Time to metastable state in SIS epidemics on graphs

Ruud van de Bovenkamp
Network Architectures and Services
Delft University of Technology
Email: R.vandeBovenkamp@tudelft.nl

Piet Van Mieghem
Network Architectures and Services
Delft University of Technology
Email: P.F.A.VanMieghem@tudelft.nl

***Abstract*—We define the spreading time in the SIS process as the average time between the start of the outbreak and the time that the number of infected nodes first reaches the average number of infected nodes in the metastable state. We show that the spreading time can be computed using a uniformised embedded Markov chain and give numerical results for the complete graph and the star graph. For the complete graph we derive, using the same method, an analytical expression for the spreading time starting from a single infected node. We show that the spreading time is only significantly larger for a single initially infected than when a few nodes are infected, and scales logarithmically as a function of the network size for a fixed fraction of infected nodes in the metastable state. We also show that mean-field methods predict that the spreading time in regular graphs is independent of the degree. For graphs with a high epidemic threshold, the spreading time is lower than for graphs with a low epidemic threshold. The spreading time seems to be related to the average hopcount in the graph. For graphs that have a relatively low average hopcount, the spreading time scales logarithmically, but for graphs with a high average hopcount, such as the rectangular grid and the ring graph, this is not the case.**

I. Introduction

When faced with the outbreak of an infectious disease, at least two concerns arise immediately: (i) How many people are expected to become infected and (ii) How much time, on average, will elapse before the disease has reached the highest number of victims, given that it started with only a small number of infected individuals. In this article, we will elaborate on the second question, by confining ourselves to the relatively simple Markovian Susceptible-Infected-Susceptible (SIS) epidemic model [1], [2], [3], [4]. In the SIS model, a node can be in one of two states: healthy or infected. Nodes in the healthy state can move to the infected state, while nodes in the infected state can move back to the healthy state. The infection process describes how nodes pass from the healthy state to the infected state, and the curing process describes the reverse movement from the infected to the healthy state. In the classic Markovian SIS model, both the curing and infection processes are Poisson processes. The Poissonian curing process has a rate δ and is a nodal process, which is not influenced by the viral state of the neighbours of the infected node. The infection process, however, is a per link Poisson process with a rate β between each healthy and infected node. The total rate of change from the healthy state to the infected state for a node i is given by β times the number of infected neighbours. The ratio between the infection rate β and the curing rate δ is called the effective infection rate $\tau = \frac{\beta}{\delta}$.

In the Susceptible-Infected-Susceptible (SIS) epidemic model on a given contact network G with N nodes and L links, the above first question can be answered by computing the average number of infected nodes in the metastable state, either by simulations or by approximations, most often of a mean-field type [5]. Only for a few graphs, an exact analysis is possible. As is well-known [15], [16], [17], the SIS model features a sharp transitional behaviour around the epidemic threshold at τ_c: infections with an effective infection rate lower than the threshold ($\tau < \tau_c$) will die out very quickly and infect only a very limited portion of the nodes, but viruses with an effective infection rate above the epidemic threshold ($\tau > \tau_c$) will stay in the network for a very long time and infect a sizeable portion of the population. That average fraction of infected nodes is called the metastable fraction of infected nodes and is denoted by $y_\infty(\tau) \in [0, 1]$. A considerable research effort [7] over the past decades has been devoted to determining the epidemic threshold τ_c in networks.

Due to the absorbing state in the SIS model, the real steady-state of any outbreak in any *finite* network is the all-healthy state [5]. The time until the network reaches the all-healthy state is called the survival time, or alternatively, the time to absorption or the extinction time. The survival time for the SIS process on a general graph [8], [9] is difficult to compute analytically or even numerically for general graphs that are larger than about 10 nodes. In general, the survival time distribution has an exponential tail that is dominated by the second largest eigenvalue of the infinitesimal generator of the Markov chain. In the case of the complete graph and the star graph, however, the survival time can be computed numerically as a result of the much smaller state-space of the Markov chain [8], [9]. The survival time in graphs is unrealistically long, compared to the actually observed time that an epidemic lasts. Therefore, the real interest in epidemiology lies in the quasi-stationary or metastable state, rather than in the exact steady-state (which is the absorbing state) of the SIS process.

Therefore, in addition to the survival time of the virus outbreak, the second question above asks for the time until an outbreak reaches the metastable state. The *average* time to first reach a number of infected nodes equal to the average number of infected nodes in the metastable state is called the *spreading time*. The spreading time of a virus in a particular graph is an important metric, since it limits the time in which the virus can still be contained with relative ease. In the extreme case of only a single infected individual, quarantining is easy and highly effective. Once the virus reaches the metastable state ($\tau > \tau_c$), it will be more difficult to eradicate the infection. Hence, the spreading time is generally the longest time for an authority (government, agency, etc.) to react and to start immunization actions (such as vaccination and quarantining).

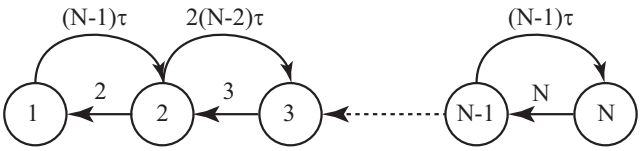

Fig. 1. Markov graph of the modified SIS process on a complete graph. The state numbers coincide with the number of infected nodes. The transition rates are indicated next to the arrows.

In this paper, we numerically compute the spreading time of the Markovian SIS process on the complete graph and the star graph, and investigate the scaling behaviour both in terms of the graph size N and effective infection rate τ. For the complete graph, we derive an analytical expression for the spreading time starting from a single node. For other graph types, we use simulations to derive the spreading time.

II. Spreading Time via the Hitting Time

In general, the spreading time in a particular graph can be found using the Markov description of the SIS process. Let $\bar{y}$ be the average number of infected nodes in the metastable state, rounded to the nearest integer, i.e. $\bar{y} = \lfloor N y_\infty(\tau) \rceil$. The spreading time can be defined as the average hitting time of the set of states for which the total number of infected nodes equals $\bar{y}$, starting from any state. Unfortunately, this method is infeasible for a general graph due to the exploding state space. For the complete graph and the star graph, however, as a result of the symmetry in the infectious state of the network, the state space scales linearly in the number of nodes, which enables us to use the hitting time to determine the spreading time.

We use an embedded Markov chain to transform the continuous-time SIS Markov process into a discrete-time one. The embedded Markov chain of a continuous-time process contains the transition probabilities at the time of a transition, but no longer contains the precise timing of the events [5]. The transition matrix of the uniformised embedded Markov chain in units of ϕ of a continuous-time Markov chain is given by $S(\phi) = I + \frac{Q}{\phi}$, where I is the identity matrix, and Q is the infinitesimal generator of the continuous-time Markov chain [5], which has elements q_{ij} containing the transition rate from state i to state j, for $i \neq j$ and $-q_i = \sum_{j=1, j\neq i}^{N} q_{ij}$ for q_{ii}. Transitions in the uniformised Markov chain all occur with the same rate $\phi \geq \max_i q_i$. Note that the embedded Markov chain contains self loops to uniformise the transition rate from state to state.

Before proceeding, we modify the SIS Markov chain so that it has a well defined steady-state. As explained in [10], [11], the metastable state for finite graphs can be defined in two ways: either we add a nodal self-infection [10] or we remove the absorbing state [11]. The latter modified SIS model, denoted by MSIS, obeys the same evolution rules as the SIS model, except that when there is only one infected node in the network, this node is forbidden to heal. In both cases, the Markov chain is irreducible and features a unique steady-state, that corresponds (by definition) to the metastable state in the original SIS Markov process. In this paper, we modify the SIS process by removing the absorbing state, ensuring that the virus/infection will always stay in the network and can never die out. Thus, as the infection is prevented to die out, our modification will return an upper bound to the exact SIS spreading time in a graph.

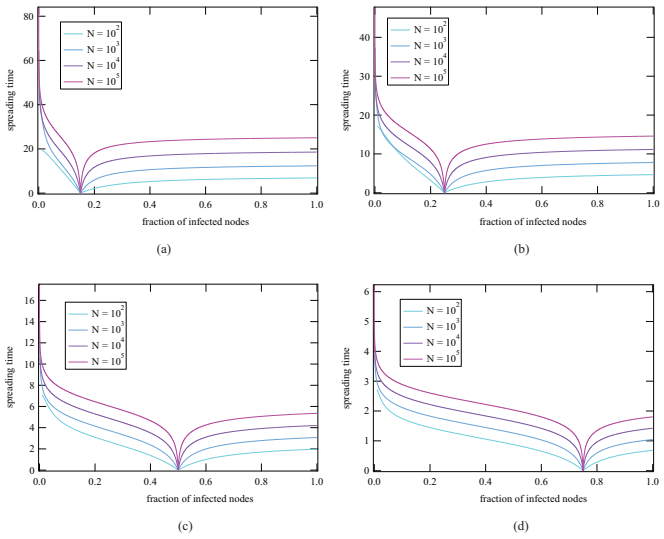

Fig. 2. Spreading time in the complete graph as a function of the fraction of initially infected nodes for four values of the normalised equilibrium point: 0.15 (a), 0.25 (b), 0.5 (c), and 0.75 (d).

The average hitting time of the set of states $\mathcal{A}$ starting from any state in the Markov chain can be found as the minimal non-negative solution of the following system [19], [5]:

$$\begin{cases} w_i = 0 & \text{for } i \in \mathcal{A} \\ w_i = 1 + \sum s_{ij}(\phi) w_j & \text{for } i \notin \mathcal{A} \end{cases}, \tag{1}$$

where w_i is the average hitting time of the set $\mathcal{A}$ starting from state i, and $s_{ij}(\phi)$ is the transition probability from state i to j in the uniformised embedded Markov chain. System (1) can be written as $Kw = b$, where $K = I - \widetilde{S}(\phi)$, and $\widetilde{S}(\phi)$ is the transition matrix of the uniformised embedded Markov chain with each row $i \in \mathcal{A}$ replaced with the standard basis vector e_i multiplied by some scalar. Because of the structure of b, the scalar can take any value, but for numerical reasons its convenient to take ϕ. In short, the $\widetilde{S}(\phi)$ matrix differs from the $S(\phi)$ matrix in that every row corresponding to a state in $\mathcal{A}$ has a non-zero element in the diagonal position, and zeros in all other positions. The vector b is defined as $b = \boldsymbol{u} - \sum_{i\in\mathcal{A}} e_i$ where $\boldsymbol{u}$ is the all-one vector. The minimal non-negative solution of the system for w_i will give the average hitting time of the absorbing state when starting in state i in the number of transitions, whereas multiplying K by the transition rate ϕ will give the average hitting *time*. Therefore, we solve the system

$$\phi K w = b, \tag{2}$$

where ϕK simplifies to $\phi K = \phi(I - I + \frac{\widetilde{Q}}{\phi}) = -\widetilde{Q}$. Again, $\widetilde{Q}$ differs from Q in that every row i corresponding to a state in $\mathcal{A}$ has been replaced by the standard basis vector e_i. In the case of the complete graph and the star graph, $\widetilde{Q}$ is a sparse matrix and matrix equation (2) can efficiently be solved numerically.

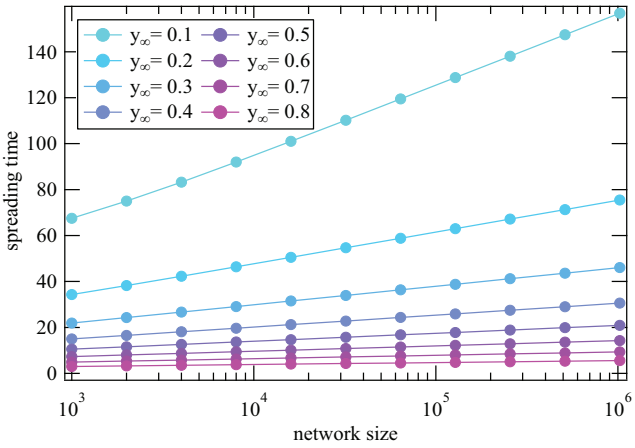

Fig. 3. Spreading time to the equilibrium point starting from a single infected node in the complete graph as a function of the network size N for various values of y_∞.

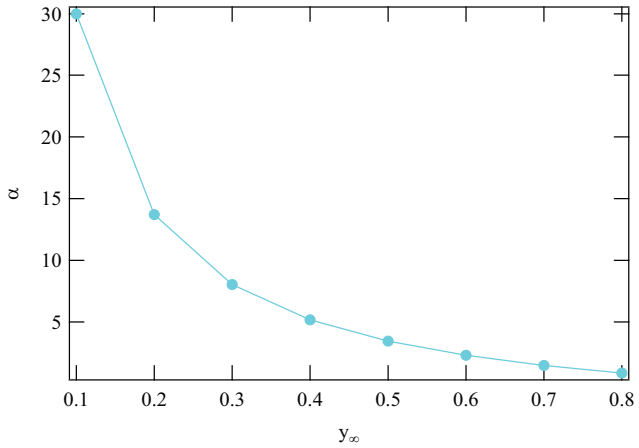

Fig. 4. Slope of the fit of the spread time in Fig. 3 as a function of the equilibrium point y_∞.

III. Spreading Time in K_N

In this section, we investigate the spreading time in the complete graph K_N with N nodes. The metastable state is specified via the equilibrium point in K_N, which is defined as the number I of infected nodes for which $\beta(N-I)I = \delta I$ holds (from here onwards I will always mean the number of infected nodes). The equilibrium point does not necessarily lie at an integer value for I, which is why we also use y_∞ to signify the (normalised) equilibrium point. At the equilibrium point in K_N, and generally in the metastable state for any graph, the total infection and curing rate are in balance. The time to the equilibrium point can be found as the hitting time of the state where $\beta I(N-I) = \delta I$ holds, starting from any state. In the case of a real virus outbreak, the spreading time is indicative of how quickly measures to counter the infection need to be taken. Figure 1 shows the Markov graph of the modified SIS process on the complete graph. Since the state numbers coincide with the number of infected nodes, the spreading time or time to equilibrium is found by solving system (2) for $\mathcal{A} = \{\bar{y}\}$. The infinitesimal generator Q can be derived by inspection from Fig. 1 and is given in the appendix. From Q, we create $\widetilde{Q}$ by replacing row $\bar{y}$ by $e_{\bar{y}}$.

Figure 2 shows the spreading time as a function of the fraction of initially infected nodes in the complete graph for various network sizes N, for various values of the normalised equilibrium point. The normalised equilibrium point was chosen at $15\%, 25\%$, 50% and 75% of infected nodes for all sizes. The spreading time drops sharply from a single initially infected node infected to a few initially infected nodes. The sharp drop is caused by the fact that, when initially a single node is infected, the probability that the virus dies out is the highest. However, in the MSIS model, dying out is forbidden, so that the time to reach the metastable state or equilibrium point is long. As soon as a few more nodes are infected, the infection is very likely to reach the equilibrium point. Usually, infectious diseases start at a few individuals/nodes and Fig. 2 illustrates that *it is crucial to detect an outbreak in a very early state*, because only at a very early state (with few infected) the spreading time is significantly larger than for other states. When initially far more nodes are infected than in the metastable state, the time to the equilibrium point is almost constant.

Figure 3 shows the spreading time starting from a single infected node as a function of the network size N for various positions of the normalised equilibrium point y_∞. The spreading time scales logarithmically with the network size for K_N. The logarithmic scaling in the number of nodes is especially interesting, because the scaling of the infection rate to keep the equilibrium point, for example, at 50% infected is linear, as follows from $\tau(N-I)I = I$. Which can be written as $(\tau N - 1)I - \tau I^2 = 0$, and solving for I gives $I = 0$ and $I = N - \frac{1}{\tau}$. Despite the fact that the effective infection rate τ decays with the reciprocal of the network size, the time to reach the equilibrium increases only logarithmically. This means that, in a larger network, a weaker virus can infect the same fraction of nodes in roughly the same time, again emphasizing the importance to spot an outbreak early.

Figure 3 shows that the spread time in the complete graph scales as $\alpha \log(N) + b$, Fig. 4 shows the fits of α in Fig. 3 as a function of the equilibrium point y_∞. With an increasing equilibrium point, α reduces quickly, indicating that a stronger virus does not only infect more individuals, but also infects them quicker than a weaker virus.

Figure 5 shows the spreading time in the complete graph starting from a single infected node as a function of the equilibrium point for various network sizes. We observe from Fig. 5 that for small values of the equilibrium point, the time to reach that equilibrium point starting from 1 node infected increases. This is caused by the smaller values of the effective infection rate τ needed to reach the equilibrium point. Interestingly, the time to reach the equilibrium point peaks due to the probability that the virus dies out when it starts with only a single infected node and a low effective infection rate τ. Because the absorbing state is removed, the process spends more time in the state with only one node infected for low effective infection rates, which increases the time to reach the equilibrium point. With an increasing number of infected nodes at the equilibrium point, the effective infection rate τ increases

until the infection rate is so large that dying out becomes unlikely. This illustrates that an increasing effective infection rate τ does not only increase the number of infected nodes in the metastable state, but also increases the probability that the metastable state will be reached from a single infection, and thus reduces the time to reach the metastable state.

A. Analytic Expression for the Spreading Time in K_N

By modifying the Markov chain in Fig. 1, and solving matrix equation (2) for the modified chain, we derive an exact expression for the spreading time in K_N starting from a single infected node to z infected nodes. In this section we slightly abuse the definition of spreading time and use it to denote the average time until a virus reaches a state with z nodes infected, starting from a single infected node in the modified SIS process. The difference here is that z does not have to be equal to the average number of infected nodes in the metastable state. The chain is modified by making the state with z infected nodes absorbing. The modified Markov chain is shown in Fig. 6, the modified infinitesimal generator $-\widetilde{Q}$ can be found by inspection and is also given in the appendix. Matrix equation (2) can be solved by reducing the augmented matrix $[-\widetilde{Q}|b]$ to row echelon form. To reach row echelon form, we iteratively perform the following row operation for $k \geq 2$.

$$r_{k+1} \to r_{k+1} + \frac{(z-k)(N-z+k)\tau}{d_k} r_k$$

where r_k indicates row k and d_k is given recursively by

$$d_k \begin{cases} (z-1)+(z-1)(N-z+1)\tau & \text{for } k=2 \\ (z-k+1)[1+(N-z+k-1)\tau \\ \quad -\frac{(z-k+2)(N-z+k-1)\tau}{d_{k-1}}] & \text{for } 2<k<z \\ (N-1)\tau - \frac{2(N-1)\tau}{d_{k-1}} & \text{for } k=z \end{cases}$$

The elements in column vector b can be found recursively as

$$b_k = 1 + \frac{(z-k+1)(N-z+k-1)\tau}{d_{k-1}} b_{k-1}$$

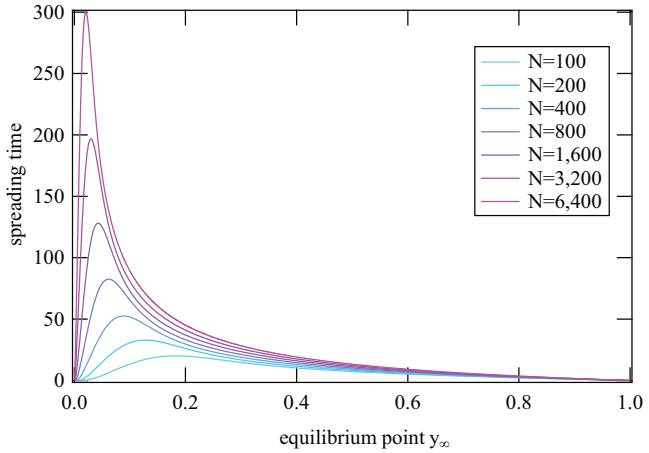

Fig. 5. The time to first reach the equilibrium point in a complete graph starting from a single infected node as a function of the equilibrium point for various graph sizes N.

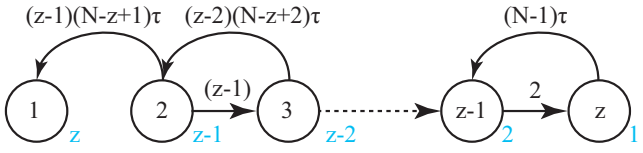

Fig. 6. Modified Markov graph of the modified SIS process on a complete graph. The number of infected nodes is indicated in blue next to the state. Note that the number of infected nodes decreases from left to right.

The spreading time $T_S(z)$ is given by b_z/d_z. Iterating the recursive relation for b_k from $k=z$ backwards leads to

$$b_z = \sum_{j=1}^{z-1} \frac{(j-1)!(N-1)!\tau^{(j-1)} \prod_{i=2}^{i=z-j} d_i}{(N-j)! \prod_{i=1}^{i=z-1} d_i},$$

and dividing by d_z gives

$$T_S(z) = \sum_{j=1}^{y-1} \frac{(j-1)!(N-1)!\tau^{(j-1)} \prod_{i=2}^{i=z-j} d_i}{(N-j)! \prod_{i=1}^{i=z} d_i} \tag{3}$$

Writing $p_j = \prod_{i=2}^{j} d_i$ and using $p_j = d_j p_{j-1}$, we can write

$$p_j = (z-j+1)p_{j-1} + (z-j+1)(N-z+j-1)\tau p_{j-1} \\ -(z-j+2)(z-j+1)(N-z+j-1)\tau p_{j-2}$$

for $2 < j < z$. Iterating the recursive relation for p_j from j backwards leads to

$$p_j = \frac{(z-1)!}{(z-j)!(N-z)!} \sum_{i=0}^{j-1} (N-z+i)!\tau^i$$

and using the correct expression for d_z,

$$p_z = \frac{(z-1)!(N-1)!\tau^{z-1}}{(N-z)!}$$

Filling in the expressions for p_j and p_z into (3) leads to our final result

$$T_S(z) = \sum_{j=1}^{z-1} \sum_{i=0}^{z-j-1} \frac{(N-z+i)!\tau^{i+j-z}}{j(N-j)!}$$

B. Mean-Field Spreading Time in Regular Graphs

We briefly move from the exact world of Markov chains to the approximate world of mean-field theory, and show that mean-field predicts a spreading time in r-regular graphs that is independent of the degree r. In the N-intertwined Mean-Field Approximation (NIMFA) [15], [14] the probability v_i that node i is infected is given by the following first-order nonlinear ordinary differential equation

$$\frac{dv}{dt} = r\beta v(t)(1-v(t)) - \delta v(t), \tag{4}$$

which also appeared in [1], [3], and has the solution (also derived in [18], [6]):

$$v(t) = \frac{1}{(\frac{1}{v_0} - \frac{1}{1-\frac{1}{r\tau}}) \exp(-(r\tau-1)\frac{t}{\delta}) + \frac{1}{1-\frac{1}{r\tau}}} \tag{5}$$

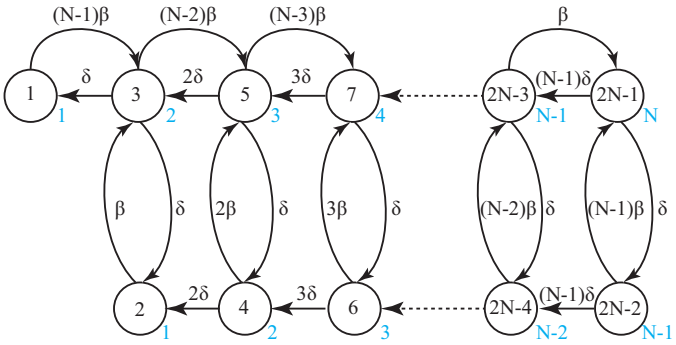

Fig. 7. Markov graph of the modified SIS process on a star graph. In the top row (odd states) the centre node is infected, whereas on the bottom row (even states) the centre is healthy. The number of infected nodes in each state is indicated in blue next to the state.

According to NIMFA, the fraction of infected nodes in the metastable state for an r-regular graph is given by

$$y_\infty = 1 - \frac{1}{r\tau},$$

Alternatively, the effective spreading rate τ that is needed to achieve y_∞ nodes infected in the metastable state is given by

$$\tau = \frac{1}{r(1-y_\infty)}$$

Substituting τ in (5) and starting from a single infected node ($v_0 = \frac{1}{N}$) leads to:

$$v(t) = \frac{1}{(N - \frac{1}{y_\infty})\exp(-\frac{y_\infty}{1-y_\infty}\frac{t}{\delta}) + \frac{1}{y_\infty}}$$

which suggests that the time to reach the metastable state is independent of r. The NIMFA spreading time is found by defining a distance between $v(t)$ and y_∞ and finding the time for which the distance is small. Starting from

$$v(t) = \frac{1}{f(t) + \frac{1}{y_\infty}}$$

with $f(t) = (N - \frac{1}{y_\infty})\exp(-\frac{y_\infty}{1-y_\infty}\frac{t}{\delta})$, we define the distance to metastable state as $y_\infty - \frac{1}{f(t)+\frac{1}{y_\infty}}$, and the spreading time as the time for which the distance to the metastable state is smaller than ε. From $y_\infty - \frac{1}{f(t)+\frac{1}{y_\infty}} \leq \varepsilon$, we find that

$$t \geq \frac{(1-y_\infty)\delta}{y_\infty}(\ln(N - \frac{1}{y_\infty}) + \ln(y_\infty^2 - y_\infty\varepsilon) - \ln(\varepsilon))$$

Although the NIMFA spreading time is not always accurate, and in some cases not even predicts the right scaling of the spreading time, as shown in Sec. V, at least for the complete graph K_N, it correctly shows a logarithmic scaling in N.

IV. Spreading Time in $K_{1,N-1}$

In this section, we investigate the spreading time in the star graph $K_{1,N-1}$ with N nodes. The Markov graph of the modified SIS process on a star graph, as adapted from [11] is shown in Fig. 7. The Markov chain consists of $2N-1$ states, to make a distinction between the situation where the centre node is infected (all odd states), and the situation where the centre node is healthy. The number of infected nodes in each

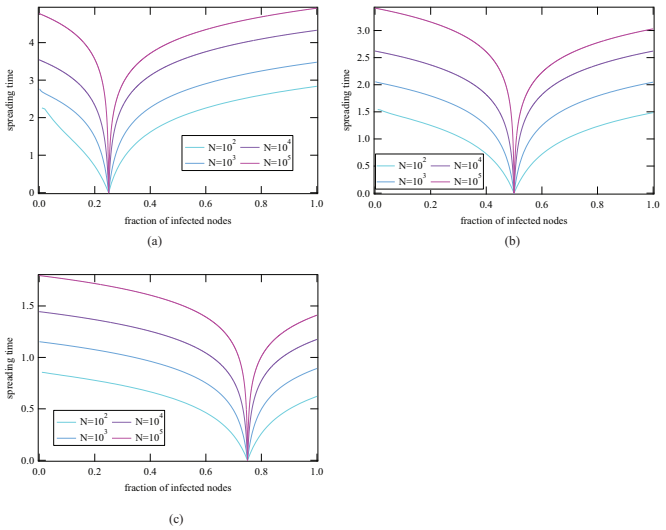

Fig. 8. Spreading time in the star graph for the case that the centre is infected as a function of the fraction of infected nodes for three different values of $\bar{y}/N$: 0.25 (a), 0.5 (b), and 0.75 (c).

state is indicated in blue next to the state in Fig. 7. For each possible number of infected nodes (with the exception of N infected nodes), the network can be in one of two states. As a result, the spreading time can be computed by solving system (2) with $\mathcal{A} = \{2\bar{y}-1, 2\bar{y}\}$, for $\bar{y} < N$, or $\mathcal{A} = \{2N-1\}$ for $\bar{y} = N$.

Figure 8 shows the spreading time in the star graph for various network sizes and values of $\bar{y}$, under the condition that the centre node is infected. The spreading time in the star graph is smaller than for the complete graph. This is caused by the higher effective spreading rate τ that is needed to reach the same $\bar{y}$ in the star graph compared to the complete graph. In a graph with a high epidemic threshold, a virus will generally infect fewer nodes compared to a network with a low epidemic threshold. However, if a virus infects the same fraction of nodes in both networks (of course, this is not possible for the same effective spreading rate τ), it infects those nodes quicker in the network with the higher epidemic threshold. Paradoxically, the better we protect our networks, the quicker a virus will reach the metastable state. Of course, for a virus with a fixed effective spreading rate τ, protecting the network by, for example, link removals will have an effect on the number of infected nodes in the metastable state. However, in the context of computer viruses or other engineered infectious processes, better network protection will have to lead to stronger viruses to achieve the same goal (infecting nodes). A similar drive might be present in biological evolutionary processes.

The spreading time in Fig. 8 is under the condition that the centre node is infected. An infection starting in the centre node has a better chance of spreading through the network than an infection starting in one of the leaf nodes. Yet, the spreading time for a virus starting in a leaf node does not differ too much from that of a virus starting in the centre, as shown in Fig. 9. Figure 9 shows the spreading time as a function of the number of initially infected nodes for three different network sizes, both under the assumption that the centre node is infected (round markers) and under the assumption that the

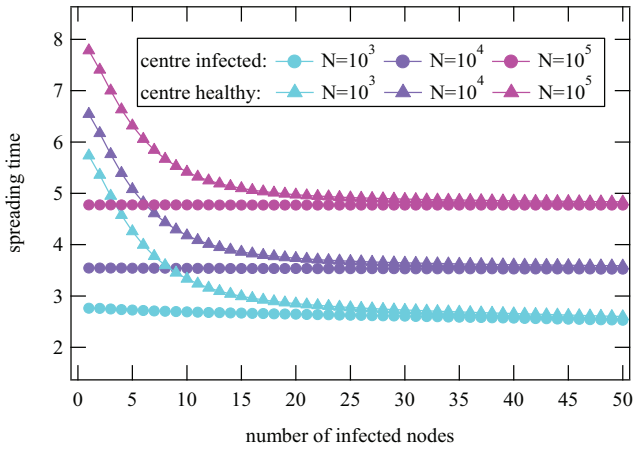

Fig. 9. Spreading time in the star graph as a function of the number of infected nodes for three different network sizes, both for the case that the centre is infected (round markers) and for the case that the centre is healthy (triangular markers).

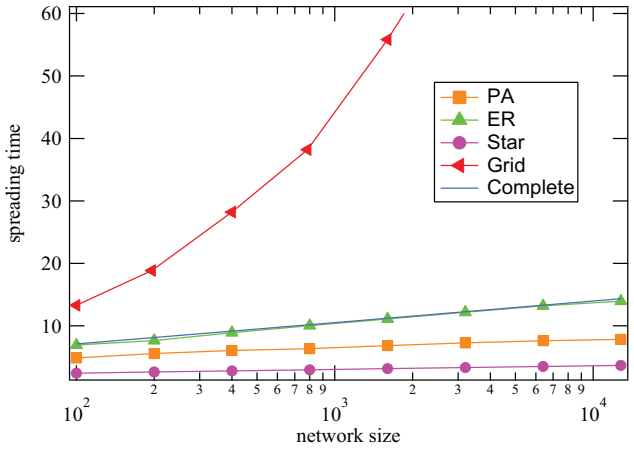

Fig. 10. Simulated spreading time starting from a single initially infected node as a function of the network size for five different graph types. The fraction of infected nodes in the metastable state is 50%.

centre node is healthy (triangular markers). In terms of the modified Markov process in Fig. 7: the round markers indicate the spreading time starting in an odd state, the triangular markers indicate the spreading time starting from an even state. For all three network sizes, the difference between starting in an odd state (centre infected) or an even state is roughly the same and diminishes quickly. If more than 20 nodes are initially infected, the spreading time under the assumption that the centre is infected is approximately the same as under the assumption that it is healthy.

V. Spreading Time in Other Graphs

The spreading time in a general graph cannot be determined using the Markov transition probability matrix as for the complete graph K_N or the star, due to the 2^N state space in general [5]. In the case of a general graph, we simulate the time until the process first reaches the number of infected nodes in the metastable state. We first determine the metastable state as a function of the effective infection rate for each graph and select the effective infection rate τ that corresponds to 50% infected in the metastable state. The data points for each graph type and size are averages over $100,000$ runs.

Figure 10 shows the spreading time as a function of the network size N for 5 different graph types, described and studied in [5] and [12]: the connected Erdős-Rényi random graph $G_p(N)$ with a link density or link probability p at the connection threshold $p_c = \frac{\ln N}{N}$, the preferential attachment graph (PA) with m links per new node (m is chosen so that the link density is equal to that in the ER graph), the star graph, the square lattice and the complete graph. Curiously, the connected Erdős-Rényi random graph $G_p(N)$ behaves almost identical to the complete graph $K_N = G_1(N)$: the spreading time seems hardly to depend on the link density $p = \frac{2L}{N(N-1)}$. Probably, the difference in link density is compensated for by the higher effective spreading rate τ that is needed to reach the same $\bar{y}$. The presence of hub nodes in the preferential attachment graph, however, leads to a smaller spreading time, especially for larger graphs. The extreme case in terms of hub nodes is the star graph. Indeed, in the star graph, the time to reach the metastable state is shortest. These four graph types all show a logarithmic scaling of the spreading time with the number of nodes, but the slope at the log-lin scale is smaller for graphs with more hub nodes.

The square grid does not show the logarithmic scaling in N that is observed in the other graph types. Indeed, Fig. 11 shows the spreading time for just the rectangular grid and a fit, suggesting that the spreading time scales as $\sqrt{N}$ with the size N. This is most likely caused by the large average hopcount in this graph type. The average hopcount in the lattice [5, p.630] equals $E[H_{\text{lattice}}] \simeq \frac{2}{3}\sqrt{N}$, for large N, whereas $E\left[H_{G_{p_c}(N)}\right] \sim \frac{\ln N}{\ln \ln N}$ (see [5, p. 428]) increases somewhat slower than logarithmically in N. Also for sparse scale-free graphs of which the Barabasi-Albert preferential attachment graph is an instance [5, p. 428], the average hopcount is $E[H_{\text{PA}}] \sim \frac{\ln N}{\ln(d_{av}-1)}$, where $d_{av} = \frac{2L}{N}$

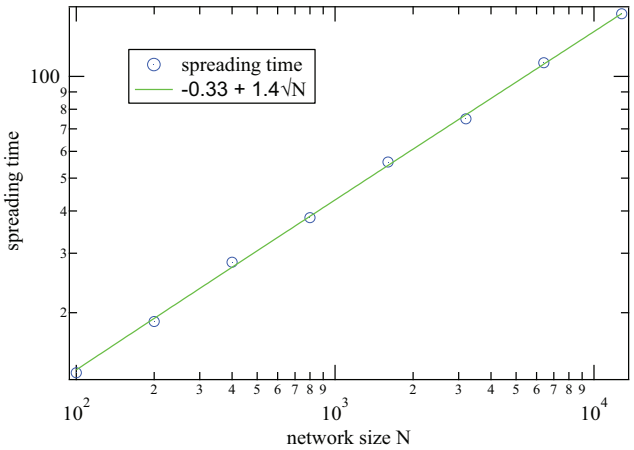

Fig. 11. Simulated spreading time starting from a single initially infected node in the rectangular grid. The fit shows that the spreading time scales as $O(\sqrt{N})$

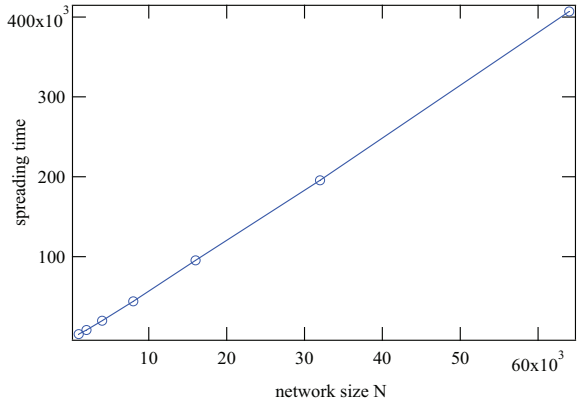

Fig. 12. Simulated spreading time starting from a single initially infected node in the ring graph.

is the average degree. Hence, for the lattice, the connected Erdős-Rényi random graph $G_{p_c}(N)$ and scale-free graphs, the scaling of the spreading time and the average hopcount agree. However, the relation between spreading time and hopcount ceases for the star and complete graph: the average hopcount in K_N is $E[H_{K_N}] = 1$ and in the star $K_{1,N-1}$ with N nodes is $E\left[H_{K_{1,N-1}}\right] = 2 - \frac{2}{N}$, thus both bounded by a constant, whereas logarithmic increases is observed in Fig. 10. In any graph, the average time [13],[5, p. 460] to hit the absorbing state in SIS epidemics for $\tau < \tau_c$ scales logarithmically in N, illustrating that, on average, the epidemic process cannot tend faster to the metastable state than $O(\log N)$, irrespective of the topology. The influence of the network topology on the SIS dynamic process is measured via the average hopcount which is a good representative measure for the diffusive spread of the epidemic. Above the epidemic threshold ($\tau > \tau_c$), the epidemic diffuses over the network. The spreading time thus reflects the combined logarithmic scaling (as a lower bound for the average epidemic extinction time and, hence, of the entire SIS dynamic process) and the average diffusion time, which is related to the average hopcount of the graph. In conclusion, only when the average hopcount exceeds a logarithmic scaling, the spreading exhibits the hopcount scaling, else logarithmic scaling is observed. As a final example of a spreading time that scales with the hopcount, Fig. 12 shows the linear scaling in N of the spread time on a ring graph.

VI. Conclusion

The temporal properties of the SIS model have received less attention in the literature than properties such as the epidemic threshold and the average fraction of infected nodes in the metastable state. This paper is devoted to the spreading time of an SIS epidemic. We define the spreading time as the average time to first reach the number of infected nodes in the metastable state, measured from the beginning of the outbreak.

The spreading time can be computed using the uniformised embedded Markov chain of the modified SIS process. In the general case, however, the state space of the Markov chain describing the SIS process on a graph scales as $O(2^N)$. Only for graphs that allow for a smaller state space, such as the complete graph and the star graph can we compute the spreading time directly. Moreover, for the complete graph, we derive an analytic expression for the spreading time starting from a single node. Mean-field approximations on regular graphs predict a spreading time that is independent of the degree, which can be far from the true value for low degrees.

Our results show that in the complete graph, and to a lesser extent in the star graph, the spreading time as a function of the fraction of infected nodes drops sharply from a single node infected and then stabilises. This shows that any preventive action to an outbreak is better taken quickly after the initial infection.

The spreading time for a fixed fraction of infected nodes scales logarithmically as a function of the network size in the complete graph. In growing networks or populations this is worrying, as it means that an increasingly weaker virus can infect the same fraction of nodes in roughly the same time.

The spreading time in the star graph as compared to the complete graph is much shorter for the same fraction of infected nodes in the metastable state. As the star graph has a much higher epidemic threshold, the effective spreading rate has to be higher in the star graph to reach the same number of infected nodes in the metastable state than in the complete graph. This leads to the shorter spreading time.

For other graphs than K_N and $K_{1,N-1}$, we have used simulations to determine the spreading time. For the Erdős-Rényi random graph and the scale-free graph the spreading time scales logarithmically, just as for the complete graph and star graph. The logarithmic scaling of the spreading time might be connected to the fact that the average hopcount is relatively low compared to the network size in these graphs. In the case of the rectangular grid (where the average hopcount scales with $O(\sqrt{N})$), no logarithmic scaling of the spreading time is observed, but a scaling as $O(\sqrt{N})$. Similarly, in the ring graph, the spreading time scales with $O(N)$, just as the hopcount.

Knowing that many real-world epidemic processes are not described well by Markovian SIS, the spreading time in non-Markovian SIS is of great interest in future research. The inter-arrival time distribution of the infection and curing processes have a great influence on the survival time and the average fraction of infected nodes in the metastable state, and will most likely influence the spreading time as well.

Acknowledgement

This work has been partially supported by the European Commission within the framework of CONGAS (project FP7-ICT-2011-8-317672) and EU Network of Excellence EINS (Grant FP7-ICT 288021).

Appendix
Matrices

The Q matrix for the modified SIS process on the complete graph corresponding to Fig. 1 is given by eq. 6. The $\widetilde{Q}$ matrix for the modified Markov chain of the modified SIS process in the complete graph K_N corresponding to Fig. 6 is given by eq. 7. In general, the Q matrix is the infinitesimal generator of the continuous-time Markov process and has elements q_{ij} containing the transition rate from state i to state j, for $i \neq j$

$$-Q = \begin{bmatrix} (N-1)\tau & -(N-1)\tau & 0 & 0 & 0 & 0 & 0 \\ -2 & 2+2(N-2)\tau & -2(N-2)\tau & 0 & 0 & 0 & 0 \\ 0 & -3 & 3+3(N-3)\tau & -3(N-3)\tau & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & -(N-1) & N-1+(N-1)\tau & -(N-1)\tau \\ 0 & 0 & 0 & 0 & 0 & -N & N \end{bmatrix} \tag{6}$$

$$-\widetilde{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -(y-1)(N-y+1)\tau & (y-1)+(y-1)(N-y+1)\tau & -(y-1) & 0 & 0 & 0 & 0 \\ 0 & -(y-2)(N-y+2)\tau & (y-2)+(y-2)(N-y+2)\tau & -(y-2) & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & -2(N-2)\tau & 2+2(N-2)\tau & -2 \\ 0 & 0 & 0 & 0 & 0 & -(N-1)\tau & (N-1)\tau \end{bmatrix} \tag{7}$$

and $-q_i = \sum_{j=1, j\neq i}^{N} q_{ij}$ for q_{ii}. The step from Q to $\widetilde{Q}$ involves replacing all the rows corresponding to states in $\mathcal{A}$ with their standard basis vector.

REFERENCES

[1] N. T. J. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications*. Charlin Griffin & Company, London, 2nd edition, 1975.

[2] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford, U.K., 1991.

[3] D. J. Daley and J. Gani. *Epidemic modelling: An Introduction*. Cambridge University Press, Cambridge, U.K., 1999.

[4] O. Diekmann, H. Heesterbeek, and T. Britton. *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton University Press, Princeton, USA, 2012.

[5] P. Van Mieghem. *Performance Analysis of Complex Networks and Systems*. Cambridge University Press, Cambridge, U.K., 2014.

[6] P. Van Mieghem. SIS epidemics with time-dependent rates describing ageing of information spread and mutation of pathogens TU Delft technical report 20140615, 2014

[7] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Review of Modern Physics*, arXiv:1408.2701, to appear 2014.

[8] P. Van Mieghem. Decay towards the overall-healthy state in SIS epidemics on networks. *arXiv:1310.3980*, 2013.

[9] R. van de Bovenkamp and P. Van Mieghem. Survival time of the SIS infection process on a graph. unpublished 2014.

[10] P. Van Mieghem and E. Cator. Epidemics in networks with nodal self-infections and the epidemic threshold. *Physical Review E*, 86(1):016116, July 2012.

[11] E. Cator and P. Van Mieghem. Susceptible-Infected-Susceptible epidemics on the complete graph and the star graph: Exact analysis. *Physical Review E*, 87(1):012811, January 2013.

[12] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, U.K., 2011.

[13] M. Draief and L. Massoulié. *Epidemics and Rumours in Complex Networks*. London Mathematical Society Lecture Node Series: 369. Cambridge University Press, Cambridge, UK, 2010.

[14] P. Van Mieghem, "The N-intertwined SIS epidemic network model," *Computing*, vol. 92, no. 2, pp. 147–169, 2011.

[15] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, 2009.

[16] C. Castellano and R. Pastor-Satorras, "Thresholds for epidemic spreading in networks," *Physical Review Letters*, vol. 105, no. 21, p. 218701, 2010.

[17] R. Parshani, S. Carmi, and S. Havlin, "Epidemic threshold for the susceptible-infectious-susceptible model on random networks," *Physical Review Letters*, vol. 104, no. 25, p. 258701, 2010.

[18] J. O. Kephart and S. R. White, "Direct-graph epidemiological models of computer viruses, *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 343–359, 1991.

[19] J. Norris, *Markov Chains*, ser. Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, 2005.