Math. Model. Nat. Phenom.
Vol. 9, No. 2, 2014, pp. 89–107
DOI: 10.1051/mmnp/20149206

The Effect of Graph Structure on Epidemic Spread in a Class of Modified Cycle Graphs

A. Szabó-Solticzky, P.L. Simon *

Institute of Mathematics, Eötvös Loránd University, Budapest, Hungary and Numerical Analysis and Large Networks Research Group, Hungarian Academy of Sciences

Abstract. In this paper, an SIS (susceptible-infected-susceptible)-type epidemic propagation is studied on a special class of 3-regular graphs, called modified cycle graphs. The modified cycle graph is constructed from a cycle graph with N nodes by connecting node i to the node $i + d$ in a way that every node has exactly three links. Monte-Carlo simulations show that the propagation process depends on the value of d in a non-monotone way. A new theoretical model is developed to explain this phenomenon. This reveals a new relation between the spreading process and the average path length in the graph.

Keywords and phrases: SIS epidemic, theoretical approximation, network process

Mathematics Subject Classification: 05C82, 37N25, 60J28, 90B15

1. Introduction

Dynamical processes running on networks were widely studied in the last decade [1, 3, 5]. Among many others, a few examples are epidemic spread, such as the SIR (susceptible-infected-recovered) and SIS (susceptible-infected-susceptible) dynamics, spread of opinions through a population, propagation of neuronal activity on a neural network. Many models are studied in monograph [1], binary-state dynamics are investigated in [4]. The common feature of these models is that a graph with N nodes is given, and the nodes can be in one of M states. The transitions between these states are described by independent Poisson processes with corresponding transition rates determined by the state of the neighbouring nodes. One of the most important mathematical questions is to establish a relation between the network structure and the dynamical behaviour of the system [2, 6, 8]. A direct way to tackle this problem is to carry out Monte-Carlo simulations of the stochastic process to measure and determine quantities such as the average number of nodes in a given state (in the epidemic context, the average number of infected nodes) as a function of time. Simulations are relatively easy to carry out, however, these are not suitable to get analytical insight about the relation between the network structure and the dynamical behaviour of the system. Therefore several analytical approximating models were derived. For configuration random graphs mean-field and pair-wise approximation models have been derived as low-dimensional approximations, see e.g. [4, 5]. In these approximating models some structural properties of the network, such as

*Corresponding author. E-mail: simonp@cs.elte.hu

© EDP Sciences, 2014

Article published by EDP Sciences and available at http://www.mmnp-journal.org or http://dx.doi.org/10.1051/mmnp/20149206
https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

the average degree of the nodes, the degree correlation and the clustering of the network can be reflected by certain parameters. However, for finite, non-random networks having a special structure, the use of the above parameters is not sufficient to specify the graph to a sufficient degree of accuracy. Moreover, the applicability of mean-field and pair-wise approximations is restricted to random networks or to higher dimensional lattices. In this paper we present the results of our theoretical study in the case of an artificially constructed network with strong one dimensional character, for which two requirements hold. On one hand, the structure of the graph can be tuned by a single parameter, and on the other hand, the dynamic of the process can be determined analytically. In order to achieve this goal we introduce a class of networks, called modified cycle graphs, the structure of which can be tuned by a single parameter d and investigate how this parameter effects the SIS-type epidemic spread on the network. The subject of our study is the time dependence of the average number of infected nodes. As it turns out, this number increases in time with a rate depending in a non-trivial way on the value of d. We find that this rate is increasing as d is varied from $d = 2$ to a certain value, which is of order $\sqrt{N}$, and then it is decreasing as d is increased to $N/2$, see Figure 2 (N is the number of nodes). We derive an analytical approximate model that can explain this behaviour. It is remarkable that the average path length of the graph depends in the same way on d as the slope of the $I(t)$ function, where $I(t)$ is the expected number of infectious nodes at time t. This phenomenon will be briefly expanded on in the Discussion section and it will be the subject of future work.

The structure of the paper is as follows. In Section 2 we introduce the class of modified cycle graphs and present simulation results that show the time dependence of the epidemic spread for different values of the parameter d. In Section 3 a theoretical model is derived describing the epidemic propagation in a subgraph of size $2d$. Then the process in the whole graph is studied in Section 4. The paper concludes with some remarks in the Discussion section.

2. SIS epidemic on modified cycle graphs

Now we define a special class of graphs that we will call "modified cycle graphs". These graphs are regular, the degree of all nodes is three. A modified cycle graph is a non-random, structured counterpart of the well-known Watts-Strogatz graph, hence a motivation of our investigation is to derive analytical models of dynamical processes running on graphs having similar properties to the Watts-Strogatz model. In our model the average path length can be tuned by the parameter d and this parameter is involved in the analytical model. Hence a possible future goal may be to involve the average path length as a parameter in dynamical models. We note that cycle graphs with long-range links has already been studied by several authors. Saramäki and Kaski [9] investigated the effect of randomly varying long-range links on the spreading of an SIR epidemic. In our work the long-range links are fixed and positioned to a well defined rule. (Moreover, in our case the epidemic is of SIS-type.)

2.1. The construction of modified cycle graphs

Let us start from C_N, the undirected cycle graph with N nodes, where the vertices are numbered from 1 to N, namely the vertex set is $V = \{1, 2, \ldots, N\}$. Now we add more edges to the graph in a systematic way, along which the degree of each node becomes three. Let us take an integer $d \geq 2$ and connect first node 1 to node $1 + d$, then node 2 to node $2 + d$ and finally node d to node $2d$. At this stage nodes $1, 2, \ldots, 2d$ have degree three. Now repeat this process starting from node $2d + 1$. Generally, starting from the cycle graph C_N, we add a link between node i and j if and only if, $deg(i) = deg(j) = 2$ and $j = (i + d) \bmod N$. In order to complete this procedure at the end point of the circle it is necessary that N is an integer multiple of $2d$, that will be assumed from now on. For simplicity, we will mostly use the values $N = 2^k$ where k is a positive integer, and $d = 2, 4, \ldots, 2^{k-1}$ throughout the paper. Let $G_{N,d}$ denote the modified cycle graph obtained by the procedure above, with N nodes and parameter d. For $N = 16$ and $d = 2, 4, 8$ these graphs are shown in Figure 1.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

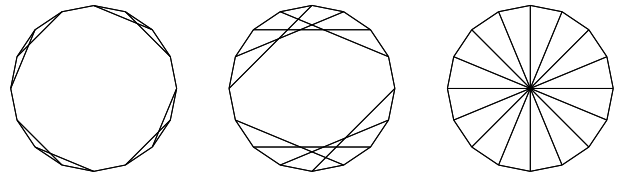


FIGURE 1. $G_{16,d}$ graphs for $d = 2, 4, 8$

2.2. Simulation results and their comparison to well-known ODE approximations

Let us present first the main phenomenon that can be observed as the value of d is changed. On a modified cycle graph an SIS type epidemic will be studied with infection rate τ and recovery rate γ. Both infection and recovery are modelled as independent Poisson processes. This means that in a short time interval δt, a susceptible individual with k infectious neighbours becomes infected with probability $1 - \exp(-k\tau\delta t)$, and an infectious individual recovers with probability $1 - \exp(-\gamma\delta t)$, independently of the state of its neighbours. This process can be investigated by individual based discrete-time stochastic Monte-Carlo simulation with synchronous updating. To be specific let us consider the graphs $G_{128,d}$ with $d = 2, 4, 8, 16, 32, 64$. The time dependence of the number of infected nodes obtained as the average of a suitably large number of simulations is shown in Figure 2. We can see that the slope of the curve increases as d is varied from $d = 2$ to $d = 16$ and then decreases as d is varied from $d = 16$ to $d = 64$. Let us compare first these simulation results to mean-field models obtained by pair and triple closure approximations. To formulate these models let us introduce $[I](t)$ as the expected number of infected nodes at time t, and similarly $[S](t)$ the expected number of susceptible nodes at time t. The following equations are exact, but not self-contained, since the expected number of ordered SI pairs is involved.

$$[\dot{S}] = \gamma[I] - \tau[SI] \tag{2.1}$$

$$[\dot{I}] = \tau[SI] - \gamma[I] \tag{2.2}$$

In order to get a self-contained system we have to approximate $[SI]$ in terms of $[S]$ and $[I]$. The simplest approximation is the pair closure that assumes statistical independence in the distribution of S and I nodes and takes the form $[SI] \approx [S][I]\frac{n}{N-1}$, where n is the average degree of a node. The above mentioned statistical independence is a good approximation when the infection reaches a stationary level, hence in the stationary state this yields a relatively good result. In our case $n = 3$, so we have the following mean-field model.

$$[\dot{S}] = \gamma[I] - \tau[S][I]\frac{3}{N-1} \tag{2.3}$$

$$[\dot{I}] = \tau[S][I]\frac{3}{N-1} - \gamma[I] \tag{2.4}$$

In fact, it is enough to use the second equation, since $[S] + [I] = N$.

Instead of closing equation $[\dot{I}] = \tau[SI] - \gamma[I]$ at the level of pairs, one can derive a differential equation for the expected value $[SI]$ and to $[II]$ that gives the following system of ODEs.

$$\begin{aligned}
[\dot{S}] &= \gamma[I] - \tau[SI] \\
[\dot{I}] &= \tau[SI] - \gamma[I] \\
[\dot{SI}] &= \gamma([II] - [SI]) + \tau\Big([SSI] - [ISI] - [SI]\Big) \\
[\dot{II}] &= -2\gamma[II] + 2\tau\Big([ISI] + [SI]\Big) \\
[\dot{SS}] &= 2\gamma[SI] - 2\tau[SSI]
\end{aligned}$$

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

We note that the derivation of this system starting from the master equations was given in [11]. Here $[AB]$ $(A, B \in \{S, I\})$ denotes the expected number of $A - B$ type pairs and $[ABC]$ $(A, B, C \in \{S, I\})$ denotes the expected number of edge pairs, if there is a link between $A - B$ and $B - C$. In order to get a self-contained system we close the equations at the level of triples by using the approximation

$$[ABC] \approx \frac{n-1}{n} \frac{[AB][BC]}{[B]},$$

that will be referred to as triple closure.

Now we compare the pair and triple closure approximations to simulation results. These approximations use only the average degree of the network that is the same for all values of d, hence obviously the d-dependence of the curves cannot be explained by these models. Moreover, the mean-field models assume the randomness of the network, therefore they are not expected to perform well for our graphs. Having these drawbacks in mind it is not surprising what Figure 2 shows for the graphs $G_{128,d}$ with $d = 2, 4, 8, 16, 32, 64$. The figure shows $I(t)$ curves obtained from simulation starting from $I(0) = 1$, that is there is only one infected node at the initial instant. We can see that neither the pair closure nor the triple closure approximation gives a good result.

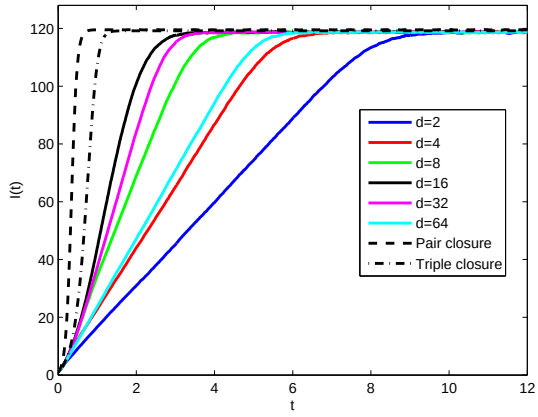


FIGURE 2. The value of $I(t)$ for an SIS epidemic on $G_{128,d}$ graphs for different values of d and the pair and triple closure approximations. The computations are started with initial condition $I(0) = 1$, the values of the parameters are $\tau = 5$, $\gamma = 1$.

If $I(0)$ is higher, e.g. the $I(0) = N/10$, i.e. 10% of the population is infected at the initial instant, then the curves getting closer to each other as it is shown in Figure 3.

We can see that the effect of graph structure on the dynamical process can only be observed with a low number of initially infected nodes, hence from now on we will consider only the case when $I(0) = 1$. Our aim is to derive an analytical model that exhibits the above observed property.

One way to compute the probability that node i is infected at time t is to solve the master equations that form a system of 2^N ordinary differential equations for the probabilities of all possible states of the graph. However, for typical graph sizes this system is too large for numerical computations. We note that using the automorphisms of the graph this system can be considerably reduced for graphs with special structure, see [10]. In our case, however, the automorphism group of the graph is not large enough to yield a useful reduction, hence we will derive a new analytic approximation that is tailored especially to the modified cycle graph.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

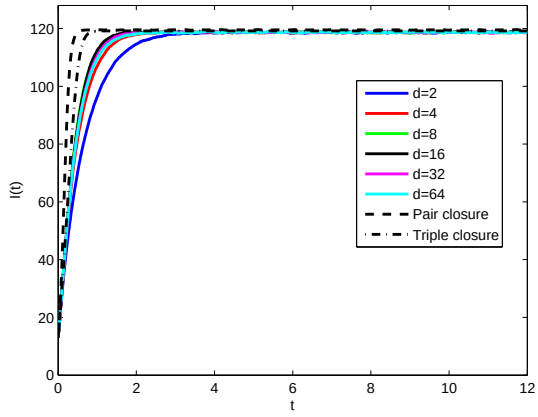


FIGURE 3. The value of $I(t)$ for an SIS epidemic on $G_{128,d}$ graphs for different values of d and the pair and triple closure approximations. The computations are started with initial condition $I(0) = N/10$, the values of the parameters are $\tau = 5$, $\gamma = 1$.

The key observation in our approximation is that our graph $G_{N,d}$ consists of $N/2d$ subgraphs of size $2d$ that are lying along a circle, see e.g. $G_{24,4}$ in Figure 4. More precisely, let $\widetilde{G_d}$ denote the subgraph that contains nodes from 1 to $2d$. In the next section we develop a model that describes the epidemic propagation in a subgraph. These results will be applied in Section 4 to model the process in the whole graph $G_{N,d}$.

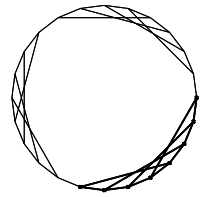


FIGURE 4. $G_{24,4}$ and a $\widetilde{G_4}$ subgraph denoted by bold lines and nodes

3. Modeling the epidemic process in a subgraph

Let us consider now the $\widetilde{G_d}$ subgraph defined above. We use the numbering $1, \ldots, 2d$ for the nodes and assume that the infection reaches the subgraph at node 1, that is at the initial instant node 1 is infected and the other nodes $2, \ldots, 2d$ are susceptible. Let us investigate now how the infection spreads in this $\widetilde{G_d}$ subgraph. Our first observation is that there is a shortest path for the infection to pass through the subgraph that is shown in Figure 5.

The second observation is that once the nodes shown in Figure 5 are infected then the infectious process runs in three separate arcs in the left, middle and right part of the subgraph as it is shown in Figure 6.

These arcs will be referred to infected arcs. We note that this approach works only in the case when d is large enough, since for small d values there is not enough space in the subgraph for these arcs to

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

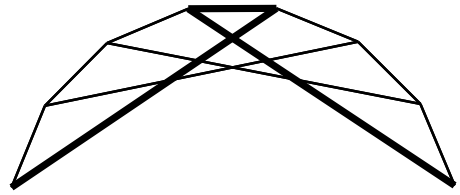


FIGURE 5. $\widetilde{G_8}$ subgraph and the shortest path along which the infection pass this subgraph.

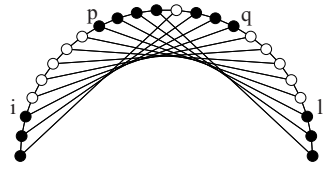


FIGURE 6. The infected arcs as the infectious process evolves in a $\widetilde{G_d}$ subgraph. Infected nodes are black, susceptible ones are white. The end points of the infected arcs are denoted by i, p, q, l.

establish. As the epidemic process evolves these arcs will finally merge and the whole subgraph will be infected. Hence we will divide the infectious process into two stages. In the first stage the infected arcs are separated from each other, and in the second stage the process reaches its steady state and the number of infected nodes becomes stationary. The two stages can be clearly seen in Figure 7, where the time dependence of the expected number of infected nodes is shown for $\widetilde{G_d}$ subgraphs with different d values. (The curves were obtained as the average of several simulations.) We can see that in the first stage the curves are linear and their slope do not depend on d. This fact is explained by the evolution of the infected arcs that obviously does not depend on the size of the subgraph. The size of the subgraph plays a role in determining the stationary value of the infection.

Therefore in approximating the curves shown in Figure 7 we use a linear function for the increasing part and a constant function for the stationary part. Hence let us introduce the approximating function $I_{std}(t)$ as the following piecewise linear function.

$$I_{std}(t) = \begin{cases} at & \text{if,} \quad 0 \le t \le t_{stat} \\ c & \text{if,} \quad t_{stat} \le t \end{cases}$$

where a and c are positive constants to be determined below and the time t_{stat} is determined simply by the intersection point of the two lines. The function $I_{std}(t)$ will be called the standard infection function. Thus in order to give an analytic approximation of the expected number of infected nodes in a $\widetilde{G_d}$ subgraph we have to give analytic formulas for the constants a and c. This will be carried out in the next two subsections.

3.1. Determination of the stationary level

In the second stage of the process the recovery of infected nodes and the reinfection of susceptible nodes are in steady state, yielding that $\tau[SI] = \gamma[I]$, where $[SI]$ is the average number of SI edges inside the subgraph. Since the recovery of infected nodes happens randomly we can assume that the distribution

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

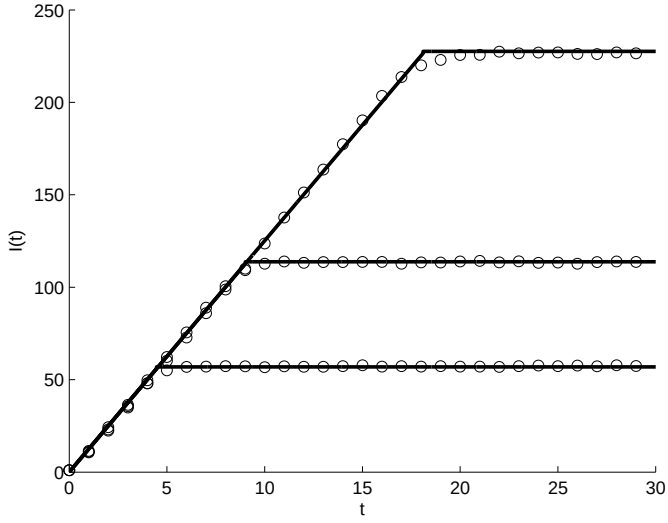


FIGURE 7. The expected number of infected nodes in $\widetilde{G_d}$ as a function of time, for $d = 32, 64, 128$. The parameter values are $\tau = 5$, $\gamma = 1$. The circles represent simulation results, the continuous lines are given by the theoretical approximation $I_{std}(t)$.

of infected and susceptible nodes is statistically independent, hence $[SI]$ can be approximated in terms of $[I]$ as $[SI] = 3[I][S]/2d$. This formula is explained by the simple fact that an S node has 3 neighbours and the proportion of infected nodes among its neighbours is the same as the proportion of infected nodes in the whole graph, which is $[I]/2d$. Then equation $\tau[SI] = \gamma[I]$ yields

$$[I](t) = 2d\left(1 - \frac{\gamma}{3\tau}\right),$$

that is the proportion of infected nodes in the steady state is $\alpha = 1 - \frac{\gamma}{3\tau}$. Thus the value of c in the formula of $I_{std}(t)$ is

$$c = 2d\alpha = 2d\left(1 - \frac{\gamma}{3\tau}\right).$$

We note that in the case $3\tau < \gamma$ the epidemic process will stop with all nodes becoming susceptible, i.e. the epidemic will not spread, hence only the case $3\tau > \gamma$ is studied here.

3.2. Determination of the slope in the increasing stage

As it can be seen in Figure 6, in the increasing stage of the process there are three infected arcs separated by susceptible arcs. We note that this holds only in the case $8 \leq d < \frac{N}{2}$. For $d = \frac{N}{2}$ there are only two infected arcs (the left and right arcs are merged), but this case can be treated together with the case $8 \leq d < \frac{N}{2}$. The cases $d = 2$ and $d = 4$ will be studied separately. First, we have to note that inside the infected arcs some nodes become susceptible randomly, and then they are reinfected again. Following our approach developed in [7], the process will be modeled statistically as follows. The recovery of infected nodes and the reinfection of susceptible nodes is considered to be in steady state inside the infected arc, hence we assume that the expected number of infected nodes is proportional to the length of the infected arc, that is there is a constant α, such that

$$[I](t) = \alpha m(t) \tag{3.1}$$

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

where $m(t)$ is the total length of the infected arcs. Because of the above steady state assumption inside the infected arcs the value of α is the same as in the stationary state, i.e. $\alpha = 1 - \frac{\gamma}{3\tau}$, as it was determined above.

The slope to be determined is $[\dot{I}](t) = \alpha \dot{m}(t)$. Therefore we have to estimate how the total length of the infected arcs changes in time. The length of the arcs changes at their four end points by infection and by recovery, depending on the number of SI type edges starting from these end points. Hence we can approximate $\dot{m}(t)$ as

$$\dot{m}(t) = e_{SI}\tau - 4\gamma, \tag{3.2}$$

where e_{SI} denotes the average total number of SI edges, the I nodes of which are at the four end points or the S nodes of which are neighbours of the four end points. The -4γ term is simply explained by the fact that the total length of the arcs can decrease if one of the four end points recovers. Thus it remains to determine (in fact, to approximate) e_{SI}.

This number depends on the mutual position of the infected arcs. In order to determine e_{SI} let us denote the end points of the arcs by i, p, q, l. (The nodes of the subgraph are numbered as $1, 2, \ldots, 2d$.) Numerical simulation and simple probabilistic arguments show that q is close to $i + d$ and l is close to $p + d$. This can be explained as follows. Node i is connected to node $i + d$, hence once node i is infected then it may infect node $i + d$, and similarly for p and $p + d$. The number of SI edges depends on the mutual position of i, q and that of p, l. In our calculation the following cases will be taken into account, see Figure 8, where the notations $k = i + d$ and $j = l - d$ are also used.

A $q = i + d$, $l = p + d$,
B $q = i + d + 1$, $l = p + d$,
C $q = i + d + 1$, $l = p + d + 1$,
D $q = i + d + 2$, $l = p + d$,
E $q = i + d + 2$, $l = p + d + 1$,
F $q = i + d + 2$, $l = p + d + 2$.

We note that the cases like $q = i + d + 1$, $l = p + d + 2$ are omitted by symmetry reasons (it can be identified with the case $q = i + d + 2$, $l = p + d + 1$), and further cases like $q = i + d + 3$, $l = p + d$ are neglected because they have pure chance, i.e. occur rarely according to simulations. The number of SI edges can be easily determined in each case as shown in Table 1 by using Figure 8.

	A	B	C	D	E	F
Number of SI edges	4	5	6	6	7	8

TABLE 1. Number of SI type edges at different configurations.

In order to determine the average number of SI edges we have to compute the weighted average of the above number of SI type edges with the probabilities of the configurations as weights. We assume that as the process evolves the probabilities of these configurations are in equilibrium, in the sense that the system jumps from one of these to another and this process is in steady state. Therefore let us consider these configurations as the states of a discrete time Markov chain and let us determine the transition probabilities by considering only the infection (the recovery is taken into account separately by using the -4γ term). From configuration A the system jumps to configuration B with probability 1, since one of the end points infect a neighbour. In configuration B there are five infecting SI edges. If node i infects node $i + 1$ or node $k + 1$ infects node $i + 1$, then the system jumps to configuration A, hence this transition happens with probability 2/5 (two out of five cases). If node j infects node $j - 1$ or node l infects node $l - 1$, then the system jumps to configuration C, hence this transition happens with probability 2/5 (two out of five cases). Finally, if node $k + 1$ infects node $k + 2$, then the system jumps to

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

configuration D, hence this transition happens with probability 1/5 (one out of five cases). This way we get the transition probabilities from configuration B as they are shown in Figure 9. In a similar way one can easily determine the transition probabilities from other configurations. It is important to note that when considering configurations D, E and F, then there may happen transitions that take the system to new configurations that are not listed in our state space. For example, from configuration D the system can go to configuration $q = i+d+3$, $l = p+d$ with probability 1/6. Then instead of introducing the new configuration $q = i + d + 3$, $l = p + d$ we increase the probability of transition $D \to E$ by 1/6, because the configuration $q = i + d + 3$, $l = p + d$ is similar to configuration E in the sense that in both of them there are 7 SI edges.

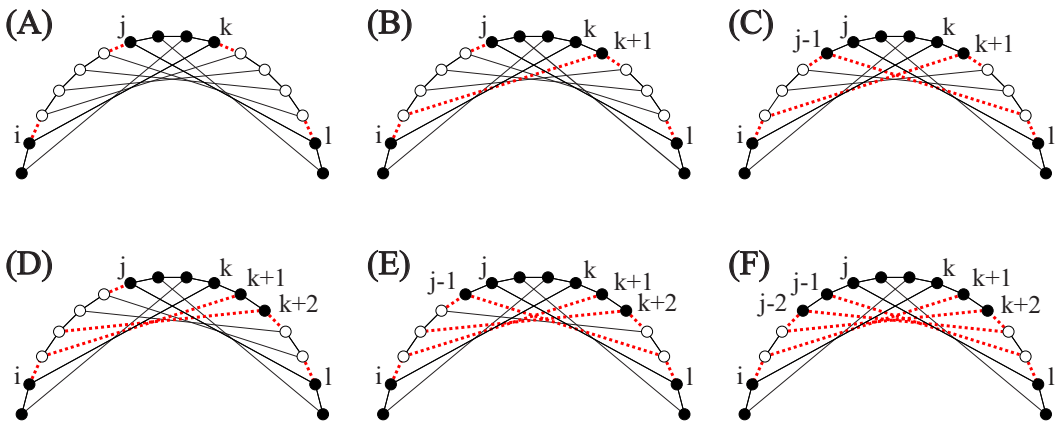


FIGURE 8. Evolution of infected arcs. The six different mutual positions of the arcs. The red dotted lines are the SI edges.

Computing the transition probabilities in the above specified way we get the following transition matrix.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{2}{5} & 0 & \frac{2}{5} & \frac{1}{5} & 0 & 0 \\ 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{7} & \frac{2}{7} & 0 & \frac{3}{7} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

The probabilities of the different configurations in the steady state are given by the left eigenvector of this matrix corresponding to the eigenvalue 1, i.e. by the eigenvector x satisfying $xA = x$. Solving this system for x (under the assumption that $\sum x_i = 1$) we get the average number of SI edges as the weighted average

$$e_{SI} = 4x_1 + 5x_2 + 6x_3 + 6x_4 + 7x_5 + 8x_6 = 6.0256.$$

Thus based on equations (3.1) and (3.2) we get that the slope of the $I(t)$ function in the first stage is

$$(6.025\tau - 4\gamma)\left(1 - \frac{\gamma}{3\tau}\right).$$

Hence for $8 \le d \le \frac{N}{2}$ we have a formula for standard infection function namely,

$$I_{std}(t) = \begin{cases} (6.0256\tau - 4\gamma)\left(1 - \frac{\gamma}{3\tau}\right) t & \text{if,} \quad 0 \le t \le t_{stat} \\ 2d\left(1 - \frac{\gamma}{3\tau}\right) & \text{if,} \quad t_{stat} \le t \end{cases}$$

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

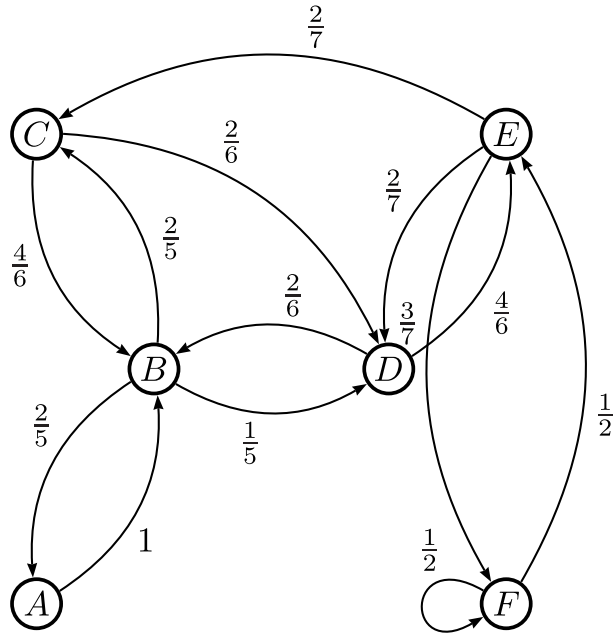


FIGURE 9. Transition probabilities of the Markov chain describing the evolution of infected arcs.

where t_{stat} is given by the intersection of the two lines.

Finally, let us consider the case when d is small. Since we are mainly interested in the cases when d is a power of 2, we consider the cases $d = 2$ and $d = 4$. In these cases the above approach using the infected arcs cannot be applied because there isn't enough space between the separated infected sets. We do not develop the theory in these cases, instead we determine the parameters from simulations. We found that the $I(t)$ curves can be approximated by a piecewise linear function. The steady state can be approximated in the same way as for large d values. For the slope in the first stage we got from simulation $a\tau - \gamma$ with $a = 1.5$ if $d = 2$ and $a = 2.5$ if $d = 4$. Thus in the cases $d = 2$ and $d = 4$ we have the following formula for standard function.

$$I_{std}(t) = \begin{cases} (a\tau - \gamma)t & \text{if,} \quad 0 \le t \le t_{stat} \\ 2d\left(1 - \frac{\gamma}{3\tau}\right) & \text{if,} \quad t_{stat} \le t \end{cases},$$

where t_{stat} is given by the intersection of the two lines and $a = 1.5$ if $d = 2$ and $a = 2.5$ if $d = 4$.

We can see in Figure 7 that our theoretical approximation $I_{std}(t)$ yields a good approximation of the simulation results, that are shown by circles in that Figure.

3.3. Analytic approximation of the passing time

In order to model the epidemic spread in the whole graph we will need the passing time that is defined as the average time needed for the infection to spread through a $2d$ size subgraph and infect the next subgraph's terminal node. Let R denote this passing time. In other words, suppose that in the subgraph $\widetilde{G_d}$ the terminal node 1 is infected and the other nodes are susceptible. Then R denotes the average time that is needed for node $1 + 2d$ (the terminal node of the next subgraph) to be infected. That is R is the time that the infection must spend in a subgraph before reaches the next subgraph. Our aim now is to approximate the passing time R.

As it was shown in Figure 5 there is a shortest path, along which the infection passes the subgraph. Hence the passing time can be approximated by the time that is needed to pass a line graph. Let us

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

denote by L_m the line graph containing m nodes, i.e. node i is connected to nodes $i-1$ and $i+1$, if $1 < i < m$ and the terminal node 1 is connected to node to 2 and node m is connected to node $m-1$. Let us denote by T_d the passing time for a subgraph $\widetilde{G_d}$. Similarly $T_{L,4}$ and $T_{L,5}$ denotes the passing times for the line graphs L_4 and L_5. The following table shows simulation results for passing times of different graphs and for different values of τ.

	T_2	T_4	T_8	T_{32}	T_{64}	T_{128}	$T_{L,4}$	$T_{L,5}$
$\tau = 5, \quad \gamma = 1$	0.47	0.65	0.81	0.84	0.84	0.84	0.71	0.93
$\tau = 7, \quad \gamma = 1$	0.34	0.45	0.55	0.58	0.58	0.57	0.45	0.65
$\tau = 10, \quad \gamma = 1$	0.23	0.32	0.37	0.39	0.39	0.39	0.32	0.43

TABLE 2. Passing times for different graphs obtained as the average of 10000 simulations.

We can see that for $8 \leq d$ the passing times are approximately the same for every subgraph and

$$T_{L,4} < T_8 \approx T_{32} \approx T_{64} \approx T_{128} < T_{L,5}$$

This confirms the idea that the infection passes a subgraph along the path shown in Figure 5, hence the passing time does not depend on d (until it is not too small). Moreover, one can observe that

$$T_8 \approx T_{32} \approx T_{64} \approx T_{128} \approx \frac{T_{L,4} + T_{L,5}}{2}$$

and

$$T_4 \approx T_{L,4}.$$

Therefore if we are able to compute the passing times T_2, $T_{L,4}$ and $T_{L,5}$, then we can approximate the others too. The main advantage of this simplification is that the graphs $\widetilde{G_2}$, L_4 and L_5 are small enough to calculate the average passing times theoretically without any approximation.

Now we show how to compute the passing time for the line graph L_4. Assume that at the initial instant node 1 is infected and assume that this node cannot recover. (If we allow this node to recover, then it will be reinfected by the terminal node of the preceding subgraph, however, we do not want to extend our model with another node.) Thus we start our process from state $ISSS$. For simplicity, we explain the process through a discrete time approximation. After time dt the second node will be infected with probability τdt, hence the system will be in state $IISS$. From this state the system can go to state $IIIS$ (infection) with probability τdt or go back to state $ISSS$ (recovery) with probability γdt. From state $IIIS$ the system can go to state $IIII$ (infection) with probability τdt or go back to state $IISS$ (recovery) with probability γdt or go to state $ISIS$ (recovery) with probability γdt. Finally, from state $ISIS$ the system can go to state $ISII$ (infection) with probability τdt or go back to state $IIIS$ (infection) with probability $2\tau dt$ or go to state $ISSS$ (recovery) with probability γdt. The states and the transition probabilities are shown in Figure 10. The continuous time Markov chain has the same states, its relation to the discrete time process is dealt with in the Appendix.

Our goal is to determine the average time that the infection needs to pass the graph L_4, i.e. the average time that is needed to arrive at states $IIII$ or to $ISII$ starting from state $ISSS$. This time is called the mean first passage time. It can be obtained by solving the master equations of the continuous time Markov chain, as it is shown in the Appendix. However, as it is explained also there, it is significantly simpler to obtain the mean first passage time from a discrete time approximation of the Markov chain. That is let us assume first that the process evolves in discrete time steps of length Δt. Let us denote by N_{ISSS} the average number of steps needed to infect the last node, i.e. to arrive at states $IIII$ or to $ISII$ starting from state $ISSS$, and introduce similarly N_{IISS}, N_{IIIS} and N_{ISIS} for the average number of steps needed to arrive at states $IIII$ or to $ISII$ starting from the different states. The advantage of

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

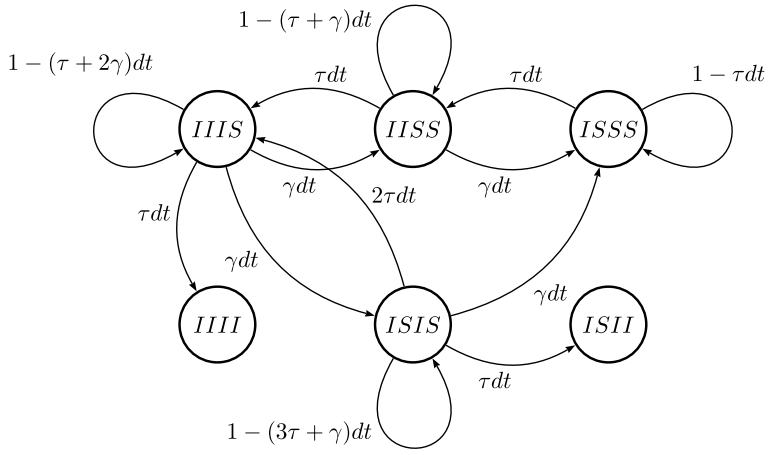


FIGURE 10. The state space for the line graph L_4 and the transition probabilities.

using the discrete time counterpart is that a linear system of equations can be derived for the variables N_{ISSS}, N_{IISS}, N_{IIIS} and N_{ISIS} directly from the transition probabilities without solving the master equations. Based on Theorem 3.2.2. in [12] this linear system is the following.

$$\begin{aligned}
N_{ISSS} &= 1 + N_{IISS}\tau\Delta t + N_{ISSS}(1-\tau\Delta t)\\
N_{IISS} &= 1 + N_{IIIS}\tau\Delta t + N_{ISSS}\gamma\Delta t + N_{IISS}(1-(\tau+\gamma)\Delta t)\\
N_{IIIS} &= 1 + (N_{IISS}+N_{ISIS})\gamma\Delta t + N_{IIIS}(1-(\tau+2\gamma)\Delta t)\\
N_{ISIS} &= 1 + 2N_{IIIS}\tau\Delta t + N_{ISSS}\gamma\Delta t + N_{ISIS}(1-(3\tau+\gamma)\Delta t)
\end{aligned}$$

The derivation of the equations can heuristically be explained as follows. Let us consider first the state $ISSS$. From this state we can move to state $IISS$ with probability $\tau\Delta t$ or we can stay in that state with probability $1-\tau\Delta t$. Hence after one time step applying the law of total probability we get the first equation. The other equations can be derived similarly based on the transitions given in Figure 10.

Let us return now to the continuous time process and introduce the variables $t_1 = N_{ISSS}\Delta t$, $t_2 = N_{IISS}\Delta t$, $t_3 = N_{IIIS}\Delta t$, $t_4 = N_{ISIS}\Delta t$. Then t_1 is the average time (in the continuous time case) that the infection needs to pass the graph L_4 that we are looking for, i.e. the mean first passage time in the continuous time system. For these variables we get the system

$$\begin{pmatrix} \tau & -\tau & 0 & 0\\ -\gamma & \tau+\gamma & -\tau & 0\\ 0 & -\gamma & \tau+2\gamma & -\gamma\\ -\gamma & 0 & -2\tau & 3\tau+\gamma \end{pmatrix} \begin{pmatrix} t_1\\ t_2\\ t_3\\ t_4 \end{pmatrix} = \begin{pmatrix} 1\\ 1\\ 1\\ 1 \end{pmatrix}$$

Solving this system we get the passing time for the graph L_4

$$R_4 := t_1 = \frac{(9\tau^3 + 12\gamma\tau^2 + 8\gamma^2\tau + 2\gamma^3)}{\tau^3(3\tau+2\gamma)}.$$

This formula gives exactly the same values as those in the 7th column of the above table obtained from simulation. In the Appendix it is shown in the simpler case of the line graph L_3 that this discrete time counterpart gives the same formula for the mean first passage time as that obtained by solving the master equation of the continuous time system, moreover, the formula obtained by the discrete time system is independent from the choice of Δt.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

This approach can be applied to the line graph L_5 in a similar way, however, now there are more states in the state spaces and the system of equations consists of 8 equations. Solving this system for N_{ISSSS} we get the passing time for the line graph L_5 as $R_5 = N_{ISSSS}\Delta t$, hence

$$R_5 = \frac{216\tau^7 + 774\gamma\tau^6 + 1313\gamma^2\tau^5 + 1422\gamma^3\tau^4 + 1071\gamma^4\tau^3 + 552\gamma^5\tau^2 + 172\gamma^6\tau + 24\gamma^7}{\tau^4(54\tau^4 + 153\gamma\tau^3 + 167\tau^2\gamma^2 + 80\tau\gamma^3 + 14\gamma^4)}.$$

Finally, let us consider the case of the graph $\widetilde{G_2}$. (We note that the method can be applied in principle for any graph, but the size of the system is increasing exponentially with size of the graph.) We number the nodes of $\widetilde{G_2}$ by $1, 2, 3, 4$ and we join to node 4 an extra node 5, since the passing time is defined as the time when the infection reaches the first node of the next subgraph. The extra node 5 represents the first node of the next subgraph. We start from state $ISSSS$ and calculate the time when node 5 becomes infected. Similarly to the above cases, we introduce N_{IABCD} as the average number of steps that are needed starting from state $IABCD$ until node 5 become infected. At the beginning we are in state (I, S, S, S, S) and we could step to (I, I, S, S, S) or (I, S, I, S, S) with probability $\tau\Delta t$, so we get

$$N_{ISSSS} = 1 + \tau\Delta t(N_{IISSS} + N_{ISISS}) + (1 - 2\tau\Delta t)N_{ISSSS}$$

where $(N_{IISSS} + 1 + N_{ISISS} + 1)$ means we have to take one step to (I, I, S, S, S) plus the time N_{IISSS} (and similarly for the state (I, S, I, S, S)). The term $(N_{ISSSS} + 1)$ represents the situation, when we are stuck in the state (I, S, S, S, S) for one step. If we are in (I, I, S, S, S) we can go to (I, I, I, S, S) with probability $2\tau\Delta t$ because the third node has two infected neighbours, or we can go back to (I, S, S, S, S) with probability $\gamma\Delta t$ or to (I, I, S, I, S) with probability $\tau\Delta t$, so the equation takes the form

$$N_{IISSS} = 1 + \tau\Delta t(2N_{IIISS} + N_{IISIS}) + (1 - (3\tau + \gamma)\Delta t)N_{IISSS} + \gamma\Delta t N_{ISSSS}.$$

Similarly, we can formulate the equations for all states, altogether there will be 8 equations. Solving that system we get N_{ISSSS}, the average number of steps needed to infect the fifth node. Then multiplying this with Δt we get the passing time for the graph $\widetilde{G_2}$ as

$$R_2 := N_{ISSSS}\Delta t = \frac{(260\tau^5 + 420\gamma\tau^4 + 316\gamma^2\tau^3 + 153\tau^2\gamma^3 + 6\gamma^5 + 45\tau\gamma^4)}{6\tau^3(20\tau^3 + 25\tau^2\gamma + 12\tau\gamma^2 + 2\gamma^3)}.$$

Summarizing, we have the following values for the passing times. Using the observations made based on Table 2 in the case $d = 2$ the passing time is R_2, in the case $d = 4$ the passing time is R_4 and in the case $d \geq 8$ the passing time is $(R_4 + R_5)/2$.

4. Theoretical model for the epidemic propagation in the whole modified cycle graph

The aim now is to derive a new approximating model that involves the parameter d and exhibits the non-monotone dependence of the slope on d. As it was mentioned above, our graph $G_{N,d}$ consists of $N/2d$ subgraphs of size $2d$ that are lying along a circle. More precisely, let $\widetilde{G_d}$ denote the subgraph that contains nodes from 1 to $2d$. The whole graph $G_{N,d}$ consists of $M = \frac{N}{2d}$ copies of $\widetilde{G_d}$ that are situated at nodes $\{2d(k-1)+1, \ldots, 2dk\}$ for $k = 1, 2, \ldots, M$. These subgraphs will be denoted by $\widetilde{G_d^k}$.

Let $I_k(t)$ denote the expected number of infected nodes in the k-th subgraph for $k = 1, 2, \ldots, M$. Then the total number of infected nodes in the whole graph is

$$I(t) = \sum_{k=1}^{M} I_k(t). \tag{4.1}$$

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

Since all the subgraphs are the same, the $I_k(t)$ functions differ only in a shift of time, that is

$$I_k(t) = I_{std}(t - t_k),$$

where the function $I_{std}(t)$ gives the number of infected nodes in a subgraph after time t measured from the instant when the infection reached the subgraph, and t_k is the time when the infection reaches the k-th subgraph. Hence our derivation consists of the following two steps.

– Describing the infection in a subgraph, i.e. deriving a formula for $I_{std}(t)$.
– Determining the time t_k when the infection reaches the k-th subgraph.

The standard infection function $I_{std}(t)$ was determined in Subsection 3.2. The reaching time t_k will be computed based on the passing time as follows. Let us determine the passing time through a subgraph $\widetilde{G_d^k}$ as it was described in the previous subsection and denote it simply by R. The modified cyclic graph is a ring of $\widetilde{G_d^k}$ subgraphs. If the infection starts inside a subgraph, then it reaches one of the two end points of the subgraph in a relatively short time and then the infection starts to spread in the next subgraph concurrently. Hence for sake of simplicity, we assume that the first infected node is the end point of the subgraph, let this node be node 1. Hence the epidemic starts to spread in the subgraphs $\widetilde{G_d^1}$ and $\widetilde{G_d^M}$. Then it reaches the subgraphs $\widetilde{G_d^2}$ and $\widetilde{G_d^{M-1}}$ and it spreads in four subgraphs concurrently, see Figure 11. The process then evolves in a similar way and after some time it reaches its steady state in the subgraphs $\widetilde{G_d^1}$ and $\widetilde{G_d^M}$ and in the other subgraphs as well. The spread of the epidemic is symmetric in the left and right part of the graph, hence we will consider the process in subgraphs numbered $1, 2, \ldots, M/2$ (we assume for simplicity that the number of subgraphs, M is an even number). Thus instead of equation (4.1) we write

$$I(t) = 2 \sum_{k=1}^{M/2} I_k(t), \tag{4.2}$$

where $I_k(t)$ denotes the number of infected nodes in the k-th subgraph. Assuming that initially node 1 is infected we have that the reaching time of the k-th subgraph is $t_k = R(k-1)$, because the infection has to pass $k-1$ subgraphs before arriving to the k-th subgraph. Therefore the number of infected nodes in the k-th subgraph can be given as

$$I_k(t) = H(t - R(k-1))I_{std}(t - (k-1)R),$$

where H is the Heaviside function whose value is zero for negative arguments and one for positive arguments. This expression means that I_k is the same as the standard infection function I_{std} when $t > t_k = R(k-1)$, i.e. when the infection reaches the k-th subgraph. Substituting this expression into (4.2) we get the total number of infected nodes as

$$I(t) = 2 \sum_{k=1}^{M/2} H(t - R(k-1))I_{std}(t - (k-1)R). \tag{4.3}$$

For given values of N, d, τ and γ this formula gives an analytical approximation of $I(t)$. As a conclusion to this theoretical section we can summarize how to apply this formula for a given set of the parameters. First, the function $I_{std}(t)$ has to be determined as it is given in Section 3.2. Then the passing time should be computed as it is explained in the last paragraph of Section 3.3. Namely, in the case $d = 2$ the passing time is R_2, in the case $d = 4$ the passing time is R_4 and in the case $d \geq 8$ the passing time is $(R_4 + R_5)/2$ (these numbers are defined in Section 3.3). Finally, $I_{std}(t)$ and R are substituted into (4.3) to get the theoretical value of $I(t)$. The results are shown in Figure 12 for a graph with $N = 1024$ nodes and for different values of d.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

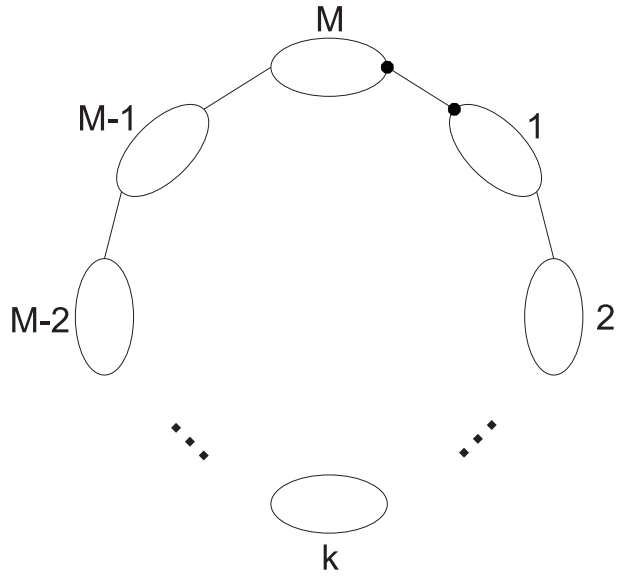


FIGURE 11. The modified cycle graph is a ring of $\widetilde{G_d^k}$ subgraphs. The infection starts at node 1, hence first the epidemic spreads in the subgraphs $\widetilde{G_d^1}$ and $\widetilde{G_d^M}$ and then in $\widetilde{G_d^2}$ and in $\widetilde{G_d^{M-1}}$.

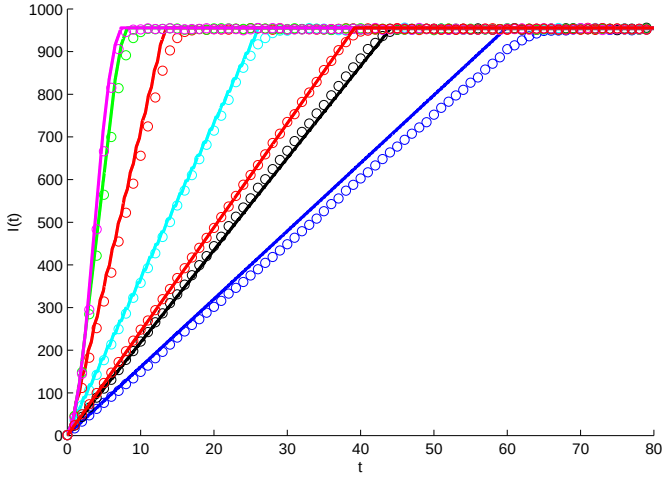


FIGURE 12. The $I(t)$ curves for a graph with $N = 1024$ nodes and for $d = 2, 4, 8, 16, 32, 64, 512$, $\tau = 5$, $\gamma = 1$. The continuous curves are obtained by the theoretical formula (4.3), the circles correspond to the average of 1000 simulations (the value of d is identified by the colors).

5. Discussion

In the paper we introduced a class of regular networks, called modified cycle graphs, the structure of which can be tuned by a single parameter d, and it was investigated how this parameter effects the

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

SIS-type epidemic spread on the network. First we showed by simulation that the number of infected nodes increases in time at a rate depending in a non-monotone way on the value of d. We found that this rate is increasing as d is varied from $d = 2$ to a certain value, which is of order $\sqrt{N}$, and then it is decreasing as d is increased to $N/2$, see Figure 12. We derived an analytical approximating model that can explain this behaviour. The construction of the theoretical model is based on the fact that our graph $G_{N,d}$ consists of $N/2d$ subgraphs of size $2d$ that are lying along a circle, as it was shown in Figure 4. Then the derivation was based on the following two observations. First, we note that a shortest path for the infection's passage through a subgraph exists as shown in Figure 5. The second observation is that once the nodes shown in Figure 5 are infected then the infectious process runs in three separate arcs along the cycle, in the left, middle and right part of the subgraph as it is shown in Figure 6. Based on these ideas, the theoretical model yields an explicit formula for the number of infected nodes, this is given in (4.3).

The theoretical prevalence curves (i.e. $I(t)$) are compared to simulation results for a graph with $N = 1024$ nodes and for different values of d in Figure 12. We can see that our theoretical model gives a good approximation for different d values. The model performs well for other values of the parameters τ and γ as well, until the the value of τ becomes too small to maintain a reasonable epidemic. We note that the performance of the theory for $d = 2$ could be amended if a more accurate value of the parameter a in the formula of $I_{std}(t)$ was derived. This could be achieved by extending the Markov-chain approach, presented in Section 3.2, to the case of $d = 2$.

It is remarkable that the average path length of the graph depends in the same way on d as the slope of the $I(t)$ function. In Table 3 the value of the average path length is given for several values of d for a graph having $N = 1024$ nodes. We can see that the average path length is decreasing with d when $2 \leq d \leq 32$ and then it is increasing for $32 \leq d \leq 512$. The slope of the $I(t)$ curve is also given in the table. It is increasing when $2 \leq d \leq 64$ and it is decreasing for $64 \leq d \leq 512$. The similarity in their behaviour can be explained qualitatively by the simple observation that smaller average path lengths lead to faster epidemics which in relatively few generations percolate through a large part of the graph. The quantitative relation between the slope of the prevalence curve and the average path length merits further attention, and will be extended to graphs with more complex but regular structure in future work.

d	2	4	8	16	32	64	128	256	512
slope	14,9	22.3	36.5	66.5	112.3	135.5	89.8	48.6	23.8
avg path length	192.2	128.6	66.8	38.8	30.5	37.1	59.1	97.3	128.6

TABLE 3. The slope of the $I(t)$ curve and the average path length of the graph for different values of d, for $\tau = 5$ and $\gamma = 1$.

Acknowledgements. Péter L. Simon acknowledges support from OTKA (grant no. 81403).

Appendix

The goal of this Appendix is twofold. On one hand we show how the mean first passage time can be computed in a continuous time Markov chain, on the other hand we show that the same formula for this time can be derived from the discrete time counter part, and also show that this latter computation is much shorter and simpler.

Consider a continuous time Markov chain with three states: ISS, IIS and III and with transition rates given in Figure 5. This models the infection in a line graph L_3 with three nodes, where the first node is fixed to be infected and the third node cannot recover. Our aim is to determine the mean first passage time, i.e. the average time that is needed for the infection to reach the third node starting from the first one. (That is the reason why the third node is assumed not to recover, since once it is infected we can stop

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

the process.) First, we will show how the mean first passage time can be computed from the continuous time model, then it will be shown that the same formula for this time can be derived from a discrete time process that is basically the time Δt discretisation of the continuous time one. It is remarkable that the same formula will be obtained independently from the choice of the discretisation time Δt. The same computation can be carried out for line graphs with more nodes, however, the computation is more tedious then, this is why we chose this simple model to illustrate our claim.

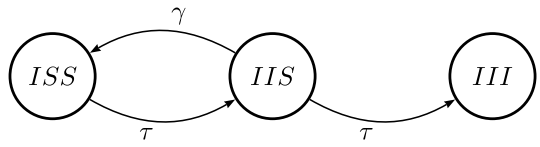


FIGURE 13. The state space for the line graph L_3 in continuous time and the transition probabilities. III is an absorbing state

5.1. Continuous time system

Let us denote by $x_1(t)$ the probability that the system is in state ISS at time t. Similarly, $x_2(t)$ and $x_3(t)$ are the probabilities of the states IIS and III. The master equations for these probabilities are

$$\begin{aligned}\dot{x_1} &= \gamma x_2 - \tau x_1\\ \dot{x_2} &= \tau x_1 - (\tau+\gamma)x_2\\ \dot{x_3} &= \tau x_2\end{aligned}$$

This system can be written in the form $\dot{x} = Ax$ with

$$A = \begin{pmatrix} -\tau & \gamma & 0\\ \tau & -\tau-\gamma & 0\\ 0 & \tau & 0\end{pmatrix}.$$

The eigenvalues of A are $\lambda_1, \lambda_2, 0$, where $\lambda_i^2 + \lambda_i(2\tau+\gamma) + \tau^2 = 0$. Computing the eigenvectors of A and using the initial condition $x_1(0) = 1, x_2(0) = 0, x_3(0) = 0$ (expressing that we start the process from state ISS), we get the solution of the system. Its third coordinate is

$$x_3(t) = \frac{\lambda_2}{\lambda_1-\lambda_2}e^{\lambda_1 t} + \frac{\lambda_1}{\lambda_2-\lambda_1}e^{\lambda_2 t} + 1.$$

The transition probability $p_{13}(t)$, that is the probability that system, starting from state 1, is in state 3 at time t is $p_{13}(t) = x_3(t)$. It is known that the mean first passage time is (see page 170 in [12])

$$T_{13} = \int_0^\infty 1 - p_{13}(t)dt.$$

This yields

$$T_{13} = \int_0^\infty 1 - x_3(t)dt = \frac{2\tau+\gamma}{\tau^2}.$$

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

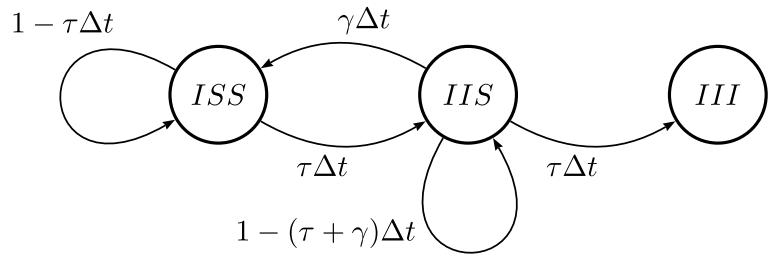


FIGURE 14. The state space for the line graph L_3 in discrete time and the transition probabilities. III is an absorbing state

5.2. Discrete time system

Fix now a positive number Δt and introduce the discrete time system by $X_i(n) = x_i(n\Delta t)$. The states of the discrete time Markov chain are the same as those of the continuous time one and the transition probabilities are given in Figure 5.2.

The master equations can be written as follows.

$$\begin{aligned} X_1(n+1) &= \gamma\Delta t X_2(n) + (1-\tau\Delta t)X_1(n) \\ X_2(n+1) &= \tau\Delta t X_1(n) + \Big(1-(\tau+\gamma)\Delta t\Big)X_2(n) \\ X_3(n+1) &= \tau\Delta t X_2(n) \end{aligned}$$

The mean first passage time can be determined again by solving the master equations. However, in the discrete time case a system of linear equations can be directly given for the mean first passage times N_{13} and N_{23} that denote the average number of steps that are needed to arrive at state 3 starting from state 1 and from state 2. This linear system is given in Theorem 3.2.2. in [12] as follows

$$\begin{aligned} N_{13} &= 1 + N_{23}\tau\Delta t + N_{13}(1-\tau\Delta t) \\ N_{23} &= 1 + N_{13}\gamma\Delta t + N_{23}\Big(1-(\tau+\gamma)\Delta t\Big) \end{aligned}$$

Eliminating N_{23} from the linear system we get

$$\tau\Delta t N_{13} = \frac{\tau}{\tau+\gamma}(N_{13}\gamma\Delta t + 1) + 1$$

yielding

$$N_{13} = \frac{2\tau+\gamma}{\tau^2\Delta t}.$$

Hence by using the relation $t = n\Delta t$, the mean first passage time in the continuous time system is

$$N_{13}\Delta t = \frac{2\tau+\gamma}{\tau^2}$$

that is the same as T_{13} independently from the discretisation time Δt.

References

[1] A. Barrat, M. Barthelemy, A. Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge, 2008.

[2] M. Boguna, R. Pastor–Satorras. *Epidemic spreading in correlated complex networks*. Phys. Rev. E., 66 (2002), 47104.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press

[3] L. Danon, A.P. Ford, T. House, C.P. Jewell, M.J. Keeling, G.O. Roberts, J.V. Ross, M.C. Vernon. *Networks and the Epidemiology of Infectious Disease, Interdisciplinary Perspectives on Infectious Diseases*. 2011:284909 special issue "Network Perspectives on Infectious Disease Dynamics".

[4] J.P. Gleeson. *High-accuracy approximation of binary-state dynamics on networks*. Phys. Rev. Letters, 107 (2011), 068701.

[5] T. House, M. J. Keeling. *Insights from unifying modern approximations to infections on networks*. J. Roy. Soc. Interface, 8 (2011), 67-73.

[6] M.J. Keeling, K.T.D. Eames. *Networks and epidemic models*. J. Roy. Soc. Interface, 2 (2005), 295-307.

[7] N. Nagy, P.L. Simon. *Monte-Carlo simulation and analytic approximation of epidemic processes on large networks*. Central European Journal of Mathematics, 11 (4) (2013), 800-815.

[8] M. E. J. Newman. *The structure and function of complex networks*. SIAM Review, 45 (2003), 167-256.

[9] J. Saramäki, K. Kaski. *Modelling development of epidemics with dynamic small-world networks*. J. Theor. Biol., 234 (2005), 413-421.

[10] P.L. Simon, M. Taylor, I.Z. Kiss. *Exact epidemic models on graphs using graph automorphism driven lumping*. J. Math. Biol., 62 (2010), 479-508.

[11] M. Taylor, P.L. Simon, D. M. Green, T. House, I.Z. Kiss. *From Markovian to pairwise epidemic models and the performance of moment closure approximations*. J. Math. Biol., 64 (2012), 1021-1042.

[12] H.C. Tijms. *A First Course in Stochastic Models*. John Wiley and Sons, 2003.

https://doi.org/10.1051/mmnp/20149206 Published online by Cambridge University Press