the time $t = 227$. Variant 1 ends at time $t = 118$. Variant 2 starts to spread at time $t = 37$ and ends at time $t = 227$. Variant 3 starts to spread at time $t = 140$ and ends at time $t = 337$. We can see the final value of the total number of infected individuals for each variant in Fig. 2c. From the total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 8250$ and the length of epidemic time of Variant 1 is $t_{e_1} = 117$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.241$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 2305$ with $t_{e_2} = 190$ and $AI_{3_{400}}^T = 992$ with $t_{e_3} = 197$. Then, we have $\rho_2 = 0.109$ and $\rho_3 = 0.049$. We can compare the number of infected neighbors to the degree of each node in Fig. 2d. The higher the degree of a node the more neighbor will be infected.

In the second case we still set $\mu_{1\to2} \sim \text{Bernoulli}(0.005)$, $\mu_{2\to3} \sim \text{Bernoulli}(0.005)$ but $\beta_1 < \beta_2 < \beta_3$. We choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$ to show how different infection rates play roles in the epidemic process. The dynamic of the epidemic process in this case is shown in Fig. 3a. We can see that Variant 2 and Variant 3 emerge earlier as shown in Fig. 3b compared to Fig. 2b. The infection peak of Variant 1 is 4241 nodes at the time $t = 46$. The infection peak of Variant 2 is 4675 nodes at the time $t = 87$. The infection peak of Variant 3 is 2510 at the time $t = 152$. Variant 1 ends at time $t = 134$. Variant 2 starts to spread at time $t = 15$ and ends at time $t = 202$. Variant 3 starts to spread at time $t = 51$ and ends at time $t = 245$. The final value of the total number of infected individuals for each variant is shown in Fig. 2c. From the total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 8181$ and the length of epidemic time $t_{e_1} = 133$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.272$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 4675$ with $t_{e_2} = 187$ and $AI_{3_{400}}^T = 2510$ with $t_{e_3} = 194$. We obtain $\rho_2 = 0.218$ and $\rho_3 = 0.122$. The increase in the infection rate for each variant can increase the total number of infected individuals and in turn, increase the infection spread number.

Now we turn to the third case where we set $\mu_{1\to2} \sim \text{Bernoulli}(0.001)$, $\mu_{2\to3} \sim \text{Bernoulli}(0.001)$ and $\beta_1 = \beta_2 = \beta_3 = 0.01$. Here we decrease the probability of the emergence of new variants. The dynamic of the epidemic process in this case is shown in Fig. 4a. The infected individuals are dominated by Variant 1 as shown in Fig. 4b. The infection peak of Variant 1 is 4482 nodes at the time $t = 46$. The infection peak of Variant 2 is 158 nodes at the time $t = 104$. The infection peak of Variant 3 is 99 at the time $t = 265$. Variant 1 ends at time $t = 119$. Variant 2 starts to spread at time $t = 65$ and ends at time $t = 240$. Variant 3 starts to spread at time $t = 232$ and ends at time $t = 306$. The final value of the total number of infected individuals for Variant 2 and Variant 3 are much smaller than the total number of infected individuals for Variant 1 as shown in Fig. 4c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 8256$ and the epidemic duration $t_{e_1} = 118$ days give the infection spread number of Variant 1 $\rho_1 = 0.243$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 496$ with $t_{e_2} = 175$ and

Figure 1. The network structure of Scenario 1. The black histogram shows the degree distribution of the network with $\bar{k} = 7$ and $\omega = 5$. It shows that the network has a lot of hubs and fewer isolated nodes. The blue histogram shows the contact distribution. More than half of the population have a contact frequency of $20 - 90$.
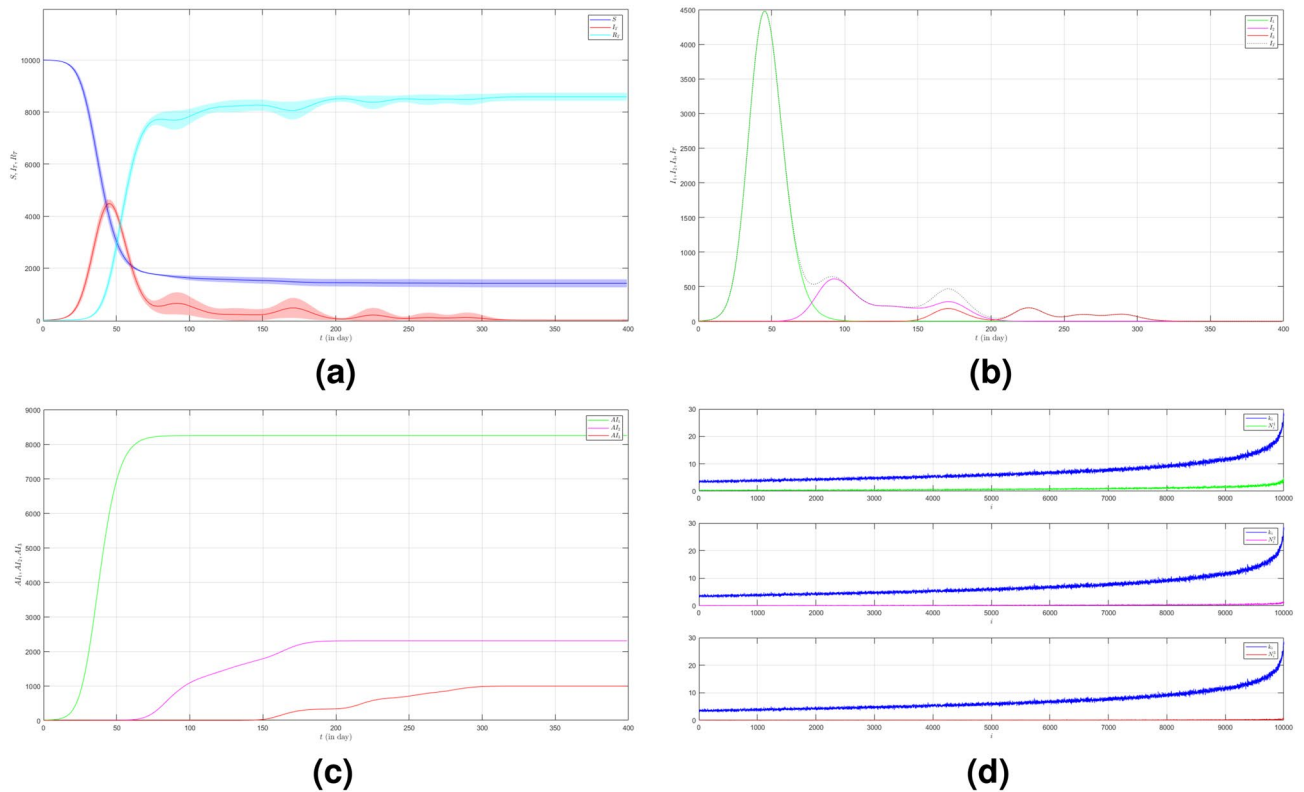
Figure 2. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the first case of Scenario 1. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{46}} = 4480$, $I_{2_{94}} = 614$, and $I_{3_{227}} = 197$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^{T} = 8250$, $AI_{2_{400}}^{T} = 2305$, and $AI_{3_{400}}^{T} = 992$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread number are $\rho_1 = 0.241$, $\rho_2 = 0.109$, and $\rho_3 = 0.049$ respectively.

$AI_{3_{400}}^{T} = 166$ with $t_{e_3} = 74$. We have $\rho_2 = 0.022$ and $\rho_3 = 0.003$. The smaller probability of the emergence of new variants can reduce the total number of infected individuals of new variants.

To see how different infection rates from each variant play roles in a condition when the probability of the emergence of new variants is smaller, we set the fourth case where $\mu_{1\rightarrow 2} \sim$ Bernoulli(0.001), $\mu_{2\rightarrow 3} \sim$ Bernoulli(0.001) and $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$. The dynamic of the epidemic process in this case is shown in Fig. 5a. The infected individuals are still dominated by Variant 1 as shown in Fig. 5b. The infection peak of Variant 1 is 4448 nodes at the time $t = 45$. The infection peak of Variant 2 is 376 nodes at the time $t = 70$. The infection peak of Variant 3 is 181 at the time $t = 192$. Variant 2 starts to spread at time $t = 30$ and ends at time $t = 147$. Variant 1 ends at time $t = 117$. Variant 3 starts to spread at time $t = 174$ and ends at time $t = 217$. The smaller probability of the emergence of new variants causes the final values of the total number of infected individuals for Variant 2 and Variant 3 to be smaller than the total number of infected individuals for Variant 1 as shown in Fig. 5c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^{T} = 8249$ and the epidemic duration $t_{e_1} = 116$ days give the infection spread number of Variant 1 $\rho_1 = 0.239$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^{T} = 939$ with $t_{e_2} = 117$ and $AI_{3_{400}}^{T} = 195$ with $t_{e_3} = 43$. We have $\rho_2 = 0.027$ and $\rho_3 = 0.002$. Although a smaller probability of the emergence of new variants can reduce the total number of infected individuals of new variants, the higher infection rates still give a significant increase in the total number of infected individuals for each variant.

Scenario 2

The network structure of Scenario 2 where $\bar{k} = 3$ and $\omega = 5$ is shown in Fig. 6. In this scenario, we contain the spread of infection in the network by implementing self-quarantine without physical distancing. We reduce individual connectivity without reducing their contact frequency.

The black histogram in Fig. 1 shows that there is no hub in this network. From the blue histogram in Fig. 1, there are fewer nodes with contacts in the range $30 - 40$ which is the high contacts in the network. The dynamic of the epidemic process is shown in Fig. 7a. The figure is the time evolution of the number of individuals of all states (S, I, R) with 95% confidence interval. Here, we still divide this scenario into four cases.

In the first case of Scenario 2, we set $\mu_{1\rightarrow 2} \sim$ Bernoulli(0.005), $\mu_{2\rightarrow 3} \sim$ Bernoulli(0.005) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. The number of infected individuals from all variants indicated by the red line is shown
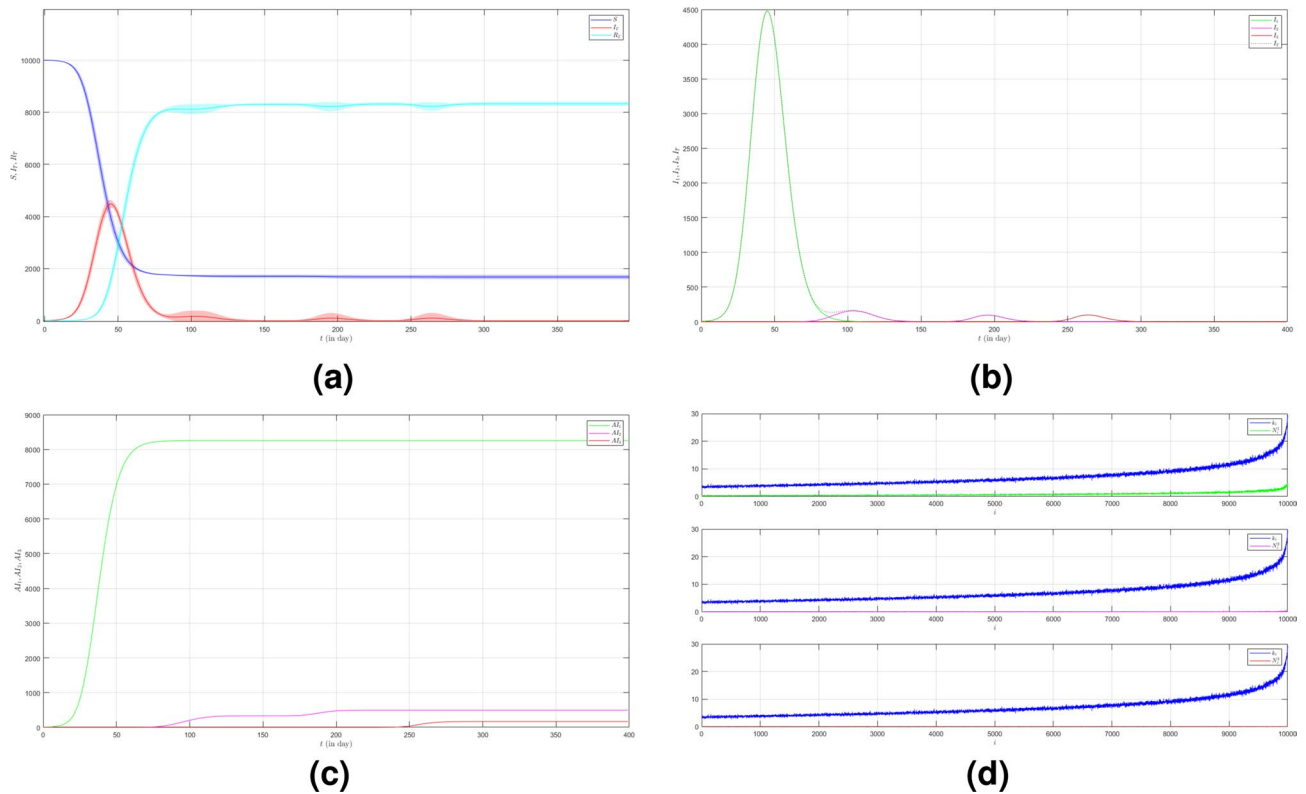
(a)

(b)

(c)

(d)
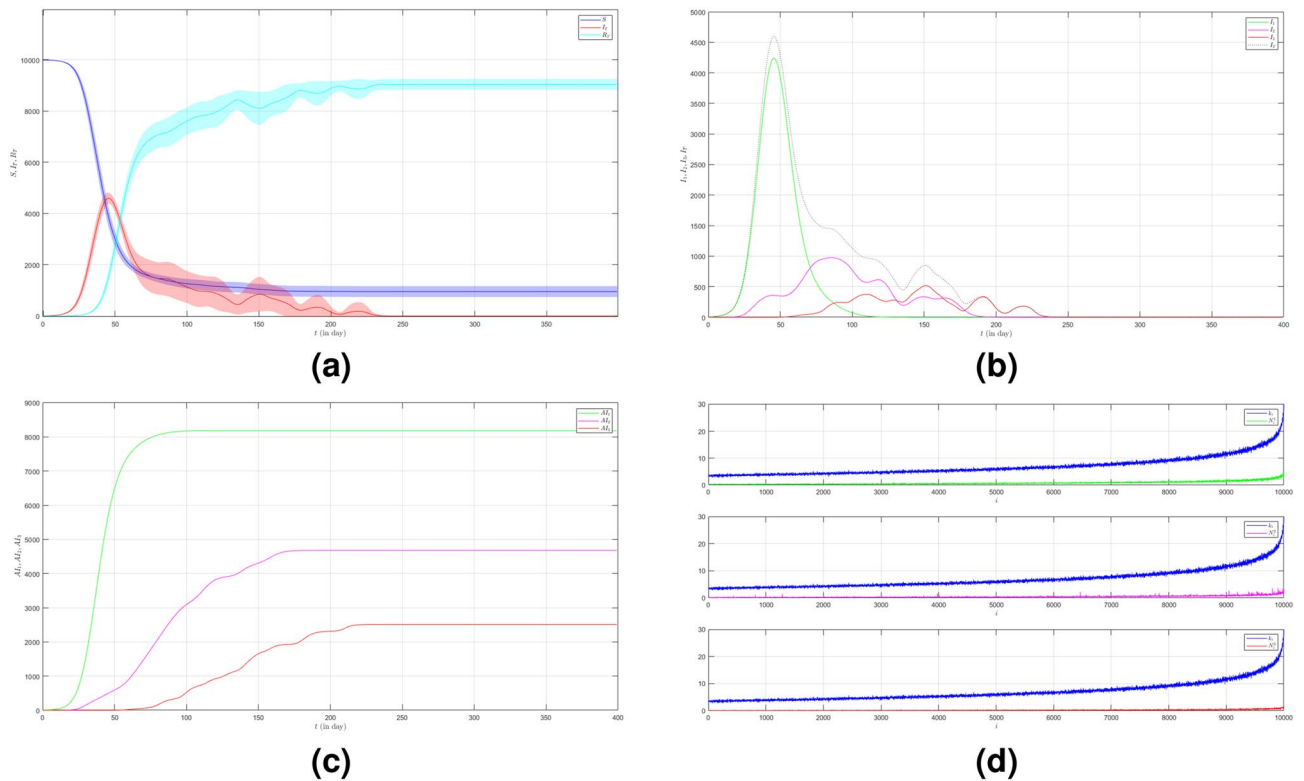
Figure 3. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the second case of Scenario 1. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{46}} = 4241$, $I_{2_{87}} = 977$, and $I_{3_{152}} = 519$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 8181$, $AI_{2_{400}}^T = 4675$, and $AI_{3_{400}}^T = 2510$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread number are $\rho_1 = 0.272$, $\rho_2 = 0.218$, and $\rho_3 = 0.122$ respectively.

in Fig. 7a. The emergence of new variants is shown in Fig. 7b. The figure is the average plot of 100 samples. The infection peak of Variant 1 is 381 nodes at the time $t = 115$. The infection of Variant 2 is 74 nodes at the time $t = 212$. The infection of Variant 3 is 30 at the time $t = 384$. Variant 1 ends at time $t = 328$. Variant 2 starts to spread at time $t = 64$ and it does not end even until the last time. Variant 3 starts to spread at time $t = 83$ and it does not end either. We can see the final value of the total number of infected individuals for each variant in Fig. 7c. From the total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 2426$ and the length of epidemic time of Variant 1 is $t_{e_1} = 327$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.198$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 688$ with $t_{e_2} = 336$ and $AI_{3_{400}}^T = 319$ with $t_{e_3} = 317$. Then, we have $\rho_2 = 0.058$ and $\rho_3 = 0.025$. In this case, reducing the connectivity of each individual by reducing the average degree prolongs the length of the epidemic time of each variant. We can compare the number of infected neighbors to the degree of each node in Fig. 7d. The higher the degree of a node the more neighbor will be infected.

In the second case of Scenario 2, we still set $\mu_{1 \to 2} \sim$ Bernoulli(0.005), $\mu_{2 \to 3} \sim$ Bernoulli(0.005) but $\beta_1 < \beta_2 < \beta_3$. We choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$ to show how different infection rates play roles in the epidemic process when the average degree is reduced. The dynamic of the epidemic process in this case is shown in Fig. 8a. There is an increase in the number of infections due to Variant 2 and Variant 3 as shown in Fig. 8b compared to Fig. 7b. The emergence of new variants is earlier than the emergence of new variants in the first case. The infection peak of Variant 1 is 337 nodes at the time $t = 105$. The infection peak of Variant 2 is 451 nodes at the time $t = 101$. The infection peak of Variant 3 is 251 nodes at the time $t = 145$. Variant 1 ends at time $t = 332$. Variant 2 starts to spread at time $t = 29$ and ends at time $t = 306$. Variant 3 starts to spread at time $t = 42$ and does not end until the last time. The final value of the total number of infected individuals for each variant is shown in Fig. 8c. From the total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 2195$ and the length of epidemic time $t_{e_1} = 331$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.182$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 2610$ with $t_{e_2} = 277$ and $AI_{3_{400}}^T = 2078$ with $t_{e_3} = 358$. We obtain $\rho_2 = 0.181$ and $\rho_3 = 0.186$.

Now we set the third case of Scenario 2 where $\mu_{1 \to 2} \sim$ Bernoulli(0.001), $\mu_{2 \to 3} \sim$ Bernoulli(0.001) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. The dynamic of the epidemic process in this case is shown in Fig. 9a. The infection of Variant 2 and Variant 3 is smaller than the two previous cases as shown in Fig. 9b. The infection peak of Variant 1 is 454 nodes at the time $t = 106$. The infection peak of Variant 2 is 29 nodes at the time $t = 334$. There is no spread of Variant 3. Variant 1 ends at time $t = 270$. Variant 2 starts to spread at time $t = 115$ and does not end until the last time. The final value of the total number of infected individuals for Variant 2 and Variant 3 are

Figure 4. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the third case of Scenario 1. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{46}} = 4482$, $I_{2_{104}} = 158$, and $I_{3_{265}} = 99$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 8256$, $AI_{2_{400}}^T = 496$, and $AI_{3_{400}}^T = 166$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread number are $\rho_1 = 0.243$, $\rho_2 = 0.022$, and $\rho_3 = 0.003$ respectively.

much smaller than the total number of infected individuals for Variant 1 as shown in Fig. 9c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 2498$ and the epidemic duration $t_{e_1} = 269$ days give the infection spread number of Variant 1 $\rho_1 = 0.168$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 277$ with $t_{e_2} = 285$ and $AI_{3_{400}}^T = 1$. We have $\rho_2 = 0.019$ and $\rho_3 = 0$.

To see how different infection rates from each variant play roles in a condition when the probability of the emergence of new variants is smaller, we set the fourth case of Scenario 2 where $\mu_{1\to 2} \sim$ Bernoulli(0.001), $\mu_{2\to 3} \sim$ Bernoulli(0.001) and $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$. The dynamic of the epidemic process in this case is shown in Fig. 10a. The time evolution of infected individuals for each variant is shown in Fig. 10b. The infection peak of Variant 1 is 429 nodes at the time $t = 102$. The infection peak of Variant 2 is 241 nodes at the time $t = 140$. The infection peak of Variant 3 is 56 at the time $t = 169$. Variant 1 ends at time $t = 303$. Variant 2 starts to spread at time $t = 59$ and ends at time $t = 303$. Variant 3 starts to spread at time $t = 125$ and ends at time $t = 226$. The total number of infected individuals for each variant is shown in Fig. 10c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 2547$ and the epidemic duration $t_{e_1} = 302$ days give the infection spread number of Variant 1 $\rho_1 = 0.192$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 1000$ with $t_{e_2} = 244$ and $AI_{3_{400}}^T = 143$ with $t_{e_3} = 101$. We have $\rho_2 = 0.061$ and $\rho_3 = 0.003$.
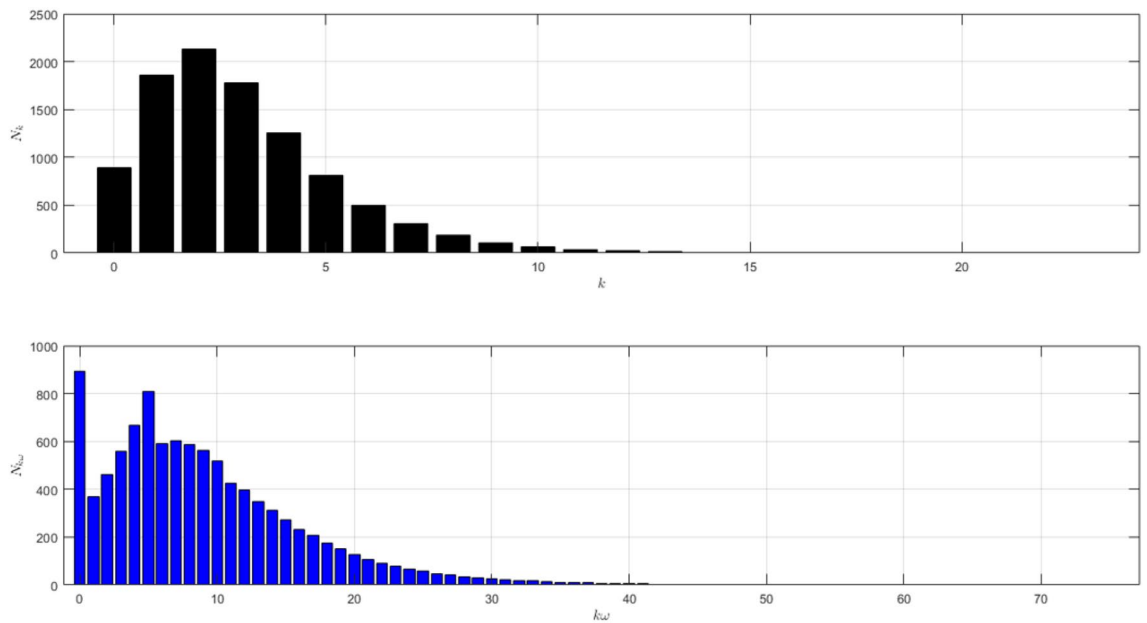
Scenario 3

We assume that the network structure of Scenario 3 where $\bar{k} = 7$ and $\omega = 1$ is the epidemic process in the network with measured physical distancing but no self-quarantine as shown in Fig. 11. In this scenario, we contain the spread of infection in the network by implementing physical distancing without self-quarantine. We reduce the individual frequency of contact without reducing the connectivity. Here, we still divide this scenario into four cases.

The black histogram in Fig. 11 is the same as the blue histogram. There are quite a lot of individuals who are hubs and have frequency contact of $14 - 28$. The dynamic of the epidemic process is shown in Fig. 12a. The figure is the time evolution of the number of individuals of all states (S, I, R) with 95% confidence interval.

In the first case of Scenario 3, we set $\mu_{1\to 2} \sim$ Bernoulli(0.005), $\mu_{2\to 3} \sim$ Bernoulli(0.005) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. The number of infected individuals from all variants indicated by the red line is shown in Fig. 12a. Although the infected individuals are still dominated by the infection of Variant 1, the total number
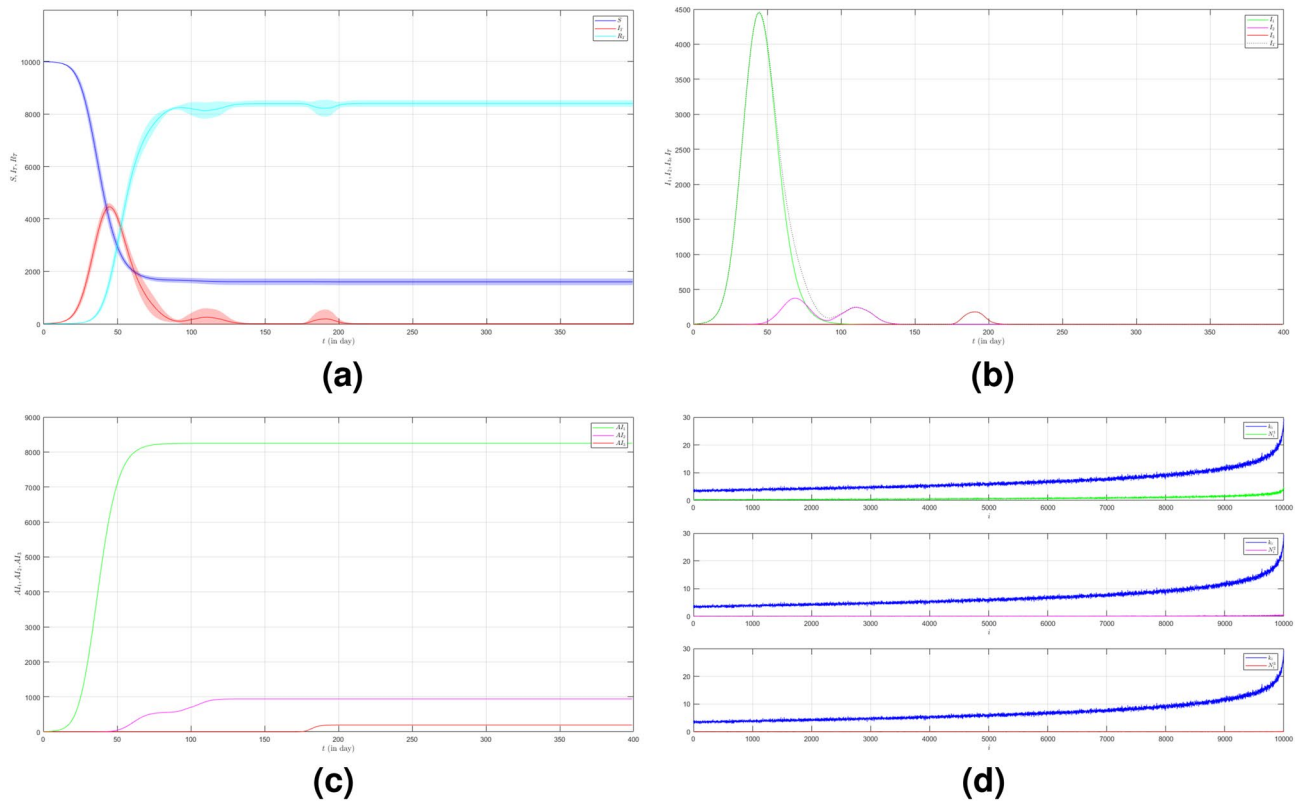
Figure 5. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the fourth case of Scenario 1. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{145} = 4448$, $I_{270} = 376$, and $I_{3192} = 181$ respectively. (**c**) The total number of infections for all variants are $AI_{1400}^T = 8249$, $AI_{2400}^T = 939$, and $AI_{3400}^T = 195$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.239$, $\rho_2 = 0.027$, and $\rho_3 = 0.002$ respectively.

Figure 6. The network structure of Scenario 2. The black histogram shows the network's degree distribution with $\bar{k} = 3$ and $\omega = 5$. The highest degree is 13. The network is dominated by the nodes with degree 2. The blue histogram shows the distribution of the total number of contacts. More than half of the population have a contact frequency of $10 - 40$.
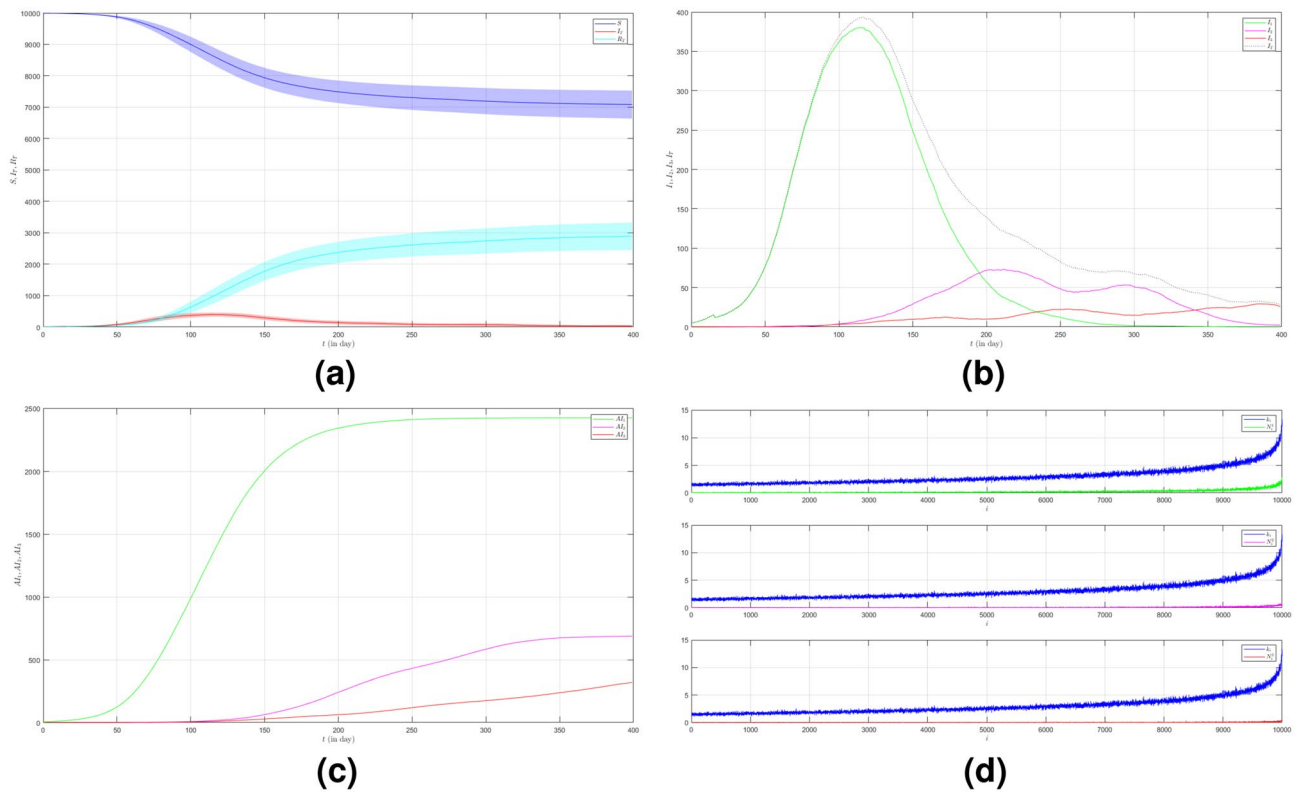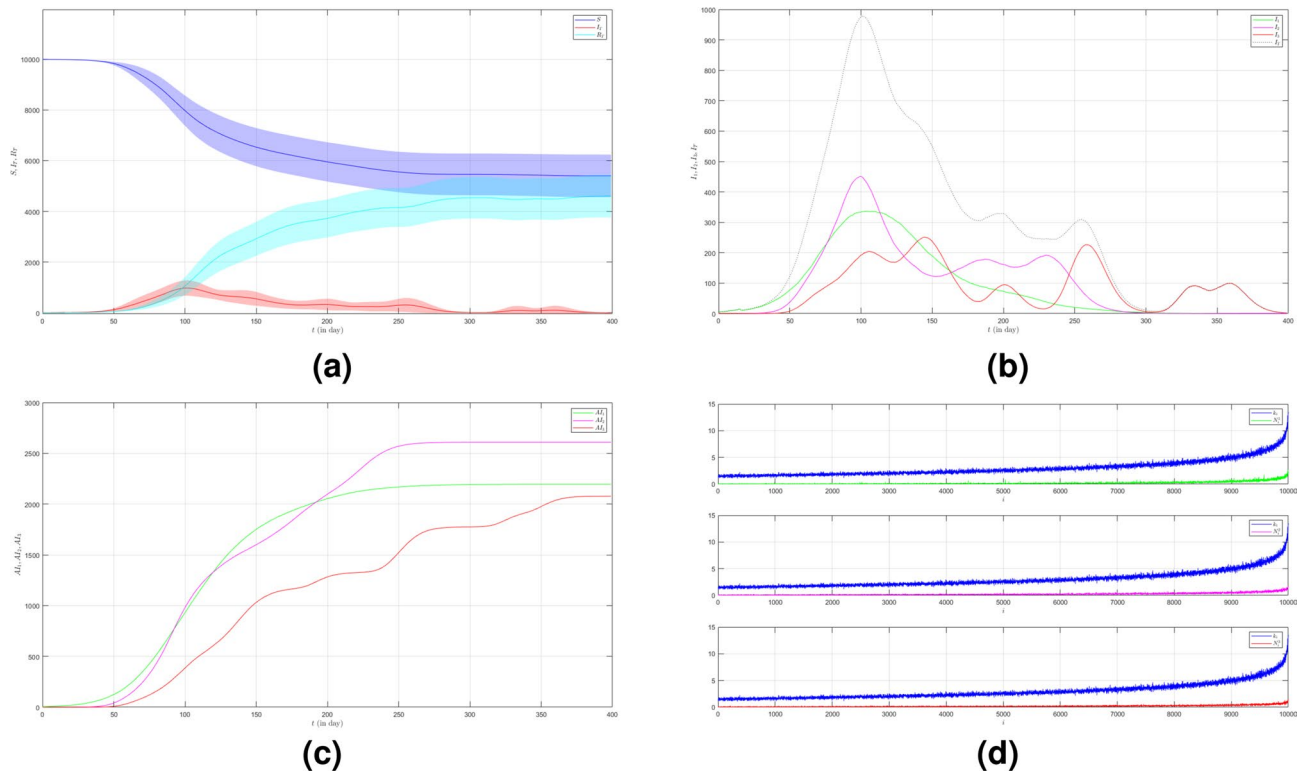
Figure 7. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the first case of Scenario 2. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{115}} = 381$, $I_{2_{212}} = 74$, and $I_{3_{384}} = 30$ respectively. (**c**) The total number of infections for all variants are $AI^T_{1_{400}} = 2426$, $AI^T_{2_{400}} = 688$, and $AI^T_{3_{400}} = 319$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.198$, $\rho_2 = 0.058$, and $\rho_3 = 0.025$ respectively.

of infected individuals is much smaller than in all previous cases we obtained above. The emergence of new variants is shown in Fig. 12b. The figure is the average plot of 100 samples. We can see that the time evolution of infection of Variant 2 has no end. The infection peak of Variant 1 is 271 nodes at the time $t = 122$. The infection of Variant 2 is 44 nodes at the time $t = 177$. The infection of Variant 3 is 18 at the time $t = 244$. Variant 1 ends at time $t = 364$. Variant 2 starts to spread at time $t = 38$ and it does not end even until the last time. Variant 3 starts to spread at time $t = 148$ and it ends at time $t = 343$. We can see the final value of the total number of infected individuals for each variant in Fig. 12c. From the total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 1925$ and the length of epidemic time of Variant 1 is $t_{e_1} = 363$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.175$. For Variant 2 and Variant 3, we obtain $AI^T_{2_{400}} = 592$ with $t_{e_2} = 362$ and $AI^T_{3_{400}} = 90$ with $t_{e_3} = 195$. Then, we have $\rho_2 = 0.054$ and $\rho_3 = 0.004$. In this case, reducing the frequency of contact also prolongs the length of the epidemic time of each variant. We can compare the number of infected neighbors to the degree of each node in Fig. 12d.

In the second case of Scenario 3, we still set $\mu_{1\to2} \sim$ Bernoulli(0.005), $\mu_{2\to3} \sim$ Bernoulli(0.005) but $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$ to show how different infection rates play roles in the epidemic process when the frequency of contact is reduced. The dynamic of the epidemic process in this case is shown in Fig. 13a. There is an increase in the number of infections due to Variant 2 and Variant 3 as shown in Fig. 13b compared to Fig. 12b. The emergence of new variants is earlier than the emergence of new variants in the first case. The infection peak of Variant 1 is 196 nodes at the time $t = 139$. The infection peak of Variant 2 is 540 nodes at the time $t = 100$. The infection peak of Variant 3 is 297 at the time $t = 211$. Variant 1 spreads until the last time. Variant 2 starts to spread at time $t = 19$ and has no end. Variant 3 starts to spread at time $t = 59$ and has no end either. The final value of the total number of infected individuals for each variant is shown in Fig. 13c. From the total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 1617$ and the length of epidemic time $t_{e_1} = 399$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.161$. For Variant 2 and Variant 3, we obtain $AI^T_{2_{400}} = 3472$ with $t_{e_2} = 381$ and $AI^T_{3_{400}} = 1794$ with $t_{e_3} = 341$. We obtain $\rho_2 = 0.331$ and $\rho_3 = 0.153$. In this case, if the infection rate of the new variant is larger than the previous one, we can see that not only does the total number of infected individuals increase but also the length of the epidemic time is longer.

Now we set the third case of Scenario 3 where $\mu_{1\to2} \sim$ Bernoulli(0.001), $\mu_{2\to3} \sim$ Bernoulli(0.001) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. In this case, we want to see what happens if we reduce the probability of the emergence of new variants. The dynamic of the epidemic process in this case is shown in Fig. 14a. We can see that the time
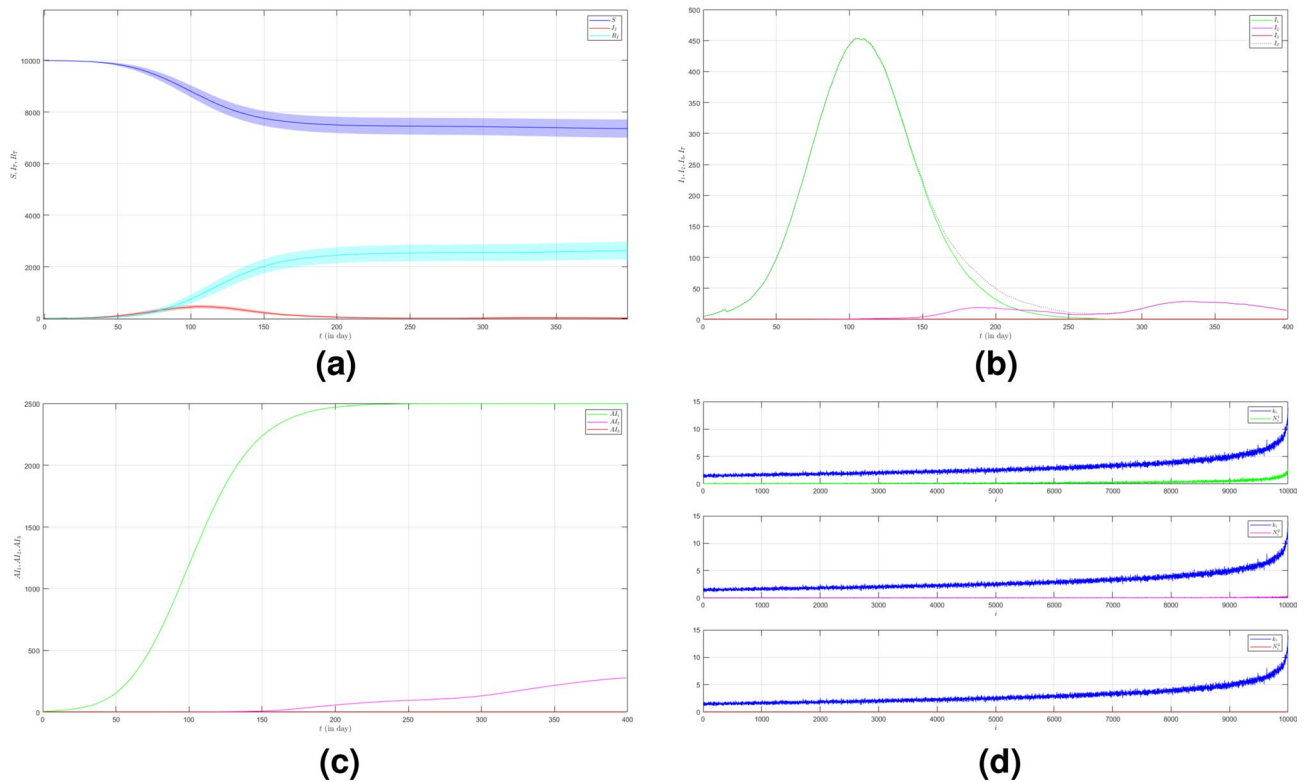
Figure 8. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the second case of Scenario 2. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{105}} = 337$, $I_{2_{101}} = 451$, and $I_{3_{145}} = 251$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 2195$, $AI_{2_{400}}^T = 2610$, and $AI_{3_{400}}^T = 2078$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.182$, $\rho_2 = 0.181$, and $\rho_3 = 0.186$ respectively.

evolution of total infections indicated by the red line increases very slowly and levels off for quite a long time, then decreases until the end. The infection number of Variant 2 and Variant 3 is much smaller as shown in Fig. 14b. The infection peak of Variant 1 is 226 nodes at the time $t = 126$. The infection peak of Variant 2 is 1 nodes at the time $t = 127$. It means there is only one infection of Variant 2. Thus, there is no spread of Variant 2. The infection peak of Variant 3 is 10 nodes at the time $t = 314$. Variant 1 ends at time $t = 304$. Variant 3 starts to spread at time $t = 259$ and does not end until the last time. The final value of the total number of infected individuals for Variant 2 and Variant 3 are much smaller than the total number of infected individuals for Variant 1 as shown in Fig. 14c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 1599$ and the epidemic duration $t_{e_1} = 303$ days give the infection spread number of Variant 1 $\rho_1 = 0.0121$. Since there is no spread of Variant 2, we write $\rho_2 = 0$. For Variant 3, we obtain $AI_{2_{400}}^T = 51$ with $t_{e_2} = 141$ and then $\rho_3 = 0.002$.

To see how different infection rates from each variant play roles in a condition when the probability of the emergence of new variants is smaller and the frequency of contact is reduced, we set the fourth case of Scenario 3 where $\mu_{1 \to 2} \sim \text{Bernoulli}(0.001)$, $\mu_{2 \to 3} \sim \text{Bernoulli}(0.001)$ and $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$. The dynamic of the epidemic process in this case is shown in Fig. 15a. The time evolution of infected individuals for each variant is shown in Fig. 15b. The infection peak of Variant 1 is 258 nodes at the time $t = 118$. The infection peak of Variant 2 is 187 nodes at the time $t = 256$. The infection peak of Variant 3 is 100 at the time $t = 226$. Variant 1 ends at time $t = 323$. Variant 2 starts to spread at time $t = 46$ and ends at time $t = 375$. Variant 3 starts to spread at time $t = 73$ and ends at time $t = 173$. The total number of infected individuals for each variant is shown in Fig. 15c. The total number of infected individuals of Variant 1 $AI_{1_{400}}^T = 1751$ and the epidemic duration $t_{e_1} = 322$ days give the infection spread number of Variant 1 $\rho_1 = 0.141$. For Variant 2 and Variant 3, we obtain $AI_{2_{400}}^T = 1357$ with $t_{e_2} = 329$ and $AI_{3_{400}}^T = 328$ with $t_{e_3} = 100$. We have $\rho_2 = 0.111$ and $\rho_3 = 0.008$.

Scenario 4
We assume that the network structure of Scenario 4 where $\bar{k} = 3$ and $\omega = 1$ is the epidemic process in the network with measured physical distancing and self-quarantine as shown in Fig. 16. In this scenario, we contain the spread of infection in the network by implementing both self-quarantine and physical distancing. We reduce both the connectivity and the individual frequency of contact.
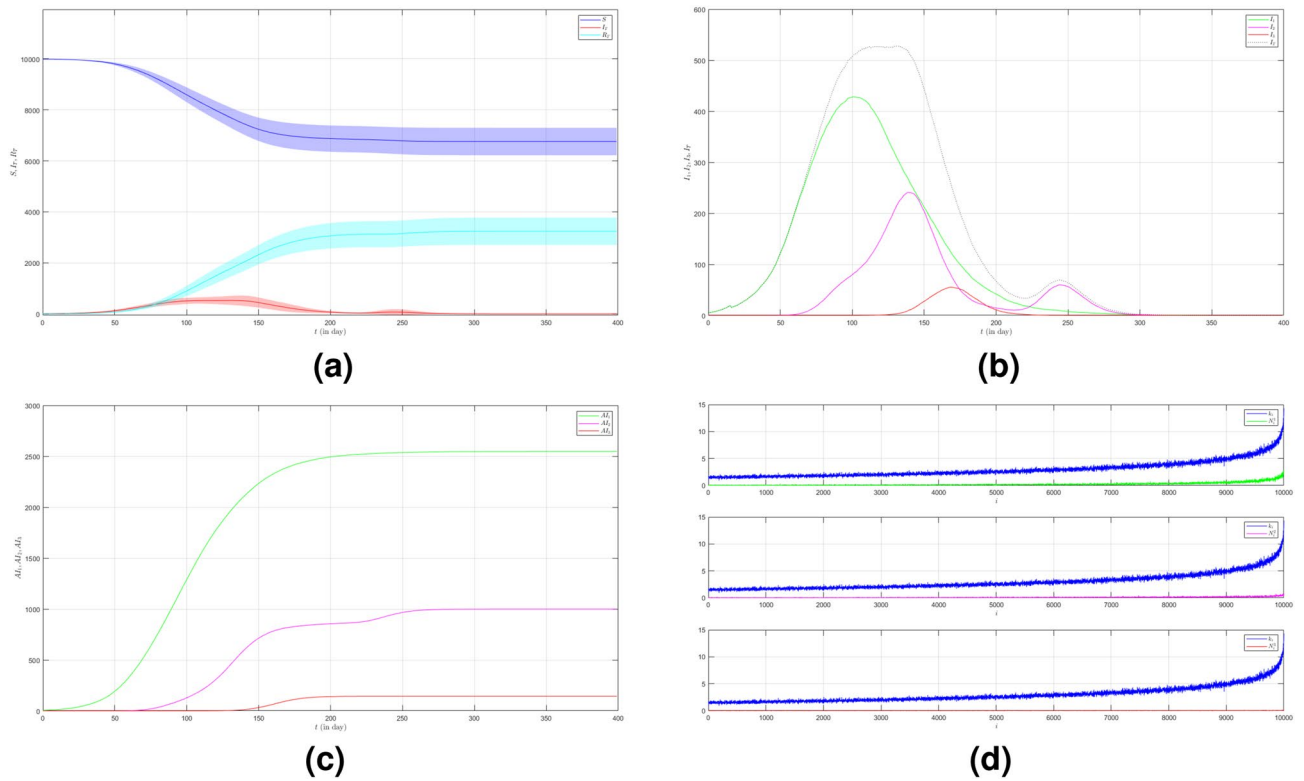
Figure 9. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the third case of Scenario 2. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{106}} = 454$, $I_{2_{334}} = 29$, and $I_{3_{139}} = 1$ respectively. (**c**) The total number of infections for all variants are $AI^T_{1_{400}} = 2498$, $AI^T_{2_{400}} = 277$, and $AI^T_{3_{400}} = 1$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.168$, $\rho_2 = 0.019$, and $\rho_3 = 0$ respectively.

The black histogram in Fig. 16 is the same as the blue histogram. There is no hub in this network. The dynamic of the epidemic process is shown in Fig. 17a. The figure is the time evolution of the number of individuals of all states (S, I, R) with 95% confidence interval. Here, we still divide this scenario into four cases.

For the first case of Scenario 4, we set $\mu_{1\rightarrow 2} \sim$ Bernoulli(0.005), $\mu_{2\rightarrow 3} \sim$ Bernoulli(0.005) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. The number of infected individuals from all variants indicated by the red line is shown in Fig. 17a. The emergence of new variants is shown in Fig. 17b. The figure is the average plot of 100 samples. The infection peak of Variant 1 is 8 nodes at the time $t = 16$. There is no spread of Variant 2 and Variant 3. The spread of Variant 1 ends at time $t = 39$. We can see the final value of the total number of infected individuals for each variant in Fig. 17c. From the total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 11$ and the length of epidemic time of Variant 1 is $t_{e_1} = 38$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.0001$. We can compare the number of infected neighbors to the degree of each node in Fig. 17d.

In the second case of Scenario 4, we still set $\mu_{1\rightarrow 2} \sim$ Bernoulli(0.005), $\mu_{2\rightarrow 3} \sim$ Bernoulli(0.005) but $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$ to show how different infection rates play roles in the epidemic process when the two containment measures are implemented. The dynamic of the epidemic process in this case is shown in Fig. 18a. There is an increase in the number of infections due to Variant 3 as shown in Fig. 18b. The infection peak of Variant 1 is 9 nodes at time $t = 16$. There is no spread of Variant 2. The infection peak of Variant 3 is 34 at time $t = 147$. Variant 1 ends at time $t = 47$. Variant 3 starts to spread at time $t = 76$ and ends at time $t = 225$. The final value of the total number of infected individuals for each variant is shown in Fig. 18c. From the total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 12$ and the length of epidemic time $t_{e_1} = 46$ days, we can find the infection spread number of Variant 1 $\rho_1 = 0.0001$. Since there is no spread of Variant 2, we write $\rho_2 = 0$. For Variant 3, we obtain $AI^T_{3_{400}} = 134$ with $t_{e_3} = 149$. We obtain $\rho_3 = 0.005$. In this case, reducing the connectivity and the frequency of contact is not strong enough to contain the spread of variants if the infection rate of the new variant is more infectious.

Now we set the third case of Scenario 3 where $\mu_{1\rightarrow 2} \sim$ Bernoulli(0.001), $\mu_{2\rightarrow 3} \sim$ Bernoulli(0.001) and $\beta_1 = \beta_2 = \beta_3 = 0.01$. In this case, we want to see what happens if we reduce the probability of the emergence of new variants in the network with reduced connectivity and frequency of contact. The dynamic of the epidemic process in this case is shown in Fig. 19a. There is no infection by Variant 2 and Variant 3 as shown in Fig. 14b. The infection peak of Variant 1 is 9 nodes at the time $t = 16$. There is no spread of Variant 2 and Variant 3. Thus $\rho_2 = \rho_3 = 0$. Variant 1 ends at time $t = 59$. The total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 14$ and the epidemic duration $t_{e_1} = 58$. We have $\rho_1 = 0.0002$.

(a)   **(b)**

(c)   **(d)**

Figure 10. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the fourth case of Scenario 2. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{102}} = 429$, $I_{2_{140}} = 241$, and $I_{3_{169}} = 56$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^{T} = 2547$, $AI_{2_{400}}^{T} = 1000$, and $AI_{3_{400}}^{T} = 143$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.192$, $\rho_2 = 0.061$, and $\rho_3 = 0.003$ respectively.
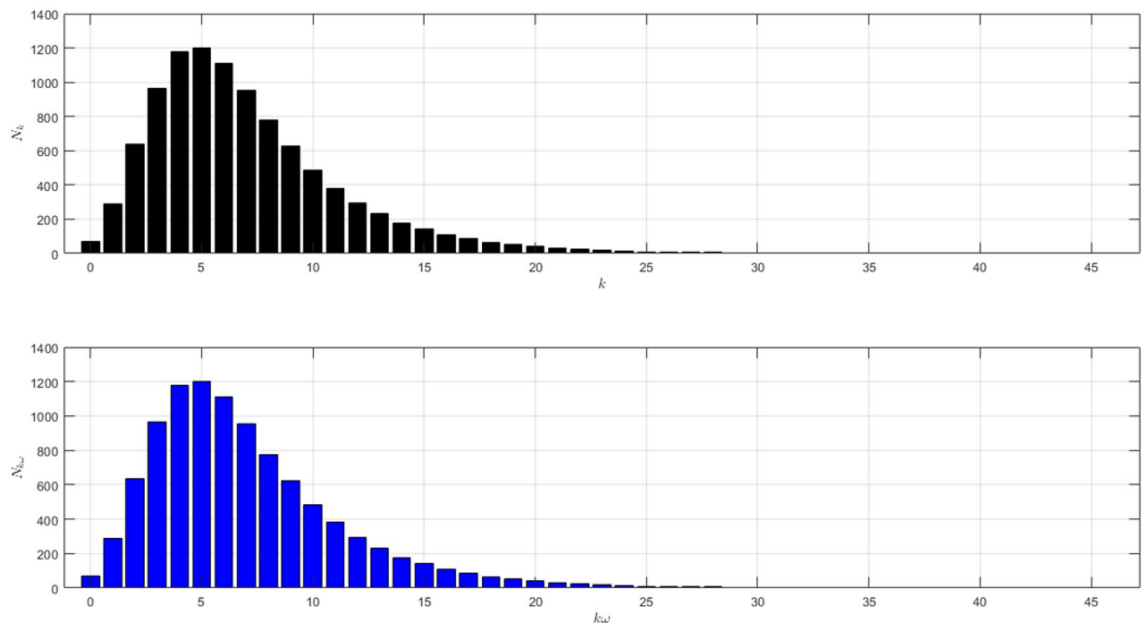
Figure 11. The network structure of Scenario 3. The black histogram shows the network's degree distribution with $\bar{k} = 7$ and $\omega = 1$. It is the same distribution as the blue histogram because the contact frequency is 1 for each link in the network. There are many hubs with a contact frequency of $14 - 28$.
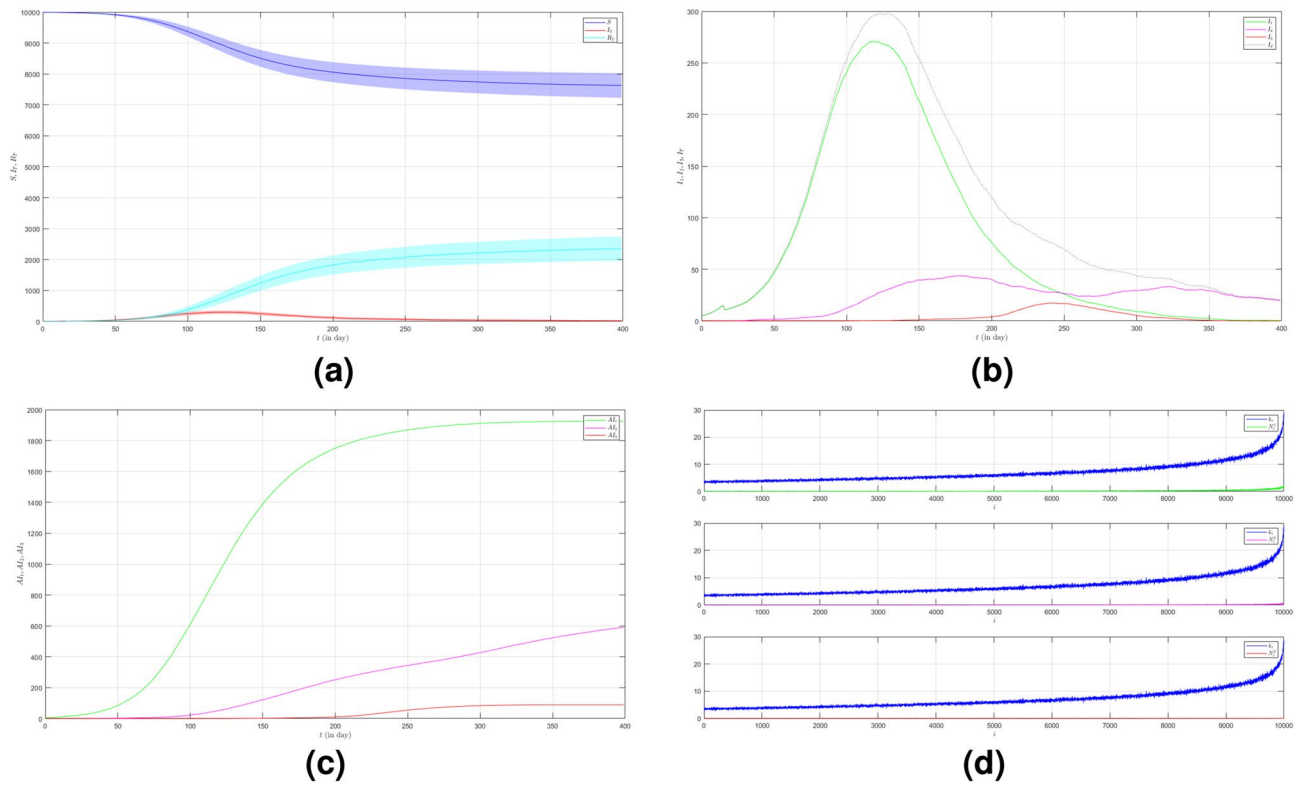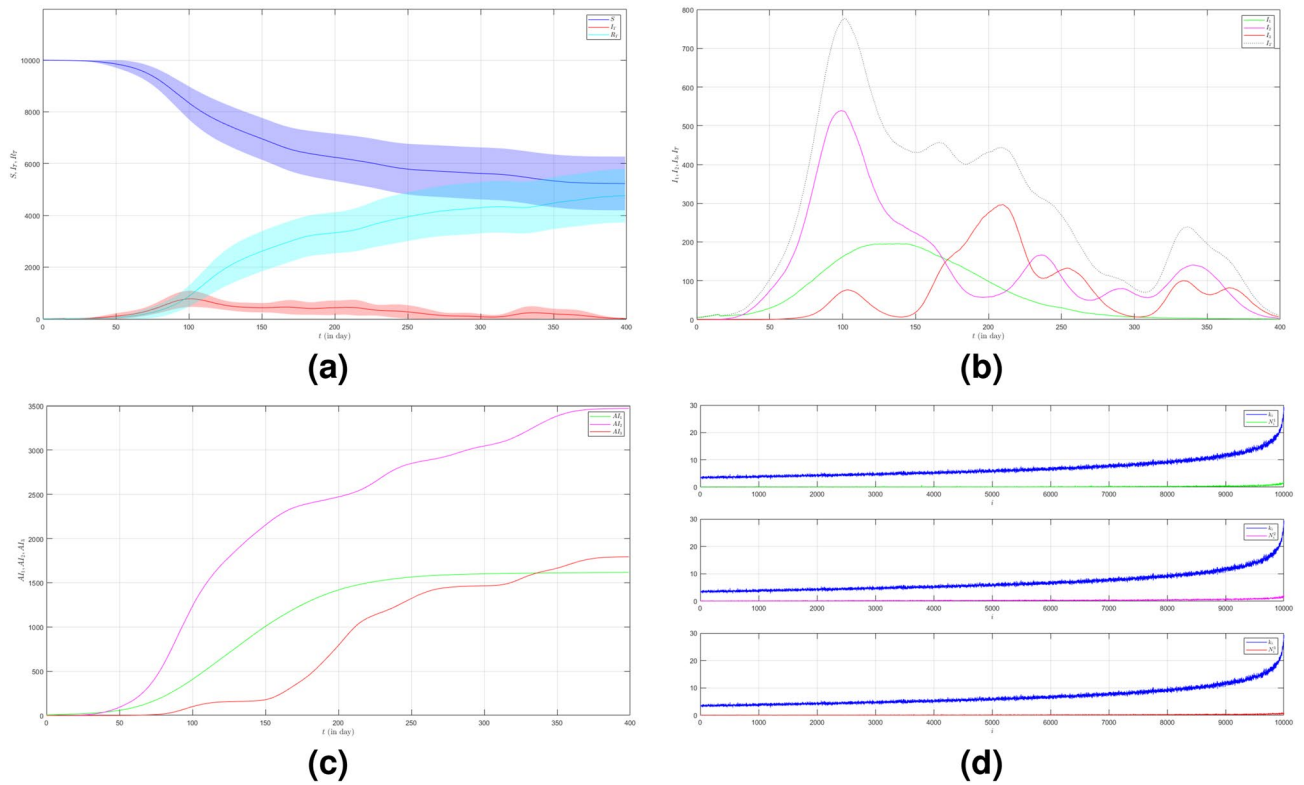
Figure 12. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the first case of Scenario 3. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{122}} = 271$, $I_{2_{177}} = 44$, and $I_{3_{244}} = 18$ respectively. (**c**) The total number of infections for all variants are $AI^T_{1_{400}} = 1925$, $AI^T_{2_{400}} = 592$, and $AI^T_{3_{400}} = 90$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.175$, $\rho_2 = 0.054$, and $\rho_3 = 0.004$ respectively.

To see how different infection rates from each variant play roles in a condition when the probability of the emergence of new variants is smaller and all contained measures are implemented, we set the fourth case of Scenario 3 where $\mu_{1\to 2} \sim$ Bernoulli(0.001), $\mu_{2\to 3} \sim$ Bernoulli(0.001) and $\beta_1 < \beta_2 < \beta_3$. We still choose $\beta_1 = 0.01$, $\beta_2 = 0.02$ and $\beta_3 = 0.03$. The dynamic of the epidemic process in this case is shown in Fig. 20a. The time evolution of infected individuals for each variant is shown in Fig. 20b. The infection peak of Variant 1 is 9 nodes at time $t = 16$. The infection peak of Variant 2 is 5 nodes at time $t = 158$. There is no infection of Variant 3. Variant 1 ends at time $t = 41$. Variant 2 starts to spread at time $t = 114$ and ends at time $t = 224$. The total number of infected individuals for each variant is shown in Fig. 20c. The total number of infected individuals of Variant 1 $AI^T_{1_{400}} = 12$ and the the length of epidemic time of Variant 1 $t_{e_1} = 40$ days give the infection spread number of Variant 1 $\rho_1 = 0.0001$. For Variant 2 we obtain $AI^T_{2_{400}} = 26$ with $t_{e_2} = 110$. We have $\rho_2 = 0.0007$.

Discussion

We developed a computational model of disease spreading with three variants that can capture the microscopic process in a synthesized human interaction network. In generating the network we need some social features in human interaction such as connectivity and contact frequency or interaction intensity. Both features determine the network structures such that every individual has an unequal chance of coming into contact with any other individual in the population. Of course, it is more realistic to describe a human population although not a real representation of the social network of the human population. Therefore, discussing the spread of contact-based diseases in a human population is more relevant in smaller-scale environments. However, we need to develop some critical aspects. Because it is a microscopic process and we brought the model to an individual scale, as a consequence, unlikely the classical SIR model, our model does not use merely infection rate β. we also do not use a recovery rate as usual we use it in the common SIR model. Incorporating more epidemiological parameters such as the average duration of the infectious period, the probability of the emergence of new variants, the length of the infection period, the length of the incubation period, and many more, has also consequences. In our model, we perform some tricks to the parameters such that they can be incorporated into the model and more applicable for both mathematics and computation purposes. We can also generalize our model to any number of variants and incorporate more epidemiological parameters such that the model is more complicated and realistic. Hopefully, there will be a study to analyze our model rigorously.

Our model has rich computational results by varying several parameters such as the number of individuals in the network, the average degree, the range of contact frequency, the values of epidemiological parameters, the

Figure 13. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the second case of Scenario 3. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{139}} = 196$, $I_{2_{100}} = 540$, and $I_{3_{211}} = 297$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^{T} = 1617$, $AI_{2_{400}}^{T} = 3472$, and $AI_{3_{400}}^{T} = 1794$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread number of all variants are $\rho_1 = 0.161$, $\rho_2 = 0.331$, and $\rho_3 = 0.153$ respectively.

initial conditions, the probability of the emergence of the new variants, and the infection rate of each variant. In the simulation, we chose to vary the average degree and contact frequency to elucidate the role of social network structures in disease spreading by keeping several epidemiological parameters and initial conditions constant. By varying the average degree and contact frequency we also address the containment measures implemented into the network to contain the spreads. The simulation results gave us a particular insight into how the social network structures determine epidemic behaviors such as the peak values of infection curves, the time to reach them, the increasing and decreasing rate of the infection curves, the infection spread numbers, and the final size of the epidemic. Here, we try to describe the disease spreading between the population that implements the containment measures and the one that does not. Next, from the model, we can develop a model that incorporates more interventions to derive efficient and realistic strategies for curbing the spread of disease throughout the network.

Method

The first step to generate a synthesized human interaction network with size N and an average degree $\bar{k}$ is using the simplest examples of a random network in which we fix only the number of nodes N and the number of edges $L = \bar{k}N/2$. We take N nodes and place L edges among them at random[19]. The next step is to attach each edge with a natural number which is randomly chosen from set $W = \{1, 2, \ldots, \omega\} \subseteq \mathbf{N}$, where ω is the maximum of contact frequency. Here, we obtain the weighted adjacency matrix $\mathbf{A} = (\omega_{ij})$ with $\omega_{ij} \in W \cup \{0\}$ to represent the network. We assume that the network is undirected. If $\omega_{ij} = 0$, then there is no interaction or link between node i and node j. From $\mathbf{A}$, we can obtain N_i as a set containing all neighbors of node i.

To simplify the discrete-time process we assume that one day is a unit time step and we write $t = 0, 1, 2, \ldots, t_f$ as a time discretization. We denote t_f as the duration of the epidemic process. At time t, a node i can be in a state X_t^i belonging to a finite set of states $\Sigma = \left\{ S_t^i, I_{1_t}^i, R_{1_t}^i, I_{2_t}^i, R_{2_t}^i, I_{3_t}^i, R_{3_t}^i \right\}$. The state $X_t^i \in \{0, 1\}$ and

$$S_t^i + I_{1_t}^i + R_{1_t}^i + I_{2_t}^i + R_{2_t}^i + I_{3_t}^i + R_{3_t}^i = 1 \text{ for all nodes } i \text{ and for all time } t \tag{1}$$

Since we put three variants into the network, we write $v = 1, 2, 3$ to denote the variants, and the original virus is Variant 1. All states are explained in the following table (Table 1).

To introduce the infections due to the Variant 1 at the initial time, we choose randomly N_0 individuals. Variant 2 will emerge from Variant 1 by moving the infected individual in the state $I_{1_t}^i$ to the state $I_{2_t}^i$ which is denoted by a Bernoulli random variable $\mu_{1 \to 2}$ with a small probability. Variant 3 will emerge from Variant 2 with a similar

(a)

(b)

(c)

(d)
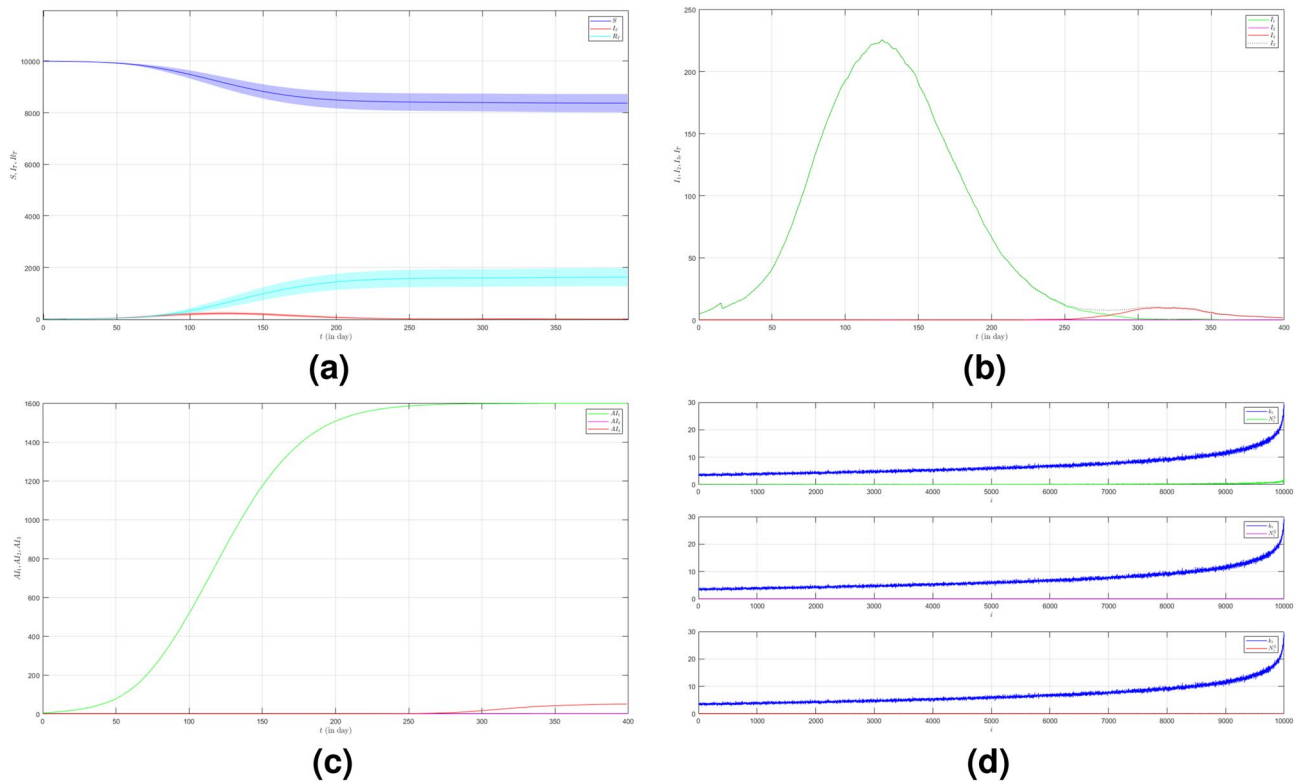
Figure 14. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the third case of Scenario 3. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{126}} = 226$, $I_{2_{127}} = 1$, and $I_{3_{314}} = 10$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 1599$, $AI_{2_{400}}^T = 1$, and $AI_{3_{400}}^T = 51$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.121$, $\rho_2 = 0$, and $\rho_3 = 0.002$ respectively.

procedure. The infected individual in the state $I_{2_t}^i$ will move to the state $I_{3_t}^i$ which is denoted by a Bernoulli random variable $\mu_{2 \to 3}$ with a small probability.

For each variant $v = 1, 2, 3$, we let ξ_{v_t} as a Bernoulli random variable with a probability β_v to denote viral transmission between two nodes. We define that β_v is the average rate of infection in one contact for Variant v. The process of viral transmission from one of the infected neighbors of node i to node i can be illustrated in Fig. 21. The viral transmission to node i at time t for each variant can be expressed by the following equation.

$$\phi_{v_t}^i = 1 - \prod_{j \in N_i} \prod_{k=1}^{\omega_{ij}} \left(1 - \xi_{v_t} I_{v_{t-1}}^j\right) \tag{2}$$

Note that the Eq. (2) describes how viral transmission occurs at the individual level or microscopic scale. Not only does the equation consider all neighbors for each individual but also considers the frequency of contact with all neighbors. It also tells us that the bigger the number of neighbors and weight, the more likely a node i to get an infection.

To make clear when the neighbors of node i contain the infected individuals with more than one variant, the viral transmission to node i occurs when the first infected individual succeeds in transmitting the infection. With this rule, we avoid an individual getting more infections at one time. Thus, we write

$$\phi_{1_t}^i + \phi_{2_t}^i + \phi_{3_t}^i = 1 \quad \text{for all node } i \text{ and for all time } t \tag{3}$$

We assume that each individual will gain permanent immunity after she or he recovers from one variant but lose immunity if she or he gets infected by other variants.

Unlike the common SIR model, we do not use the recovery rate. Instead, we use the parameters t_v^i, τ_v and the unit-step functions $u(t, t_v^i)$ and $u(t, t_v^i + \tau_v)$ for each variant v. The parameter t_v^i is the time when a node i gets infected by Variant v for the first time. Meanwhile, the parameter τ_v is the average duration of the infectious period for each variant. The functions $u(t, t_v^i)$ and $u(t, t_v^i + \tau_v)$ will move the individual from the susceptible state to the infected state and from the infected state to the recovery state respectively. The process of state transition for each node in the network is shown in Fig. 22 and the process is governed by the following system of equations.
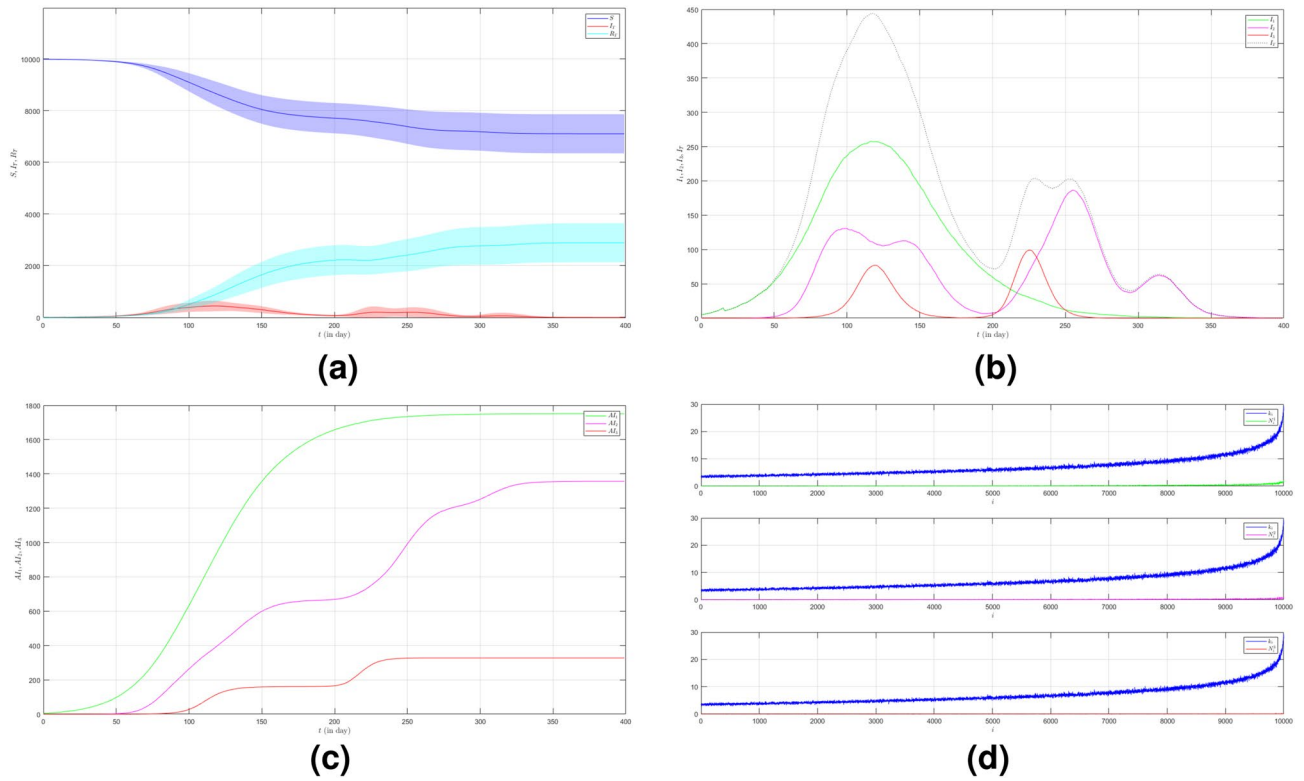
Figure 15. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the fourth case of Scenario 3. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{118}} = 258$, $I_{2_{256}} = 187$, and $I_{3_{226}} = 100$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^{T} = 1751$, $AI_{2_{400}}^{T} = 1357$, and $AI_{3_{400}}^{T} = 328$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.141$, $\rho_2 = 0.111$, and $\rho_3 = 0.008$ respectively.
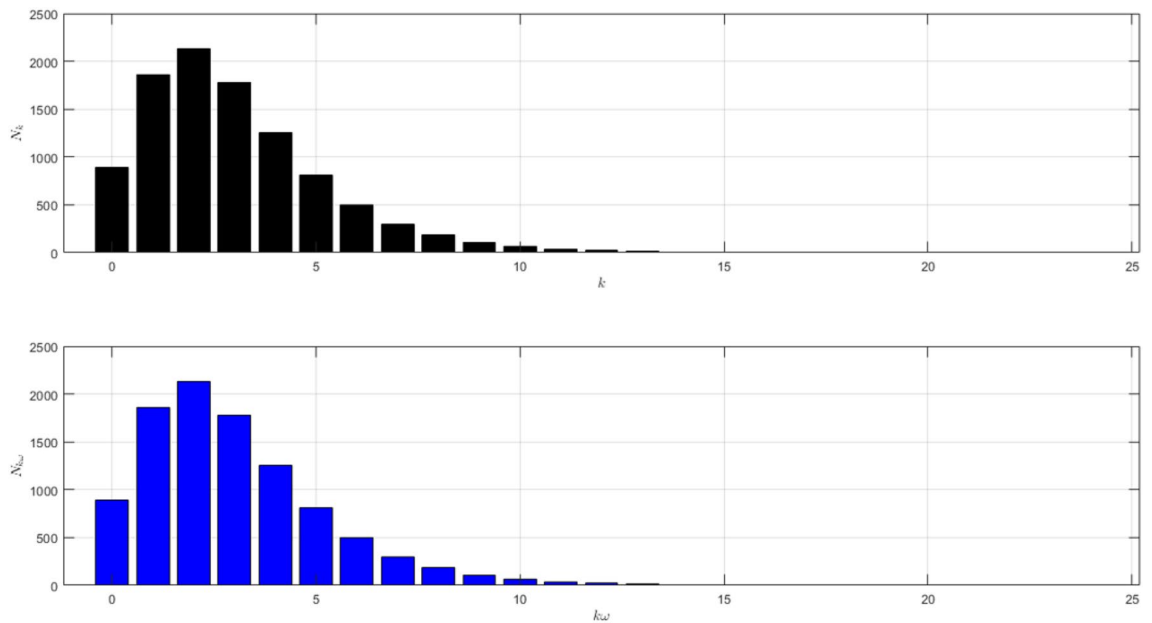
Figure 16. The network structure of Scenario 4. The black histogram shows the network's degree distribution with $\bar{k} = 3$ and $\omega = 1$. The black histogram is the same as the blue one. There is no hub in this network.
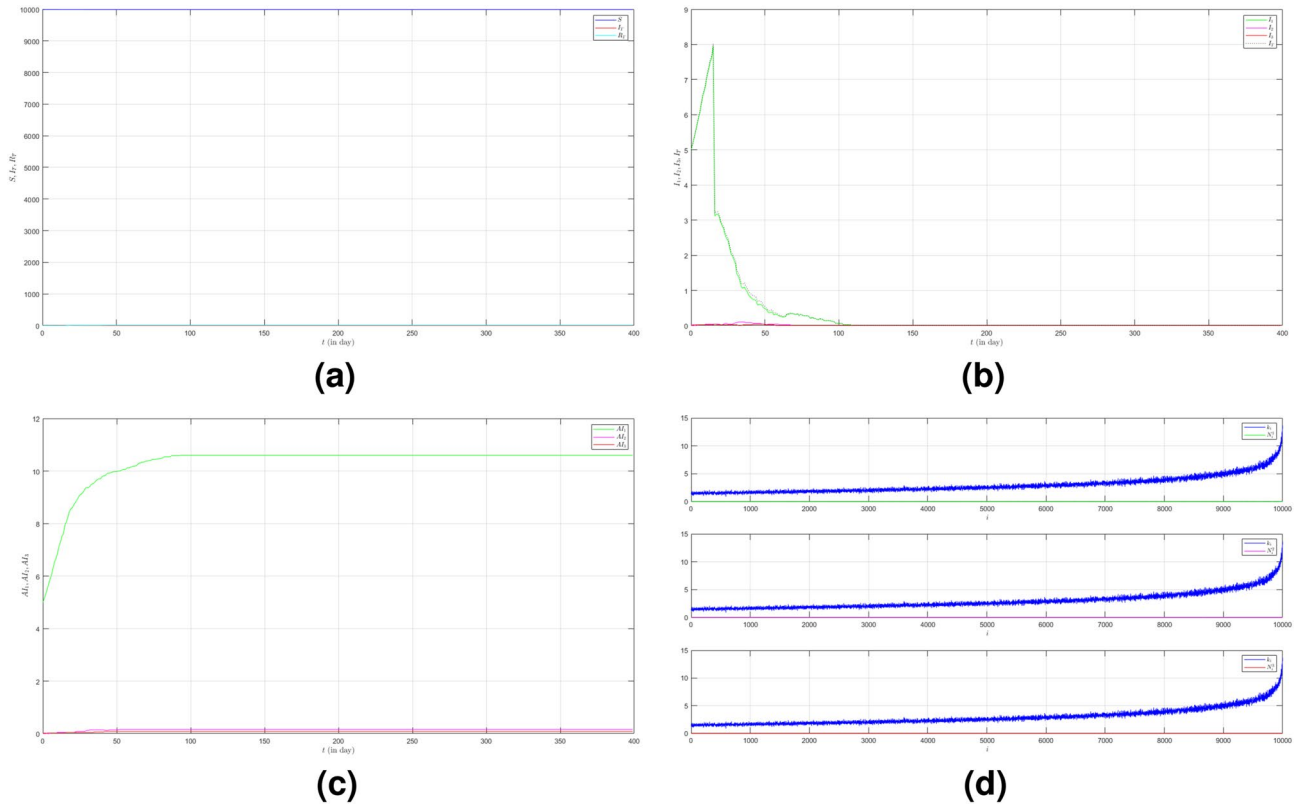
Figure 17. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the first case of Scenario 4. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of Variant 1 and Variant 2 are indicated by $I_{1_{16}} = 8$ and $I_{2_{33}} = 1$. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 11$, $AI_{2_{400}}^T = 1$, and $AI_{3_{400}}^T = 0$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.0001$, $\rho_2 = 0$, and $\rho_3 = 0$ respectively.

$$S_{t+1}^i = S_t^i \left(1 - \left(\phi_{1_{t+1}}^i + \phi_{2_{t+1}}^i + \phi_{3_{t+1}}^i \right) \right) \tag{4}$$

$$I_{1_{t+1}}^i = I_{1_t}^i \left(1 - u\left(t, t_1^i + \tau_1\right) - \mu_{1\to 2} \right) + \phi_{1_{t+1}}^i \left(S_t^i + R_{2_t}^i + R_{3_t}^i \right) \tag{5}$$

$$I_{2_{t+1}}^i = I_{2_t}^i \left(1 - u\left(t, t_2^i + \tau_2\right) - \mu_{2\to 3} \right) + \mu_{1\to 2} I_{1_t}^i + \phi_{2_{t+1}}^i \left(S_t^i + R_{1_t}^i + R_{3_t}^i \right) \tag{6}$$

$$I_{3_{t+1}}^i = I_{3_t}^i \left(1 - u\left(t, t_3^i + \tau_3\right) \right) + \mu_{2\to 3} I_{2_t}^i + \phi_{3_{t+1}}^i \left(S_t^i + R_{1_t}^i + R_{2_t}^i \right) \tag{7}$$

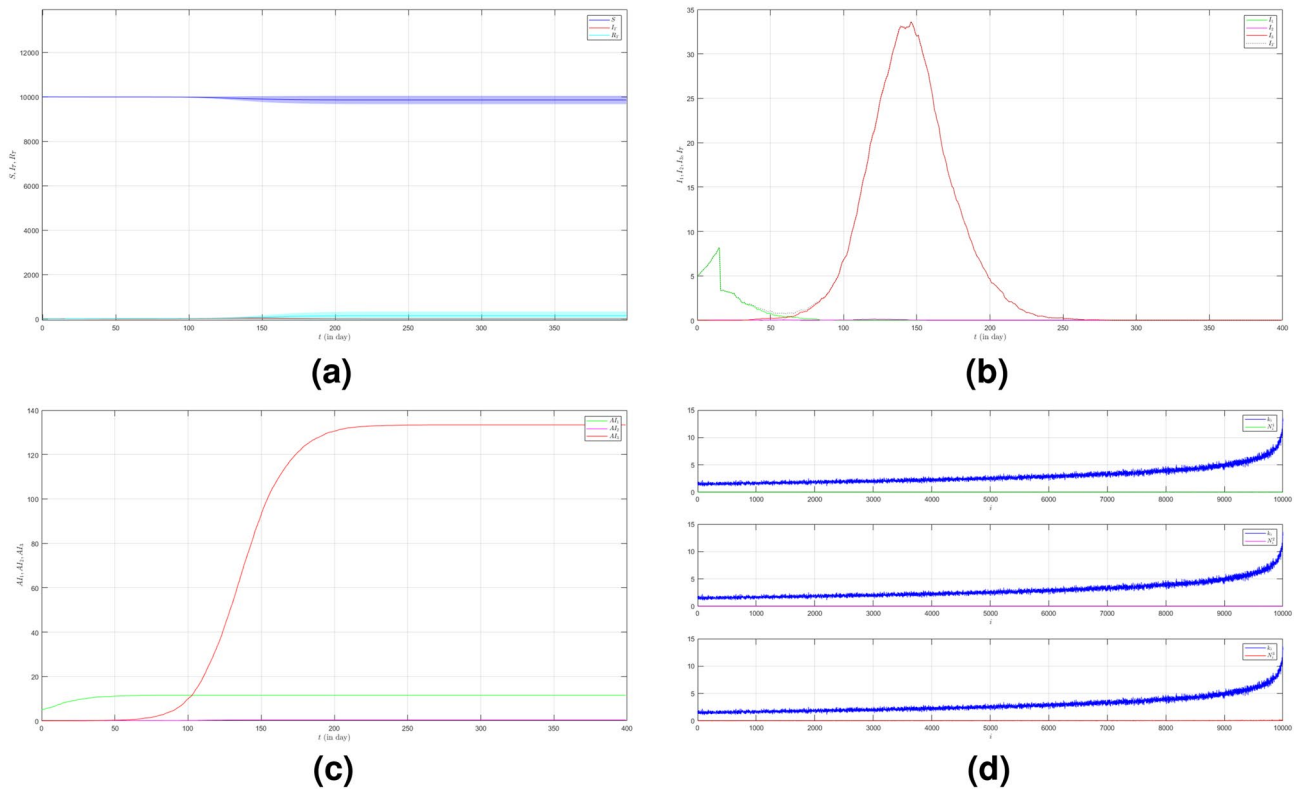$$R_{1_{t+1}}^i = R_{1_t}^i \left(1 - \left(\phi_{2_{t+1}}^i + \phi_{3_{t+1}}^i \right) \right) + u\left(t, t_1^i + \tau_1\right) I_{1_t}^i \tag{8}$$

$$R_{2_{t+1}}^i = R_{2_t}^i \left(1 - \left(\phi_{1_{t+1}}^i + \phi_{3_{t+1}}^i \right) \right) + u\left(t, t_2^i + \tau_2\right) I_{2_t}^i \tag{9}$$

$$R_{3_{t+1}}^i = R_{3_t}^i \left(1 - \left(\phi_{1_{t+1}}^i + \phi_{2_{t+1}}^i \right) \right) + u\left(t, t_3^i + \tau_3\right) I_{3_t}^i \tag{10}$$

To indicate whether a node i has been infected, we use the following equations.

$$AI_{1_{t+1}}^i = AI_{1_t}^i + \phi_{1_{t+1}}^i \left(S_t^i + R_{2_t}^i + R_{3_t}^i \right) \tag{11}$$

$$AI_{2_{t+1}}^i = AI_{2_t}^i + \phi_{2_{t+1}}^i \left(S_t^i + R_{1_t}^i + R_{3_t}^i \right) \tag{12}$$
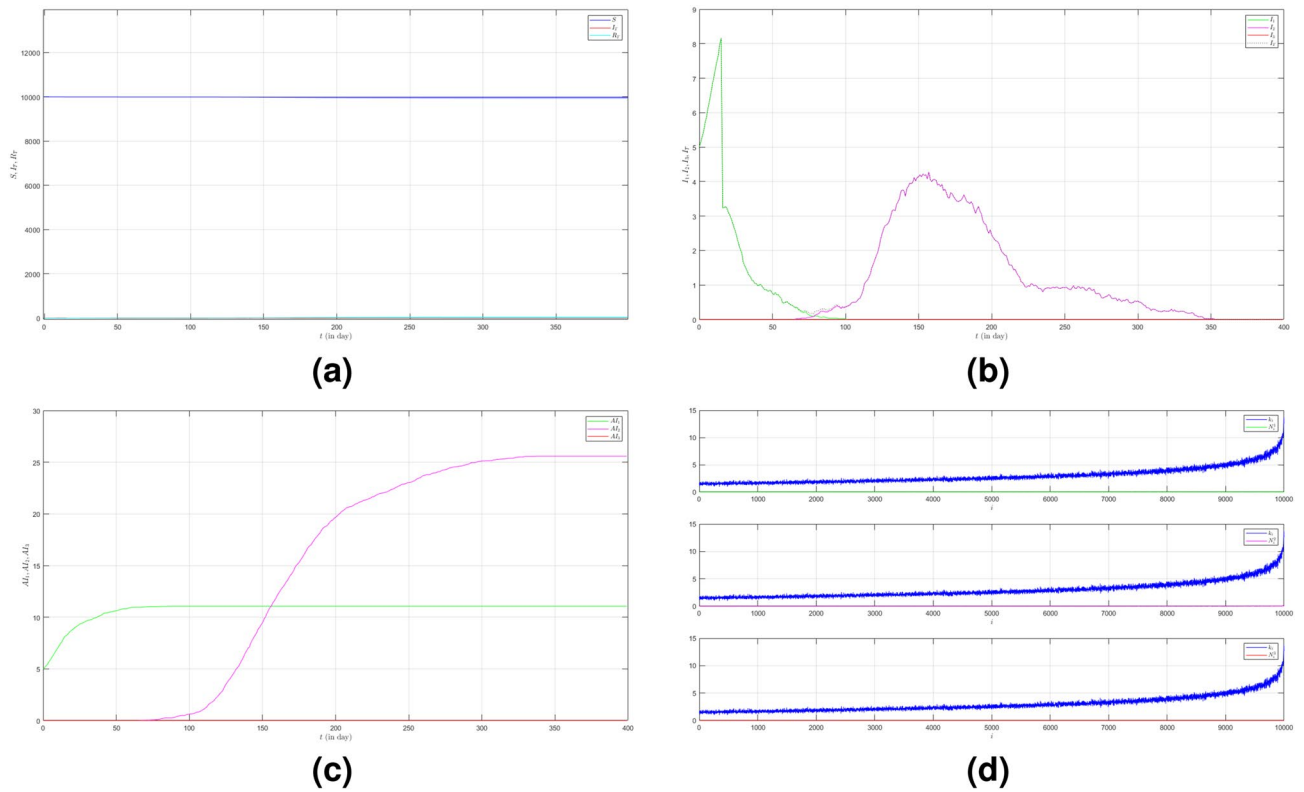
Figure 18. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the second case of Scenario 4. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of all variants are indicated by $I_{1_{116}} = 9$, $I_{2_{121}} = 1$, and $I_{3_{147}} = 34$ respectively. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 12$, $AI_{2_{400}}^T = 1$, and $AI_{3_{400}}^T = 134$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.0001$, $\rho_2 = 0$, and $\rho_3 = 0.005$ respectively.

$$AI_{3_{t+1}}^i = AI_{3_t}^i + \phi_{3_{t+1}}^i \left(S_t^i + R_{1_t}^i + R_{2_t}^i \right) \tag{13}$$

Based on Eqs. (4)–(10), we can calculate the number of nodes for each state at time *t* with the following equations.

$$S_t = \sum_{i=1}^N S_t^i, \ I_{v_t} = \sum_{i=1}^N I_{v_t}^i, \ R_{v_t} = \sum_{i=1}^N R_{v_t}^i, \ v = 1, 2, 3 \tag{14}$$

The final number of infections for each variant at the final time can be calculated by the following equations.

$$AI_1^T = \sum_{i=1}^N AI_{1_t}^i, \ AI_2^T = \sum_{i=1}^N AI_{2_t}^i, \ AI_3^T = \sum_{i=1}^N AI_{3_t}^i \tag{15}$$
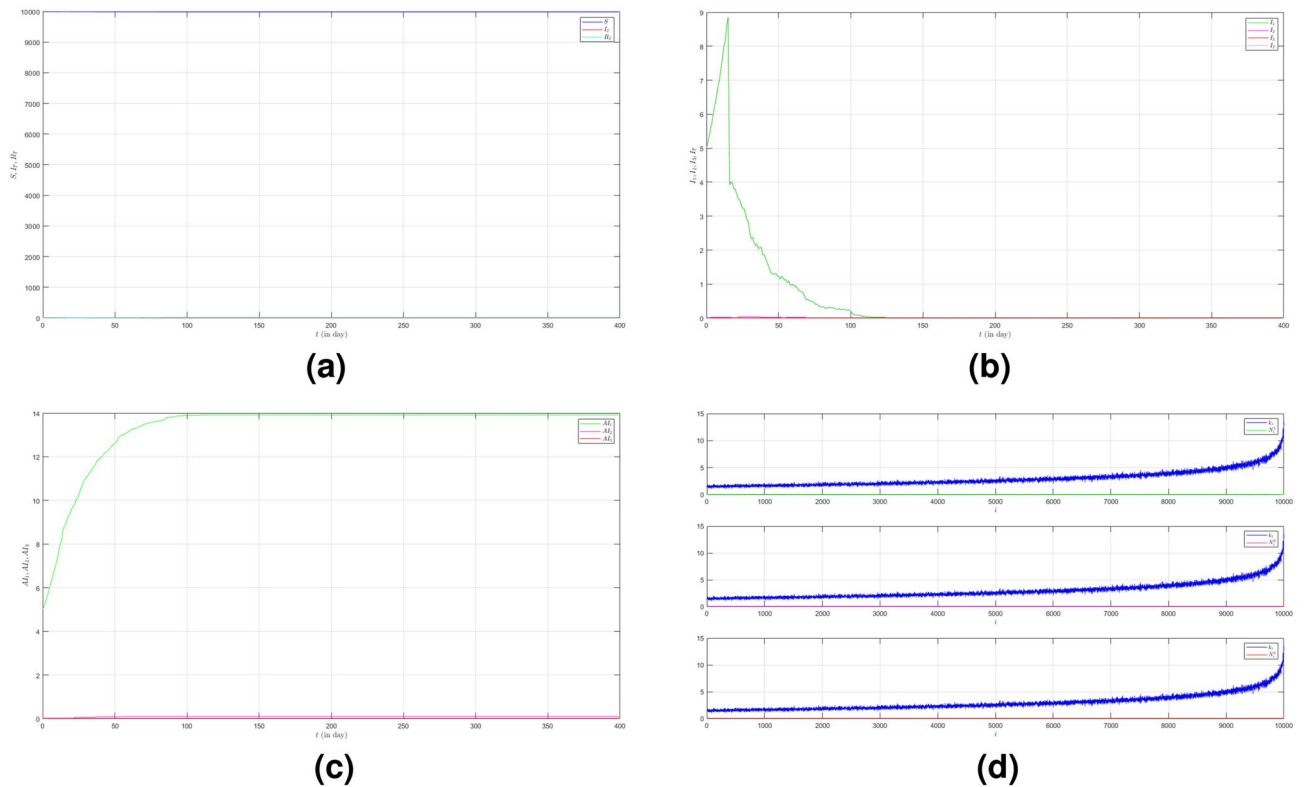
Figure 19. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the third case of Scenario 4. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of Variant 1s is $I_{1_{16}} = 9$. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 14$, $AI_{2_{400}}^T = 0$, and $AI_{3_{400}}^T = 0$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.0002$, $\rho_2 = 0$, and $\rho_3 = 0$ respectively.

In epidemiology, we used to consider basic reproduction number $\mathscr{R}_0$ to describe the ability of the disease to spread in a population. It is an average number of additional infected individuals produced by such individual passes the disease onto before it recovers[19]. Referring to the classic SIR model, when $\mathscr{R}_0 > 1$, infected individuals will grow geometrically, causing an epidemic[26]. When $\mathscr{R}_0 < 1$, the disease will die out. It can be calculated straightforwardly by finding the ratio of infection and recovery rate. However, it is not easy to find the number from our model. We do not have a recovery rate for each variant. In addition, our model uses three variants and relies on the microscopic process in discrete time. We consider the epidemic process on an individual scale. The size of the population is limited to a certain number which is much smaller than the size of the population in the common SIR model. Trying to find the basic reproduction number of the model in this research is no longer relevant. Thus, we do not deal with the basic reproduction number for our model because the model we developed is fully different from the common SIR model. Instead of using it, we define a number that represents how big and long an epidemic spreads. We call the number the infection spread number and denote it with ρ. We define it explicitly for each variant as the following.

$$\rho_v = \frac{AI_{v_{t_f}}^T \, t_{e_v}}{N \, t_f}, \; v = 1, 2, 3 \tag{16}$$

where t_{e_v} is the length of epidemic time which is defined as when the infection starts to spread and when it ends. The infection spread number has no dimension. We can interpret the number to infer how severe the epidemic is. The larger its value takes, the bigger and the longer, the size and duration of the epidemic will be.

(a)

(b)

(c)

(d)

Figure 20. The dynamics of the epidemic process with three variants on a synthesized human interaction network for the fourth case of Scenario 4. (**a**) The time evolution of all states with 95% confidence interval. The red lines indicate all infected nodes from all variants. (**b**) The average of 100 infection plots for each variant. The infection peaks of Variant 1 and Variant 2 are $I_{1_{16}} = 9$ and $I_{2_{158}} = 614$. (**c**) The total number of infections for all variants are $AI_{1_{400}}^T = 12$, $AI_{2_{400}}^T = 26$, and $AI_{3_{400}}^T = 0$. (**d**) Comparison between the number of infected neighbors and each node's degree for all variants. The infection spread numbers are $\rho_1 = 0.0001$, $\rho_2 = 0.0007$, and $\rho_3 = 0$ respectively.

State	Description
S_t^i	State at time t when a node i is susceptible ($S_t^i = 1$) or not ($S_t^i = 0$)
$I_{v_t}^i$	State at time t when a node i is infected by Variant j ($I_{v_t}^i = 1$) or not ($I_{v_t}^i = 0$)
$R_{v_t}^i$	State at time t when node i is recovered from Variant j ($R_{v_t}^i = 1$) or not ($R_{v_t}^i = 0$)

Table 1. The Description of states.
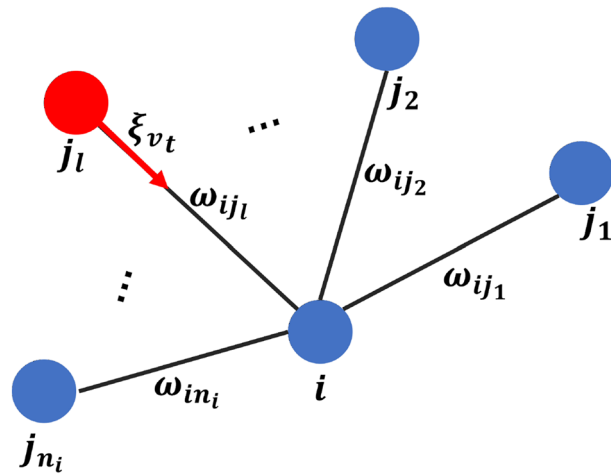
Figure 21. The viral transmission to node i from its neighbours. The red node represents an infected one. The blue nodes can be susceptible or recovered from previous variants.
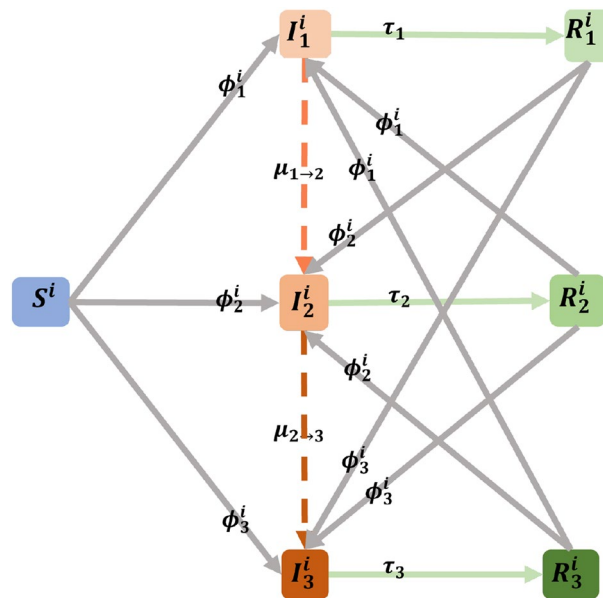
Figure 22. The process of all possibilities of state transition for each node in the network at a single time t.

Data availability

References

1. WHO. Covid-19 new variants: Knowledge gaps and research (2020).
2. Cacciapaglia, G. *et al.* Epidemiological theory of virus variants. *Phys. A*https://doi.org/10.1016/j.physa.2022.127071 *(2022)*.
3. Vasireddy, D., Vanapharty, R., Mohan, G., Malayala, S. V. & Atluri, P. Review of Covid-19 variants and Covid-19 vaccine efficacy: What the clinician should know?. *J. Clin. Med. Res.* **12**(6), 317–325. https://doi.org/10.14740/jocmr4518 (2021).
4. Thorne, E. Understanding the Covid-19 variants of concern (2021).
5. Roy, B., Dhillon, J. K. & Habib, N. Global variants of Covid-19: Current understanding. *J. Biomed. Sci.* **8**(1), 8–11. https://doi.org/10.3126/jbs.v8i1.38453 (2021).
6. ECDC. Threat assessment brief: Implications of the further emergence and spread of the SARS CoV 2 B. 1.1. 529 variant of concern (Omicron) for the EU/EEA first update (2021).
7. Worldometer. Corona updates (2021).

8. Charumilind, S. *et al.* When will the Covid-19 pandemic end?. *Indian J. Physiol. Allied Sci.* **74**, 19–24 (2022).
9. Sanjuan, R. & Domingo-Calap, P. Mechanism of viral mutation. *Cell. Mol. Life Sci.* **73**(23), 4433–4448. https://doi.org/10.1007/2Fs00018-016-2299-6 (2016).
10. Cruickshank, S. How viruses mutate and what can be done about it (2021).
11. Finucane, M. Will the coronavirus ever stop mutating? Experts explain how omicron and other variants emerge (2021).
12. Sabin, N. S. *et al.* Implications of human activities for (re)emerging infectious diseases, including Covid-19. *J. Physiol. Anthropol.* https://doi.org/10.1007/2Fs00018-016-2299-6 *(2020)*.
13. Meng, X., Cai, Z., Si, S. & Duan, D. Analysis of epidemic vaccination strategies on heterogenous networks: Based on Seirv model and evolutionary game. *Appl. Math. Comput.* **403**, 200. https://doi.org/10.1016/j.amc.2021.126172 (2021).
14. Rella, S. A., Kulikova, Y. A., Dermitzakis, E. T. & Kondrashov, F. A. Rates of SARS-CoV-2 transmission and vaccination impact the fate of vaccine-resistant strains. *Sci. Rep.* **11**, 200. https://doi.org/10.1038/s41598-021-95025-3 (2021).
15. Lobinska, G., Pauzner, A., Traulsen, A., Pilpel, Y. & Nowak, M. A. Evolution of resistance to Covid-19 vaccination with dynamic social distancing. *Nat. Hum. Behav.* **6**, 193–206. https://doi.org/10.1038/s41562-021-01281-8 (2022).
16. Nuraini, N., Khaerudin, K., Susanto, H., Hasan, A. & Sumarti, N. Mathematical models for assessing vaccination scenarios in several provinces in Indonesia. *Infect. Dis. Model.* **6**, 1236–1258. https://doi.org/10.1016/j.idm.2021.09.002 (2021).
17. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **115**, 700–721 (1927).
18. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans* 1st edn. (Oxford University Press, 1991).
19. Newman, M. E. J. *Networks: An introduction* 1st edn. (Oxford University Press, 2010).
20. Keeling, M. J. & Eames, K. T. D. Networks and epidemic models. *J. R. Soc. Interface* https://doi.org/10.1098/rsif.2005.0051 *(2005)*.
21. Okabe, Y. & Shudo, A. Spread of variants of epidemic diseases based on the microscopic numerical simulations on networks. *Sci. Rep.* https://doi.org/10.1038/s41598-021-04520-0 *(2022)*.
22. Mossong, J., Hens, N., Jit, M., Beutels, P. & Auranen, K. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, 381–391. https://doi.org/10.1371/journal.pmed.0050074 (2008).
23. Dai, J. *et al.* Modeling the spread of epidemics based on cellular automata. *MDPI* **9**, 55. https://doi.org/10.3390/pr9010055 (2021).
24. Herrmann, H. A. & Schwartz, J. M. Why Covid-19 models should incorporate the network of social interactions. *Phys. Biol.* **17**(6), 065008 (2020).
25. Okabe, Y. & Shudo, A. Microscopic numerical simulations of epidemic models on networks. *Mathematics* **9**(9), 932. https://doi.org/10.3390/math9090932 (2021).
26. Brauer, F., Driessche, P. & Wu, J. *Mathematical Epidemiology* (Springer, 2008).

Author contributions

Seprianus write the main manuscript and prepare all figures. All authors reviewed the manuscript and give some feedbacks.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.