Research Article

Epidemic Spreading Combined with Age and Region in Complex Networks

Xu Zhang,1 Yurong Song,2 Haiyan Wang,3 and Guo-Ping Jiang2

1School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
2College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China
3School of Mathematical and Natural Sciences, Arizona State University, Phoenix, AZ 85069, USA

Correspondence should be addressed to Yurong Song; songyr@njupt.edu.cn

Received 6 April 2020; Accepted 16 May 2020; Published 22 June 2020

Guest Editor: Michael Z. Q. Chen

Copyright © 2020 Xu Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In social networks, the age and the region of individuals are the two most important factors in modeling infectious diseases. In this paper, a spatial susceptible-infected-susceptible (SIS) model is proposed to describe epidemic spreading over a network with region and age by establishing several partial differential equations. Numerical simulations are performed, and the simulation of the proposed model agrees well with real influenza-like illness (ILI) in the USA reported by the Centers for Disease Control (CDC). Moreover, the proposed model can be used to predict the infected density of individuals. The results show that our model can be used as a tool to analyze influenza cases in the real world.

1. Introduction

Epidemic spreading [1], which was dramatic historical events, continues to pose health threats to humans today. Recently, epidemic outbreaks have caused the death of many people, such as during the spread of SARS [2] and H1N1 [3, 4] influenza. Thus, to reduce the danger of epidemic spreading, the study of the dynamics of epidemics is an important issue and has raised a great deal of concern. The spread of epidemics in complex networks [5] has been extensively studied by many researchers from different disciplines, including computer science, mathematics, biology, and physics.

Mathematical modeling is a useful tool that has been used to reveal many phenomena of disease propagation in complex networks. One of the most widely used models for the spread of an infectious disease is the susceptible-infected-susceptible (SIS) model [6–8], where the disease is transmitted from infected individuals to their susceptible neighbors with an infection rate, and the infected individuals can recover to become susceptible again with a recovery rate.

In the real world, many studies on real diseases, such as cholera [9, 10], have revealed that diseases might have different infection rates and mortality rates for different age groups [11]. Individuals of different ages might also have different behaviors, and behavioral changes are crucial in the control and prevention of many infectious diseases. Young individuals tend to be more active in interactions with or between populations and in disease transmissions [12]. Thus, many investigators have developed the age-structured models, composed of partial differential equations [13], for the spread of the epidemics. The age-structured models [14, 15], where the density of the infected individuals was expressed as a function of multiple independent variables, and the epidemic process were modeled as partial differential equations. Kuniya [16] studied the global asymptotic stability by discretizing the age-structured multigroup model. Inaba et al. [17] established an age-structured model of epidemic spreading for the demographic transition and obtained the stability condition using reproduction numbers. So et al. [18] derived the equation of a reaction-diffusion model for a single species population with the age structure.

Meanwhile, the other significant factor is spatial location [19]. Individuals in different regions may have different reproductive and survival capacities. Especially, disease or
information spreading has different behaviors in different spatial locations. Therefore, several researchers have proposed epidemic models based on partial differential equations on a network whose underlying edge represents the physical distance between nodes. Bustamante-Castaneda et al. [20] extended a Kermack–McKendrick model to a geographical network, and different parameters influenced this model and obtained a simple criterion for the onset of the epidemic. The characteristics of influenza are more diverse in subtropical and tropical regions [21]. Wang et al. [22] established a number of partial differential equations of the second order over networks to characterize information spreading in temporal and spatial dimensions.

From the above, studies usually focus on the epidemic models with either the region or age factor based on partial differential equations separately. The main contribution of this study is to combine age and region together based on partial differential equations. Thus, in this paper, we will establish several partial differential equations to describe the epidemic spreading in combination with age and region.

This paper has the following structure: the classification of individuals in complex networks is introduced in Section 2, and the mathematical model is constructed in Section 3. The numerical simulations are given in Section 4. In Section 5, we provide concluding remarks.

2. Embedding of a Network with Age and Region to the X-Axis

In social networks, individuals are likely to come into contact with individuals of the same age, especially during the teenage years and in old age. Students of the same age are always in the same grade, and apart from their parents and teachers, they are most exposed to their peers. Additionally, individuals of different ages have different behaviors in epidemic spreading. Therefore, the individuals in complex networks are divided into different communities based on an age structure. Meanwhile, individuals are always divided into different communities by distance. For example, a person is always in contact with his neighbors, friends, colleagues, or others who are in the same place or same region, and together, they form a community in a contact network. Thus, from the beginning, individuals in complex networks can be divided into different communities based on their spatial location.

First, in a social network or complex network, we consider the cluster of individuals of different classifications: age and region. We group individuals of the same age and region together in the complex network.

Suppose that $P(t)$ denotes the total population in the networks at time $t$. Individuals in networks have their age factors and spatial locations or regions. We combine region and age as group $\{r_n, a_m\}$, where the region $r$ is divided into several groups $r = \{r_1, r_2, \ldots, r_N\}$ and the age $a = \{a_1, a_2, \ldots, a_M\}$. Here, $r_N$ is the maximum region group, and $a_M$ is the maximum age group. Based on this group, one could divide the total population into a set of groups, i.e., $P(t) = \{P_{r_n, a_m}(t)\}$. Individuals in group $P_{r_n, a_m}(t)$ share the same age $a_m$ in the same region $r_n$, where we denote group $\{r_n, a_m\}$ as the combination of region and age. Then, we use the $x$-axis as the group and embed the density $P_x(t)$ at the location $x$, where $x$ denotes the combination of region and age, satisfying

$$x = m + (n - 1) \times M.$$  \hspace{1cm} (1)

Let $\rho(x, t)$ represent the density $P_x(t)$ at the location $x$, where $x \in (0, X]$, and $X \in (0, +\infty)$ is the upper bound of $x$. By setting $m = M$ and $n = N$ into equation (1), we obtain

$$X = MN.$$ \hspace{1cm} (2)

Then, we have

$$\int_0^X \rho(x, t)dx = P(t).$$ \hspace{1cm} (3)

3. Partial Differential Equation Model

This section describes the system of differential equations that describes the multifactor epidemic model. We aim to establish a realistic model that can provide a broad perspective for the prediction and control of disease propagation in real-world networks.

Individuals are classified into different groups. Based on this classification, we establish several equations that describe the evolution of the individuals in the classification system. Then, we establish a new SIS model to analyze the spread of the epidemic.

Generally, we consider a given human population that is divided into two classes: susceptible and infected. At each time step, each individual adopts one of these two states. During one time step, a susceptible individual will be infected when he comes into contact with an infected individual. Meanwhile, the infected individual will become susceptible when he has recovered.

For each classification, we consider both the density of the combination of region and age and the population changes over time. Thus, we describe the density of humans in each class as $s = s(x, t)$ and $i = i(x, t)$, which are susceptible density and infected density of the combination of region and age $x$ and time $t$, respectively.

$$\rho(x, t) = s(x, t) + i(x, t).$$ \hspace{1cm} (4)

The dynamic equation can be written as

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)s(x, t) = \gamma(x) i(x, t) - \lambda(x) s(x, t) i(x, t) - \mu(x) s(x, t),$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)i(x, t) = -\gamma(x) i(x, t) + \lambda(x) s(x, t) i(x, t) - \mu(x) i(x, t),$$

$$s(x, 0) = \phi(x), i(x, 0) = \varphi(x), \quad x > 0,$n

$$s(0, t) = \xi(t), i(0, t) = \zeta(t), \quad t > 0.$$ \hspace{1cm} (5)
where \( \mu(x) \) is the age-region-specific mortality rate, \( \gamma(x) \) is the recovery rate which is a function of \( x \), and \( \lambda(x) \) is the infection rate as a function of \( t \). \( \phi \) and \( \varphi \) and \( \xi \) and \( \zeta \) are the initial distributions of the susceptible and infected individuals, respectively. Additionally, there is the birth condition, which is assumed that all newborns are susceptible:

\[
s(0, t) = \int_0^{\infty} \beta(x)p(x, t)dx,
\]

where \( \beta(x) \) is the birth rate.

In the real world, the number of deaths due to the spreading of disease is far less than the number of infected individuals. Therefore, we can ignore the death rates and assume that there is no migration.

For simplicity, we assume that \( \rho(x, t) =: \rho^*(x) \) is independent of time, and the population is \( \int_0^X \rho(t, x)dx = \int_0^X \rho^*(x)dx \), which is in a stationary state.

\[
\frac{\partial i(x, t)}{\partial t} + \frac{\partial i(x, t)}{\partial x} = -\gamma(x)i(x, t) + \lambda(x)(\rho^*(x) - i(x, t))i(x, t),
\]

where the boundary conditions are \( i(x, 0) = \varphi(x), x > 0 \), and \( i(0, t) = \xi(t), t > 0 \). The first term in equation (7) considers the infected individuals whose location is \( x \). The second term considers the probability that an individual with \( x \) is healthy \( s(x, t) \) and will become infected via a connection with an infected individual. The transmission dynamics of the disease are governed by equation (7), where a susceptible individual with location \( x \) becomes infected with the probability \( \lambda(x) \) when the individual connects to an infected one, while an infected individual becomes susceptible with the recovery rate \( \gamma(x) \) spontaneously. It is worth mentioning that the conclusion in this paper can be extended to other epidemic models, such as the SIR model.

4. Simulation

To support our model, we verify it with real weekly data on influenza-like illness (ILI) in the community from the 42nd week of 2017 to the 4th week of 2018, as reported by the CDC. This paper uses fifteen classifications and fifty classifications to conduct the simulations. To simplify the simulation, the unit of time is one week.

4.1. Simulation Results of Fifteen Classifications. The regions of the USA [19] defined by the CDC are shown in Figure 1. In our model simulation, we rezone the entire area into three new regions. Here, we set \( M = 5 \) and \( N = 3 \). The 1st, 2nd, 3rd, and 4th regions are reazoned into Region 1; the 5th, 6th, and 7th regions are reazoned into Region 2; and the 8th, 9th, and 10th regions are reazoned into Region 3. Meanwhile, the ages are divided into five intervals, namely, [0, 4], [5, 24], [25, 49], [50, 64], and [65, \( \infty \)], where \( A \) is the upper bound of individuals. Here, setting \( M = 5 \) and \( N = 3 \) into equation (1), we obtain \( x = 1, 2, \ldots, 15 \), and the details are shown in Table 1. By combining the age and region, fifteen classifications are derived.

The initial guess parameter values are

\[
\lambda(x) = 0.01 + \exp[-5.20 \times (x - 2.70)],
\]

\[
\gamma(x) = 0.30 + \exp[-0.60 \times (x - 3.00)],
\]

where the exponent form is chosen based on the distribution of displacement lengths of individuals [23].

The final parameter values are

\[
\lambda(x) = 0.01 + \exp[-5.23 \times (x - 2.72)],
\]

\[
\gamma(x) = 0.30 + \exp[-0.60 \times (x - 3.02)].
\]

The information in Table 1 and Figure 2 shows how a change in the value of \( x \), the combined region and age, changes the density of the infected individuals, which means that all of the data have been divided into fifteen classifications. Based on the initial guess parameter values, we simulate the equation by adjusting the parameter values. Figure 2 gives the simulation results, where the red-dotted line is the data for the 19th week of ILI cases reported by the CDC. The blue solid line is the simulated result using our model in the 19th week. As shown in Figure 2, these two lines match well, especially for integer values of \( x \).

4.2. Simulation Results of Fifty Classifications. Here, we further set \( M = 5 \) and \( N = 10 \). The ages are divided as described in Section 4, and the regions are shown in Figure 1 according to the 10 regions of the USA. Setting \( M = 5 \) and \( N = 10 \) into equation (1), we obtain \( x = 1, 2, \ldots, 50 \), and the details are shown in Table 2. By combining the age and region, fifty classifications are obtained.

The initial guess parameter values satisfy equations (8) and (9). The information in Table 2 and Figure 3 shows that all the data have been divided into fifty classifications. Figure 3 gives the simulation results, where the red-dotted line is the data from the 18th week of ILI cases reported by the CDC. The blue solid line is the simulated result using our model in the 18th week. As shown in Figure 3, these two lines match well, especially at integer values of \( x \). The simulation of the model agrees well with the real cases provided by the CDC. Our model can very accurately simulate real conditions.
4.3 Prediction Accuracy. From the above, we can see that our model agrees well with the reported ILI cases provided by the CDC. Here, we use our model to predict the infected density of individuals in the next three weeks, and the prediction accuracy is calculated by using fifteen classifications mentioned in Section 4. The prediction accuracy of the model against the actual value is defined as

\[
\text{accuracy} = 1 - \frac{|\text{predv} - \text{actv}|}{\text{actv}},
\]

where \(\text{predv}\) is the predication value of our model and \(\text{actv}\) is the actual value of the real data reported by the CDC.

We use data from the 1\textsuperscript{st} week to the 4\textsuperscript{th} week for training, find the suitable parameters where simulation of our model agrees well with the data from the CDC, and predict the data of ILI cases reported by the CDC for the next three weeks. Following the same procedure, we train the data from the 1\textsuperscript{st} week to the 7\textsuperscript{th} week and predict the data from the 8\textsuperscript{th} week to the 10\textsuperscript{th} week; train the data from the 1\textsuperscript{st} week to the 10\textsuperscript{th} week and predict the data from the 11\textsuperscript{th} week to the 13\textsuperscript{th} week; train the data from the 1\textsuperscript{st} week to the 13\textsuperscript{th} week and predict the data from the 14\textsuperscript{th} week to the 16\textsuperscript{th} week. Figure 4 shows the average prediction accuracy for fifteen classifications. And the results demonstrate that our model can predict the data of ILI cases reported by the CDC for fifteen classifications with similar accuracy.

Figure 4 shows that a prediction accuracy of 85.01% is obtained if the data of the first 4 weeks are used for training. Similarly, a prediction accuracy of 76.95%, 83.20%, and 87.31% is obtained if the data of the first 7 weeks, first 10 weeks, and first 13 weeks are used for training, respectively. It can be seen that our model can obtain high prediction accuracy for fifteen classifications.

| \(x\) (combination of region and age) | \(m\) (age groups) | \(n\) (region groups) |
|-------------------------------------|-------------------|----------------------|
| 1                                  | From 0 to 4       | Region 1             |
| 2                                  | From 5 to 24      | Region 1             |
| 3                                  | From 25 to 49     | Region 1             |
| 4                                  | From 50 to 64     | Region 1             |
| 5                                  | From 65 to upper bound | Region 1          |
| 6                                  | From 0 to 4       | Region 2             |
| 7                                  | From 5 to 24      | Region 2             |
| 8                                  | From 25 to 49     | Region 2             |
| 9                                  | From 50 to 64     | Region 2             |
| 10                                 | From 65 to upper bound | Region 2          |
| 11                                 | From 0 to 4       | Region 3             |
| 12                                 | From 5 to 24      | Region 3             |
| 13                                 | From 25 to 49     | Region 3             |
| 14                                 | From 50 to 64     | Region 3             |
| 15                                 | From 65 to upper bound | Region 3          |

**Table 1:** Fifteen classifications for the combined ages and regions, \(x\).
4.4. Comparison and Summary. As shown in Figures 2 and 3, the curve simulated by our model has better performance with fifty classifications than with fifteen classifications. This is because in numerical simulations of differential equations, the denser the points, the more accurate the results. In summary, the empirical results agree well with the real data of weekly reported cases provided by the CDC. From Figure 4, a high accuracy is achieved to predict the reported IILI cases provided by the CDC. This indicates that our model can be used as a tool to analyze flu cases in terms of the regions or ages. Since the simulation of the model agrees well with the reported IILI cases provided by the CDC in terms of the regions or ages, we argue that our model can be used to analyze flu cases in real-world networks.

| $x$ (combination of age and region) | $m$ (age groups) | $n$ (region groups) |
|-----------------------------------|------------------|---------------------|
| 1 From 0 to 4                     | Region 1         |
| 2 From 5 to 24                     | Region 1         |
| 3 From 25 to 49                    | Region 1         |
| 4 From 50 to 64                    | Region 1         |
| 5 From 65 to upper bound           | Region 1         |
| 6 From 0 to 4                      | Region 2         |
| 7 From 5 to 24                     | Region 2         |
| 8 From 25 to 49                    | Region 2         |
| 9 From 50 to 64                    | Region 2         |
| 10 From 65 to upper bound          | Region 2         |
| 11 From 0 to 4                     | Region 3         |
| 12 From 5 to 24                    | Region 3         |
| 13 From 25 to 49                   | Region 3         |
| 14 From 50 to 64                   | Region 3         |
| 15 From 65 to upper bound          | Region 3         |
| 16 From 0 to 4                     | Region 4         |
| 17 From 5 to 24                    | Region 4         |
| 18 From 25 to 49                   | Region 4         |
| 19 From 50 to 64                   | Region 4         |
| 20 From 65 to upper bound          | Region 4         |
| 21 From 0 to 4                     | Region 5         |
| 22 From 5 to 24                    | Region 5         |
| 23 From 25 to 49                   | Region 5         |
| 24 From 50 to 64                   | Region 5         |
| 25 From 65 to upper bound          | Region 5         |
| 26 From 0 to 4                     | Region 6         |
| 27 From 5 to 24                    | Region 6         |
| 28 From 25 to 49                   | Region 6         |
| 29 From 50 to 64                   | Region 6         |
| 30 From 65 to upper bound          | Region 6         |
| 31 From 0 to 4                     | Region 7         |
| 32 From 5 to 24                    | Region 7         |
| 33 From 25 to 49                   | Region 7         |
| 34 From 50 to 64                   | Region 7         |
| 35 From 65 to upper bound          | Region 7         |
| 36 From 0 to 4                     | Region 8         |
| 37 From 5 to 24                    | Region 8         |
| 38 From 25 to 49                   | Region 8         |
| 39 From 50 to 64                   | Region 8         |
| 40 From 65 to upper bound          | Region 8         |
| 41 From 0 to 4                     | Region 9         |
| 42 From 5 to 24                    | Region 9         |
| 43 From 25 to 49                   | Region 9         |
| 44 From 50 to 64                   | Region 9         |
| 45 From 65 to upper bound          | Region 9         |
| 46 From 0 to 4                     | Region 10        |
| 47 From 5 to 24                    | Region 10        |
| 48 From 25 to 49                   | Region 10        |
| 49 From 50 to 64                   | Region 10        |
| 50 From 65 to upper bound          | Region 10        |
5. Conclusion

In real networks, age and region are two of the most important characteristics of epidemic processes. Individuals of different ages may engage in different behaviors in disease spreading. Individuals in different regions may also have different reproductive and survival capacities. Based on the SIS model, a new propagation model has been proposed to describe epidemics spreading combined with age and region as a system of partial differential equations. Then, numerical simulations have been performed to show that the simulation of this model agrees well with real influenza-like illness in the USA as reported by the CDC. This implies that our model can be a tool to analyze and predict flu cases with granularity at the regional or age level. However, some other factors, such as migration and time delay, are also important that impact the spreading of disease, which will be our near future work.

Data Availability

The data used to support the findings of this study may be released upon application to the Centers for Disease Control (CDC).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Xu Zhang, Yurong Song, and Guo-Ping Jiang were supported by the National Natural Science Foundation of China (Grant nos. 61672298, 61873326, 61802155, and 61802201) and the Philosophy Social Science Research Key Project Fund of Jiangsu University (Grant no. 2018SJZD1142).

References

[1] D. J. Daley, J. Gani, and J. M. Gani, Epidemic Modelling: An Introduction, Cambridge University Press, Cambridge, UK, 2001.
[2] C. Dye and N. Gay, “Modeling the SARS epidemic,” Science, vol. 300, no. 5627, pp. 1884-1885, 2003.
[3] J. Zarocostas, “World health organization declares A (H1N1) influenza pandemic,” BMJ, vol. 338, p. 2425, 2009.
[4] V. P. Maru, C. Nagarathna, B. S. Shakuntala, and H. K. Navin, “Mongrelised genetics of H1N1 virus: a bird’s eyewview,” Indian Journal of Dental Research, vol. 21, no. 4, pp. 586–590, 2010.
[5] R. Pastor-Satorras, C. Castellano, P. V. Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” Review of Modern Physics, vol. 87, no. 3, pp. 120–131, 2014.
[6] P. V. Mieghem, “The N-intertwined SIS epidemic network model,” Computing, vol. 93, no. 2–4, pp. 147–169, 2011.
[7] E. Cator and P. V. Mieghem, “Susceptible-infected-susceptible epidemics on the complete graph and the star graph: exact analysis,” Physical Review E, vol. 87, no. 1, Article ID 012811, 2013.
[8] C. Li, R. V. D. Bovenkamp, and P. V. Mieghem, “Susceptible-infected-susceptible model: a comparison of N-intertwined and heterogeneous mean-field approximations,” Physical Review E, vol. 86, no. 2, Article ID 026116, 2012.
[9] A. Alexanderian, M. K. Gobbert, K. R. Fister, H. Gaff, S. Lenhart, and E. Schaefer, “An age-structured model for the spread of epidemic cholera: analysis and simulation,” Nonlinear Analysis: Real World Applications, vol. 12, no. 6, pp. 3483–3498, 2011.
[10] A. Aghaksanierian and M. K. Gobbert, “Modeling the spread of epidemic cholera: an age-structured model,” 2007.
[11] R. M. Anderson and R. M. May, Infectious Diseases of Humans, Dynamics and Control, Oxford University Press, Oxford, UK, 1991.
[12] H. W. Hethcote, “The mathematics of infectious disease,” SIAM Review, vol. 42, no. 4, pp. 599–653, 2000.
[13] H. Wang and X.-S. Wang, “Traveling wave phenomena in a Kermack–McKendrick SIR model,” Journal of Dynamics and Differential Equations, vol. 28, no. 1, pp. 143–166, 2016.
[14] J. Wang, J. Lang, and X. Zou, “Analysis of an age structured HIV infection model with virus-to-cell infection and cell-to-cell transmission,” Nonlinear Analysis: Real World Applications, vol. 34, pp. 75–96, 2017.
[15] M. Cecchini, S. Cividino, R. Turco, and L. Salvati, "Population age structure, complex socio-demographic systems and resilience potential: a spatio-temporal, evenness-based approach," *Sustainability*, vol. 11, no. 7, p. 2050, 2019.

[16] T. Kuniya, "Global stability analysis with a discretization approach for an age-structured multigroup SIR epidemic model," *Nonlinear Analysis: Real World Applications*, vol. 12, no. 5, pp. 2640–2655, 2011.

[17] H. Inaba, R. Saito, and N. Bacaër, "An age-structured epidemic model for the demographic transition," *Journal of Mathematical Biology*, vol. 77, no. 5, pp. 1299–1339, 2018.

[18] J. W.-H. So, J. Wu, and X. Zou, "A reaction–diffusion model for a single species with age structure. I travelling wavefronts on unbounded domains," *Proceedings of the Royal Society A*, vol. 457, no. 2012, pp. 1841–1853, 2001.

[19] F. Wang, H. Wang, K. Xu et al., "Regional level influenza study with geo-tagged twitter data," *Journal of Medical Systems*, vol. 40, no. 8, p. 189, 2016.

[20] F. Bustamante-Castaneda, J.-G. Caputo, G. Cruz-Pacheco, A. Knippel, and F. Mouatamid, "Analysis of an epidemic model on a network," 2019, http://arxiv.org/abs/1906.07449.

[21] W. Yang, B. J. Cowling, E. H. Y. Lau, and J. Shaman, "Forecasting influenza epidemics in Hong Kong," *PLoS Computational Biology*, vol. 11, no. 7, Article ID e1004383, 2015.

[22] H. Wang, F. Wang, and K. Xu, "Modeling information diffusion in online social networks with partial differential equations," in *Surveys and Tutorials in the Applied Mathematical Sciences*, Springer, Berlin, Germany, 2020.

[23] D. H. Barmak, C. O. Dorso, and M. Otero, "Modelling dengue epidemic spreading with human mobility," *Physica A: Statistical Mechanics & its Applications*, vol. 447, pp. 129–140, 2015.