Epidemic progression on networks based on disease generation time

Bahman Davoudia, Flavia Mosera, Fred Brauerb and Babak Pourbohloula,c*

aDivision of Mathematical Modeling, British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada; bDepartment of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada; cSchool of Population & Public Health, University of British Columbia, Vancouver, British Columbia, Canada

(Received 31 July 2012; final version received 19 June 2013)

We investigate the time evolution of disease spread on a network and present an analytical framework using the concept of disease generation time. Assuming a susceptible–infected–recovered epidemic process, this network-based framework enables us to calculate in detail the number of links (edges) within the network that are capable of producing new infectious nodes (individuals), the number of links that are not transmitting the infection further (non-transmitting links), as well as the number of contacts that individuals have with their neighbours (also known as degree distribution) within each epidemiological class, for each generation period. Using several examples, we demonstrate very good agreement between our analytical calculations and the results of computer simulations.

Keywords: epidemiology; population dynamics

1. Introduction

The spread of disease in a population has recently been the subject of extensive studies. Disease spread has been studied using various models and approaches, among them are compartmental models [1,2,4,8], network models [10,14,15,19,26,27], and agent-based models [7,21,22]. Compartmental models generally assume homogeneous mixing for the contacts between individuals. This assumption implies that every infected individual is equally likely to spread the disease to a susceptible individual, which is not realistic for most infectious diseases. A further drawback of compartmental models is that they cannot describe the initial phase of a disease outbreak, in which the number of infectious individuals is small and the progression of an outbreak is dominated by stochastic fluctuations [5]. In network models, a pathogen is transmitted between connected infectious and susceptible individuals with a certain probability, generally referred to as transmissibility, \( T \). The connectivity of individuals in the network model is coded within the topology of the network. These two components – that is, disease transmissibility and the contact network structure – allow modelling of the stochasticity of a disease outbreak, in particular at
the beginning of an epidemic. In network models, connections are static, whereas in dynamic network/agent-based models they are dynamic.

Many approaches have been proposed to model disease spread on networks, including the mean value reaction equation \([3,9,13,20,23–25]\), generating function formalism \([12,16,17]\), time evolution of the generating function \([11,18]\), and computer simulations \([6,10]\). Among these, the second approach focuses only on the final outcome of an epidemic, while others study the time evolution of the disease spread. This paper presents our framework for studying time progression of disease spread on a network in terms of disease generation time. Within our context, generation time for a disease is defined as the time interval from when an individual is exposed to a pathogen, followed by a period when he/she remains infectious, until the time when he/she can no longer transmit the infection as a result of complete recovery or death. Using this concept, we define generation 0 as the initial infected individual of the outbreak or epidemic, generation 1 as the individuals becoming infected as a result of contact with the initial infected case, and nodes of generation \(g\) are those who acquire the disease from an infectious member of generation \(g - 1\). This is equivalent to a bond percolation model, where each infectious individual stays infectious for one generation. For a disease such as measles, whose incubation period is longer than infectious period, different generations can readily be distinguished from one another. Disease can transfer between a connected pair of susceptible and infectious individuals with a certain probability, called transmissibility.

This manuscript is organized as follows. In Section 2, we provide a short introduction to networks, to the concept of generation time, and to the definition of different types of links and vertices. In Section 3, we present an analytical framework for calculating the number of new infections in each generation. In Section 4, we present the numerical results of our framework and compare them with simulation results. Finally, in Section 5, we summarize our contributions, including possible extensions to the presented framework. Table 1 summarizes the variables introduced in the rest of the paper.

### Table 1. Notation.

| Symbol | Description |
|--------|-------------|
| \(p_k\) | Degree distribution: probability of a randomly chosen vertex having degree \(k\) |
| \(q_k\) | Probability of a randomly chosen stub belonging to a vertex with degree \(k\) |
| \(z\) | Average degree: \(z = \langle k \rangle_{p_k}\) |
| \(Z\) | Excess degree: average degree of vertices chosen by targeting one of their stubs excluding the targeted stub, \(Z = \langle k - 1 \rangle_{q_k}\) |
| \(T\) | Transmissibility: probability of disease transmission along a link from an infectious to a susceptible vertex |
| \(N\) | Network size: Number of vertices in network |
| \(N^g\) | Number of vertices in class \(a\) at generation \(g\) |
| \(L\) | Number of links in network |
| \(L^g\) | Number of intra-links between class \(a\) and \(b\) at generation \(g\) |
| \(L_{ab}^g\) | Number of inter-links between class \(a\) and \(b\) at generation \(g\) |
| \(p_{a,n}\) | Degree distribution of unprocessed class after \(n\) vertices have been processed |
| \(p_{b,n}\) | Degree distribution of processed class after \(n\) vertices have been processed |
| \(q_{a,n}\) | Probability that \(n\)th processed vertex has degree \(k\) |
| \(z_{a,n}\) | Expected average degree of unprocessed class after \(n\) vertices have been processed |
| \(Z_{a,n}\) | Expected average degree of class \(a\) at generation \(g\) |
| \(\lambda^g\) | Number of transmitting links during generation \(g\) |
| \(\lambda_{NT}^g\) | Number of non-transmitting links during generation \(g\). |
| \(\Lambda^g\) | Total number of non-transmitting links accumulated prior to generation \(g\) |
| \(P^g_j\) | Effective degree distribution of susceptible vertices after discounting \(j\) non-transmitting links at generation \(g\) |
| \(Z^g_j\) | Effective average degree of susceptible vertices after discounting \(j\) non-transmitting links at generation \(g\) |
2. Definitions

This paper analyses the evolution of disease spread on a network using the notion of generation time. A network consists of $N$ vertices (or nodes), which represent individuals, and links (or edges) that represent contacts between these individuals that could potentially lead to disease transmission. A link is made of two stubs, each emanating from its own vertex. Two vertices are called neighbours if they are connected by a link. The degree of a vertex is defined as the number of links connecting the vertex to its neighbours and is equal to the number of stubs emanating from it. The probability distribution of these degrees in the network is called the degree distribution, which is denoted by $p_k$. The average degree, $z$, can be calculated as $z = \sum k p_k$. We define $q_k = k p_k / z$ as the probability that a randomly chosen stub belongs to a vertex of degree $k$. The average excess degree, $Z$, is the average number of stubs emanating from vertices selected by targeting a random stub and can be calculated as $Z = \sum (k - 1) q_k$. This value is the average degree calculated by the probability $q_k \cdot \sum k q_k$, after subtracting the targeted stub.

In order to model the disease spread on a network, we divide the vertices into three classes: the susceptible, infectious, and removed classes, where removed refers to vertices that were once infectious, but no longer transmit infection. Such models are called susceptible–infected–recovered models to indicate these three classes.

We assume that the generation time is equal to the incubation period plus the infectivity period of the disease under consideration. During one generation, an infectious individual infects each of its susceptible neighbours with probability $T$ (transmissibility). In fact, this premise is the basis of contact tracing approaches in public health outbreak management, where all contacts within one disease generation time are identified and tested for potential transmission. The rate of new infections at each generation depends on the transmissibility value and the degree distribution, the number of links, and the number of vertices in each class.

In this paper, we use subscripts $\alpha$ and $\beta$ as generic indices for $s$, $i$, or $r$. We define $N^g_\alpha$ as the number of vertices in class $\alpha$ in generation $g$, and $z^g_\alpha$ as their average degree. Links are referred to as intra-links or inter-links, depending on whether they connect vertices of the same class or of different classes, respectively. The number of intra-links in class $\alpha$ is denoted by $I^g_{\alpha \alpha}$, while the number of inter-links between classes $\alpha$ and $\beta$ is represented by $I^g_{\alpha \beta}$. Note that $I^g_{\alpha \beta} = I^g_{\beta \alpha}$. The different types of links and vertices are illustrated in Figure 1.

![Figure 1. Different types of links and vertices during generation $g$ of disease spread. The quantities $\lambda^g_T$ and $\lambda^g_{NT}$ refer to the number of transmitting and non-transmitting links between classes $S$ and $I$, as defined in Section 3.3.](image)
3. Theoretical framework

This manuscript’s main contribution is a theoretical framework for calculating the time evolution of disease spread on finite-size networks based on the notion of disease generation time. Given that the time to acquire an infection, as well as the infectivity profile, may vary from person to person, accounting in continuous time for time evolution of an epidemic at the individual level would be a daunting task, both analytically and practically. To head in that direction, however, requires developing a theoretical framework where the dynamics of infection transmission are analysed at the individual level, without the need for incorporating specific details of the transmission process such as continuous-time infectivity profiles or the exact timing of exposure to infection. An analytical, discrete-time network approach potentially solves this issue. As such, understanding the dynamics of epidemics during each generation of infection becomes our main focus, with a specific goal of calculating the number of new infections produced during each generation. This number can be calculated from the number of inter-links between susceptible and infectious individuals, \( L_{si} \) (Section 3.1), after discounting multi-targeting (Section 3.3) and non-transmitting links (Section 3.4). The quantity \( L_{si} \) also depends on the class-specific average degree in each generation (Section 3.2).

3.1. Number of intra-links and inter-links

As mentioned before, the number of new infections in each generation depends on the number of links between susceptible and infectious individuals, as denoted by \( L_{si} \). In general, the number of stubs in class \( \alpha \) is given by \( N_{\alpha}^{z_{\alpha}} \), of which \( 2L_{si}^{z_{\alpha}} \) contribute to intra-links and the rest contribute to the inter-links. We can therefore establish the following equations:

\[
\begin{align*}
    z_{\alpha}^{s}N_{s}^{\alpha} &= 2L_{ss}^{\alpha} + L_{si}^{\alpha} + L_{sr}^{\alpha}, \\
    z_{i}^{s}N_{i}^{\alpha} &= 2L_{si}^{\alpha} + L_{ii}^{\alpha} + L_{ir}^{\alpha}, \\
    z_{r}^{s}N_{r}^{\alpha} &= 2L_{sr}^{\alpha} + L_{ir}^{\alpha} + L_{rr}^{\alpha}.
\end{align*}
\]

(1)

The above set of equations imposes a strong constraint on the number of intra- and inter-links that can arise. Thus, the number of inter-links can be calculated provided that the number of intra-links, the average degree in each class, and the number of vertices in each class are all given.

In the following explanation, we calculate the expected number of intra-links in class \( \alpha \) in terms of \( N_{\alpha}^{z_{\alpha}} \) and \( z_{\alpha}^{z_{\alpha}} \). First, we randomly select a stub of a vertex in class \( \alpha \). The number of stubs to which the chosen stub can be connected to from either an intra-link or inter-link is

\[
N_{\alpha}^{z_{\alpha}} - z_{\alpha}^{z_{\alpha}},
\]

(2)

where the second term is responsible for the exclusion of self-loops. The number of stubs in class \( \alpha \) to which the chosen stub can be connected is

\[
(N_{\alpha}^{z_{\alpha}} - 1)z_{\alpha}^{z_{\alpha}}.
\]

(3)

Thus, the probability that the chosen stub is connected to another stub within the same class can be calculated as the ratio of the previous two equations:

\[
\frac{(N_{\alpha}^{z_{\alpha}} - 1)z_{\alpha}^{z_{\alpha}}}{N_{\alpha}^{z_{\alpha}} - z_{\alpha}^{z_{\alpha}}}.
\]

(4)

This probability demonstrates the contribution of the first chosen stub to the number of intra-links in class \( \alpha \).
The probability that a second randomly chosen stub of the same vertex is attached to a stub of another vertex within the same class can be simply calculated by

\[ \frac{(N^\alpha - 2)z^\alpha}{Nz - 2z^\alpha}, \]

which excludes self-loops and repeated links. For the \( j \)th stub, we obtain

\[ \frac{(N^\alpha - j)z^\alpha}{Nz - jz^\alpha}. \]

Equations (2), (3), (5) and (6) are approximations that provide the related expectation values. This process is continued until all stubs of the chosen vertex are exhausted or until all other vertices within the same class have been chosen. More precisely, while

\[ j \leq \gamma^\alpha = \min\left(\frac{N^\alpha}{z^\alpha}, \frac{zg^\alpha}{zg^\alpha}\right). \]

The expected number of intra-links associated with a given vertex is the sum over the above terms and is given by

\[ L^\alpha = \frac{N^\alpha}{2} \left[ \gamma^\alpha - \left( \frac{Nz}{z^\alpha} - N^\alpha \right) \left( \psi^{(1)} \left( \frac{Nz}{z^\alpha} \right) - \psi^{(1)} \left( \frac{Nz}{z^\alpha} - \gamma^\alpha \right) \right) \right], \]

where \( \psi^{(n)}(x) \) is the first polygamma function defined as

\[ \psi^{(n)}(x) = (-1)^{n+1}n! \sum_{j=0}^{\infty} \frac{1}{(x+j)^{n+1}}. \]

Equation (8) leads to

\[ L^\alpha \approx \frac{z^\alpha}{2Nz} \left( N^\alpha \right)^2 \]

in the limit \( N \gg \gamma^\alpha \) when \( N^\alpha > z^\alpha \).

Alternatively, \( L^\alpha \) can be calculated by the probability that a chosen stub in class \( \alpha \) is not connected to another stub within the same class. The derivation is similar to the previous one. We define \( \tilde{N}_a^\alpha \) to be the number of vertices in the complement of class \( \alpha \) and \( \tilde{z}_a^\alpha \) as their average degree. Thus,

\[ L^\alpha = \frac{N^\alpha}{2} \left[ \gamma_a^\alpha - \tilde{\gamma}_a^\alpha + \left( \frac{Nz}{z_a^\alpha} - \tilde{N}_a^\alpha \right) \left( \psi^{(1)} \left( \frac{Nz}{z_a^\alpha} \right) - \psi^{(1)} \left( \frac{Nz}{z_a^\alpha} - \gamma_a^\alpha \right) \right) \right], \]

where \( \tilde{\gamma}_a^\alpha = \min(\tilde{N}_a^\alpha, \tilde{z}_a^\alpha) \).

### 3.2. Time evolution of class-specific average degrees

In the following derivation, we calculate the expected degree of a vertex in a given class in each generation, assuming that all vertices in the network are connected. This calculation is similar to the one presented in [18]. The average degree of vertices in a class changes over time since high-degree vertices are more likely to become infected during the course of an epidemic. As a result of this, the expected average degree of susceptible vertices is smaller than the expected average degree of infectious or removed vertices. Next, we derive the time evolution of the expected average degree of each class.

In each generation several vertices are infected simultaneously. Although we acknowledge this issue in our calculations presented here, in order to derive the equations of interest in a systematic manner, we have focused on one new infection at a time. We start with processing one infected...
case within each generation and continue the process by adding, one by one, other infected nodes produced during that generation, resulting in an expanding pool of ‘processed’ nodes, which hereafter we will refer to as the processed pool; similarly, we call the remainder of the infected nodes within that generation not yet processed in this manner as unprocessed pool.

We first define the two processed and unprocessed pools and then explain how the susceptible, infectious, and removed degree distributions can be calculated with knowledge of these quantities. We employ the following algorithm for calculating the degree distribution of the processed and unprocessed pools. In the beginning, all vertices belong to the unprocessed pool. We then randomly select a vertex from this pool in each step and assign it to the processed pool, as described below. Hereafter, we denote quantities related to the processed pool of individuals with a tilde (\(\sim\)) and the ones corresponding to the unprocessed pool without tilde. We use \(p_{k,j}\) to denote the degree distribution of the unprocessed vertices after \(j\) vertices have been processed. The degree distribution of the processed vertices is denoted by \(\tilde{p}_{k,j}\).

The first vertex is chosen randomly; thus, the probability that it has degree \(k\) is given by \(q_{k,1} = p_k\), the degree distribution of the network. The degree distribution of the unprocessed pool is \(p_{k,1} = p_k\). Next, we collect a second vertex by targeting a random stub of a vertex in the unprocessed pool. The probability that this second vertex has degree \(k\) is given by \(q_{k,2} = k\frac{p_k}{z_1}\). Therefore, the degree distribution of the processed pool is \(\tilde{p}_{k,2} = \frac{(q_{k,1} + q_{k,2})}{2}\) and the degree distribution of the unprocessed pool can be written as \(p_{k,2} = \frac{(Np_k - 2\tilde{p}_{k,2})}{(N - 2)}\). The probability that the \(j\)th chosen vertex has degree \(k\) is \(q_{k,j} = \frac{k\tilde{p}_{k,j}}{z_j}\). Similarly, the degree distribution of the processed and unprocessed vertices, respectively, are given by

\[
\tilde{p}_{k,j} = \frac{\sum_{l=1}^{j} q_{k,l}}{j},
\]

\[
p_{k,j} = \frac{Np_k - j\tilde{p}_{k,j}}{N - j}.
\]  

With \(\tilde{p}_{k,j}\) and \(p_{k,j}\) in place, we are able to calculate \(\bar{z}_j\) and \(z_j\). The expected average degree of susceptible, infectious, and removed classes at generation \(g\) can then be calculated based on \(\bar{z}_j\) and \(z_j\) by

- \(\bar{z}_s^g = \bar{z}N_s^g + N_r^g\),
- \(\bar{z}_i^g = \frac{(N_s^g + N_r^g)\bar{z} + N_s^g - N_s^g\bar{z}_r^g}{N_r^g}\),
- \(\bar{z}_r^g = \bar{z}N_r^g\).

### 3.3. Multi-targeting

We recall that the quantity \(L_{si}^g\) denotes the number of links between susceptible and infectious vertices. Links which led or are leading to an infection are referred to as transmitting links, whereas links that do not lead to infection are non-transmitting links. As an example, see Figure 1, where these different types of links are represented with different line styles. The expected number of transmitting and non-transmitting links between the susceptible and infected classes can be calculated by \(\lambda_s^g = TL_{si}^g\) and \(\lambda_{NT}^g = (1 - T)L_{si}^g\), respectively. This implicitly assumes that susceptibility and infectiousness are independent.

In a finite-size network, some of the transmitting links may target the same susceptible individual, such as the vertex marked with a \(Y\) in Figure 1. If a susceptible individual is infected by more
than one infectious vertex, we refer to this process as multi-targeting. The effect of multi-targeting is important when calculating the number of new infections using $\lambda$. In Equations (12)–(16), we assume that all susceptible vertices have the same degree, which is a good approximation when the degree distribution is narrow; we will relax this constraint at the end of this section. The first randomly chosen transmitting link has a contribution of one to the number of new infections. The probability that the second randomly chosen transmitting link is attached to a different susceptible vertex is roughly proportional to $\frac{N_s^g - 1}{N_s^g}$. (12)

This in fact is the contribution of the second link to the expected number of new infections. The contribution of the $j$th link is then given by

$$
N_s^g - n_{j-1} = \frac{N_s^g - n_{j-1}}{N_s^g},
$$

where $n_{j-1}$ is the expected number of new infections after having processed $(j - 1)$ links. The expected number of the new infections after using $j$ transmitting links is

$$
n_j = n_{j-1} + \frac{N_s^g - n_{j-1}}{N_s^g}.
$$

This is a relation between the sum of $j$th and $(j - 1)$th terms of a geometric series. The number of new infections after all transmitting links have been used is then given by

$$
N_{i+1} = 1 - \frac{(\beta_s^g)^{\lambda_s}}{1 - \beta_s^g},
$$

where $\beta_s^g = 1 - (1/N_s^g)$. Therefore, it is straightforward to show that

$$
N_i^{g+1} = \lambda_s - \frac{\lambda_s (\lambda_s - 1)}{2N_s^g},
$$

where $N_s^g \gg \lambda_s$. In this limit, the first term refers to the number of transmitting links, whereas the second term adjusts for multi-targeting.

In deriving Equation (15), we assumed that all vertices have the same degree. Relaxing this assumption leads us to the following equation:

$$
N_i^{g+1} = \sum_k N_i^{g+1}(k),
$$

where $N_i^{g+1}(k)$ is the number of new infections of degree $k$ at generation $g + 1$. The number of new infectious vertices of degree $k$ can be calculated by extending Equation (16) in the following way:

$$
N_i^{g+1}(k) = \begin{cases} 
\frac{1 - (\beta_k^g)^{\lambda_k^g}}{1 - \beta_k^g}, & \lambda_k^g > 1 \\
\lambda_k^g, & \lambda_k^g < 1 
\end{cases},
$$

where $\beta_k^g = 1 - (1/N_s^g p_k/n_i^g)$ and $\lambda_k^g = \lambda k p_k/n_i^g$. Note that for $\lambda_k^g < 1$ there is no need to discount for multi-targeting. Susceptible vertices with different degree have different contribution in multi-targeting. We divided these vertices into different degree groups. The number of vertices
in each degree group is given by $N_g^s p_{k,N-N_g^s}$, which explains the form of $\beta_g^s$. Moreover, those vertices with higher degree have a higher chance of being targeted. The chance is proportional to $k p_{k,N-N_g^s}/Z_g^s$, and this extra factor explains the form of $\lambda_g^s$.

The degree distribution $p_{k,N-N_g^s}$ in the above equation requires further adjustment that will be discussed in the next section.

3.4. Degree distribution adjustment

During an epidemic process, all stubs of a susceptible vertex might belong to non-transmitting links (such as the susceptible vertex marked with $X$ in Figure 1), where the links of the vertex are only connected to removed vertices. In this case, the vertex cannot be infected in further generations. In order to calculate the number of stubs from susceptible individuals that could carry future infections, the stubs belonging to non-transmitting links need to be excluded. This process leads to a so-called effective degree distribution that discounts for $j$ non-transmitting links and is denoted by $P_g^s(j)$. In this expression, the argument $j$ refers to the number of links, not vertices, as is the case with $p_k$. We also define the effective average degree as $Z_g^s(j) = \sum_k P_g^s(j) k$.

This degree distribution carries the information about stubs that can lead to infection and is derived below.

The number of non-transmitting links pointing to susceptible vertices is denoted by $\Lambda_g^s = \lambda_g^{NT} + L_g^{sr}$, where $\lambda_g^{NT}$ is the number of non-transmitting links in the current generation, and $L_g^{sr}$ is the number of non-transmitting links accumulated in previous generations.

We define $P_g^s(0) = p_{k,N-N_g^s}$ as the effective degree distribution of susceptible individuals before removing any non-transmitting links. The effective degree distribution of susceptible individuals after removing one stub belonging to a non-transmitting link is $P_g^s(1)$, which can be calculated by

$$N_g^s P_g^s(1) = N_g^s P_g^s(0) + \frac{P_{k+1}^s(0)(k+1)}{Z_g^s(0)} - \frac{P_g^s(0)k}{Z_g^s(0)}.$$ (19)

The second term reflects the contribution of vertices whose effective degree used to be $k+1$, but now have effective degree of $k$, whereas the third term reflects the vertices whose effective degree used to be $k$, but now have degree $k - 1$.

We can therefore show that

$$N_g^s P_g^s(\Lambda_g^s) = N_g^s P_g^s(0) + \sum_{j=0}^{\Lambda_g^s-1} \left[ \frac{P_{k+1}^s(j)(k+1)}{Z_g^s(j)} - \frac{P_g^s(j)k}{Z_g^s(j)} \right].$$ (20)

is the effective degree distribution of the susceptible class after removing all stubs belonging to non-transmitting links. This effective degree distribution should be incorporated into Equation (17) in order to make the degree distribution adjustment.

4. Results

In this section, we validate the analytical results derived in the previous section by comparing them to simulation results. We present results for the number of intra-links, the time evolution of the average degree, and the number of infectious and removed vertices.

We use two types of networks, namely Poisson networks, $p_k = z^k e^{-z}/k!$ (with $z = 10$) and exponential networks, $p_k = (1 - e^{-1/\kappa})e^{-k/\kappa}$ (with $\kappa = 10$). We show the results for networks with
size $N = 1000$ and perform 1000 simulations to derive an average value for each epidemic quantity. The low number of vertices increases the likelihood of amplifying any potential discrepancies between simulation and analytical results due to network finite-size effect [18].

Figure 2 shows the average number of intra-links for the removed (a) and infectious (b) classes in terms of generation time for Poisson networks with $N = 1000$ and $T = 0.3$. The green line refers to the simulation outcomes. The red circles and blue triangles are the outputs of Equations (8) and (10), respectively (for these calculations the number of vertices and average degree were taken from simulation output). The pink squares are the result of Equation (8) (the number of vertices taken from the simulation output and average degree calculated using equations in Section 3.2). Figure 3 shows the same result for exponential networks. The small deviation of the average number of intra-links for the infectious class is the result of assuming that all vertices have the same degree.

Figure 2. Poisson network: average number of intra-links for the infectious (a) and removed (b) classes in terms of generation time for $T = 0.3$ and $N = 1000$. The green line refers to the simulation outcomes. The red circles and blue triangles are the outputs of Equations (8) and (10), respectively, where the average degree were taken from the simulation output. Equation (8) gives the number of intra-links using the probability that a given stub is attached to another stub in the same class, whereas the number of intra-links in Equation (10) is calculated by using the probability that a given stub is not attached to any other stub within the same class. The pink squares are also the result of Equation (8), but this time the average degree is calculated using equations in Section 3.2 (Colour online).

Figure 3. Exponential network: average number of intra-links for the infectious (a) and removed (b) classes in terms of generation time for $T = 0.3$ and $N = 1000$. The green line refers to the simulation outcomes. The red circles and blue triangles are the outputs of Equations (8) and (10), respectively, where the average degree were taken from the simulation output. Equation (8) gives the number of intra-links using the probability that a given stub is attached to another stub in the same class, whereas the number of intra-links in Equation (10) is calculated by using the probability that a given stub is not attached to any other stub within the same class. The pink squares are also the result of Equation (8), but this time the average degree is calculated using equations in Section 3.2 (Colour online).
The agreement between the results shown with green line, triangles, and circles confirms the validity of Equations (8) and (10). In Figure 3, the deviation of the pink squares from the other analytical results in generation 6 is due to the approximation of the average degree derived in Section 3.2. The deviation, albeit minor, is due to the abundance of infectious vertices during generation 6, which happens to be the timing of the epidemic peak in this example. We will discuss this issue in more detail below.

We depict the average degree of removed vertices in terms of generation time in Figure 4. The red solid (Poisson) and blue dashed (exponential) lines represent simulation results, and the circles (Poisson) and squares (exponential) represent analytical results. The significant change in the average degree of removed vertices corresponding to the exponential degree distribution can be attributed to the large width of this distribution.

In Figure 5, we depict the number of new infections and the number of transmitting links for exponential (a) and Poisson (b) degree distributions using the same parameter values as before. The squares represent the number of new infections calculated by Equation (15) using the simulation result for the number of infections in previous generation. The red line represents the simulation result.

In the exponential network, the slight difference between the analytical (green squares) and simulation results is partly due to the asymmetry of the degree distribution, since the majority of the stubs belongs to vertices with degrees higher than the average degree (70% of the stubs belong to 35% of the vertices). However, this can be corrected using Equation (17), which accounts for the multi-targeting effect for asymmetric degree distributions. The result with corrections due to multi-targeting is shown by the pink triangles, which is in excellent agreement with simulation results.

In the Poisson network, the difference between the two analytical results (Equations (15) and (17)) is very small due to the symmetry and small width of its degree distribution. These equations count the number of transmitting links. The main difference between Equations (15) and (17) is that we assume all vertices have the same degree in Equation (15), whereas this assumption is adjusted in Equation (17).
Figure 5. Number of new infections for the exponential (a) and Poisson (b) networks for $T = 0.3$ and $N = 1000$. The squares and triangles correspond to the number of the new infections calculated by the analytical formulas. The red line corresponds to simulation results, while the blue circles correspond to the average number of transmitting links. The squares represent the number of new infections calculated by Equation (15). The result with corrections due to multi-targeting is shown by the pink triangles (Equation (17)). The effect of non-transmitting links is calculated by Equation (20) and is shown with cyan stars (Colour online).

The effect of non-transmitting links ($L_{g}^{r}$ and $\lambda_{g}^{r}$) can be important when the average degree is low, or when the majority of vertices have a low degree. This effect is considered in Equation (20) and depicted by the cyan stars. The effect of Equation (20) is even more pronounced for lower transmissibility values, in which case the number of non-transmitting links is much higher (results not shown).

The effect of multi-targeting is shown by comparing the average number of transmitting links (blue circles) with the number of new infections (cyan triangles).

Up to this point we have validated our analytical results by comparing them to the results of simulations. Now we are in a position to analytically calculate the number of removed vertices in an algorithmic manner without using any simulation results as input. Our analytical calculation starts with one infection, that is, $N_{0}^{i} = N - 1$, $N_{0}^{r} = 1$, $N_{0}^{0} = 0$, and $L_{0}^{0} = z$. Using these initial conditions, we can calculate the number of transmitting links, $\lambda^{0}$ ($\lambda^{0} = TL_{g}^{i}$), and the number of new infections in generation, $N_{1}^{i}$, that resulted from these links (Equation (15)). This procedure is continued until the end of the epidemic. In Figure 6, we show the average number of removed vertices as a function of generation time for the exponential (a) and Poisson (b) networks ($T = 0.15, 0.2$, and $0.3$; dotted green, dashed red, and solid blue, respectively). The lines and symbols represent the simulation and analytical results, respectively (Colour online).
vertices as a function of generation time for the exponential (a) and Poisson (b) networks for $T = 0.15, 0.2, \text{ and } 0.3$ (green dotted, red dashed, and blue solid, respectively).

Due to stochasticity in the initial phase, epidemics enter the exponential phase at different generations. As far as computer simulations are concerned, this leads to a shift in the timing of epidemic peak from one simulation to another. As a result, when one curve reaches its peak, another may reach the middle of its incline or decline phase. Due to this limitation inherent in simulating an epidemic process, the curves of epidemic peaks become lower and wider than analytically derived values. This explains the discrepancy between the simulation results (solid lines) and the corresponding analytical results (symbols) in Figure 6.

5. Conclusion

In this paper, we presented a new theoretical framework that studies the spread of disease in terms of disease generation time. This network-based framework enables us to calculate in detail three main values that are important for modelling the time evolution of disease spread. Firstly, this framework allows us to calculate ‘transmitting links’. These are links (contacts) within the network that are capable of producing new infectious nodes (individuals). Secondly, this framework also allows us to calculate ‘non-transmitting links’, links that represent contacts that do not result in infection transmission. Finally, this framework enables us to calculate the degree distribution (the distribution of the number of contacts that individuals have with their neighbours) for each epidemiological class. Using several examples, we demonstrated a very good agreement between our analytical calculations and results of computer simulations. The proposed method enables us to calculate physical quantities such as $L^{\alpha \beta}_{g}$ − that is, the number of links within a network between classes $\alpha$ and $\beta$ during generation $g$ − in an efficient manner, whereas collecting this information during multiple runs of a typical numerical epidemic simulation on the same network generally takes much longer time. This advantage becomes more noticeable as the size of the network increases.

To continue this avenue of research in the future, the proposed framework could be extended to account for transmission processes in continuous time. Another interesting approach would be to extend this framework to study the impact of network clustering on the spread of disease. Susceptible–infected–removed–susceptible models can also be studied using a similar approach. Finally, investigations that include variability in network structure (i.e. dynamic networks), in addition to the dynamics of epidemic processes on the network, are other approaches that would be interesting and relevant to the modelling of epidemic processes.

Acknowledgements

B.P. would like to acknowledge the support of the Canadian Institutes of Health Research (Grant No. PTL-97126), the Michael Smith Foundation for Health Research (Grant No. CI-SSH-00931(06-1)), and the British Columbia Ministry of Health (Pandemic Preparedness Modeling Project).

References

[1] R.M. Anderson and R.M. May, *Infectious Disease of Humans*, Oxford University Press, Oxford, 1991.
[2] N.T.J. Bailey, *The Mathematical Theory of Infectious Disease and Its applications*, Hafner Press, New York, 1975.
[3] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, *Dynamical patterns of epidemic outbreaks in complex heterogeneous networks*, J. Theor. Biol. 235 (2005), pp. 275–288.
[4] F. Brauer, P. van den Driessche, and J. Wu (eds.), *Mathematical Epidemiology*, Lecture Notes in Mathematics, Mathematical Biosciences Subseries 1945, Springer-Verlag Berlin Heidelberg, 2008.
[5] B. Davoudi, J.C. Miller, R. Meza, L.A. Meyers, D.J.D. Earn, and B. Pourbohloul, *Early real-time estimation of the basic reproduction number of emerging infectious diseases*, Phys. Rev. X 2 (2012), 031005 (15 pages).

[6] L.K. Gallos and P. Argyakis, *Distribution of infected mass in disease spreading in scale-free networks*, Phys. A 330 (2003), pp. 117–123.

[7] G. Hartvigsen, J.M. Dresch, A.L. Zielinski, A.J. Macla, and C.C. Leary, *Network structure, and vaccination strategy and effort interact to affect the dynamics of influenza epidemics*, J. Theor. Biol. 246 (2007), pp. 205–213.

[8] H.W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev. 42 (2000), pp. 599–653.

[9] E. Kenah and J.M. Robins, *Second look at the spread of epidemics on networks*, Phys. Rev. E 76 (2007), 036113 (12 pages).

[10] M. Kuperman and G. Abramson, *Small world effect in an epidemiological model*, Phys. Rev. Lett. 86 (2001), pp. 2909–2912.

[11] M. Marder, *Dynamics of epidemics on random networks*, Phys. Rev. E 75 (2007), 066103 (5 pages).

[12] L.A. Meyers, M.E.J. Newman, and B. Pourbohloul, *Predicting epidemics on directed contact networks*, J. Theor. Biol. 240 (2006), pp. 400–418.

[13] J.C. Miller, *Percolation and epidemics in random clustered networks*, Phys. Rev. E 76 (2007), 010101 (4 pages).

[14] C. Moore and M.E.J. Newman, *Epidemics and percolation in small-world networks*, Phys. Rev. E 61 (2000), pp. 5678–5682.

[15] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, *Epidemic outbreaks in complex heterogeneous networks*, Eur. Phys. J. B 26 (2002), pp. 521–529.

[16] M.E.J. Newman, *Spread of epidemic disease on networks*, Phys. Rev. E 66 (2002), 016128 (11 pages).

[17] M.E.J. Newman, *Properties of highly clustered networks*, Phys. Rev. E 68 (2003), 026121 (6 pages).

[18] P.-A. Noël, B. Davoudi, L.J. Dubé, and B. Pourbohloul, *Time evolution of epidemic disease on finite and infinite networks*, Phys. Rev. E 79 (2008), 026101 (14 pages).

[19] R. Pastor-Satorras and A. Vespignani, *Epidemic spreading in scale-free nets*, Phys. Rev. Lett. 86 (2001), pp. 3200–3203.

[20] R. Pastor-Satorras and A. Vespignani, *Epidemic dynamics and endemic states in complex networks*, Phys. Rev. E 63 (2001), 066117 (9 pages).

[21] H. Rahmandad and J. Sterman, *Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models*, Manag. Sci. 54 (2008), pp. 998–1014.

[22] T.C. Schelling, *Micromotives and Macrobehavior*, Norton, New York, 1978.

[23] G. Szabó, *Branching annihilating random walk on random regular graphs*, Phys. Rev. E 62 (2000) pp. 7474–7477.

[24] E. Volz, *SIR dynamics in random networks with heterogeneous connectivity*, J. Math. Biol. 56 (2008), pp. 293–310.

[25] E. Volz and L.A. Meyers, *Susceptible, infected, recovered epidemics in dynamic contact networks*, Proc. R. Soc. Ser. B 274 (2007), pp. 2925–2934.

[26] C.P. Warren, L.M. Sander, and I. Sokolov, *Epidemics, disorder, and percolation*, Phys. A: Statist. Mech. Appl. 325 (2003), pp. 1–8.

[27] C.P. Warren, L.M. Sander, I. Sokolov, C. Simon, and J. Koopman, *Continuum description of a contact infection spread in a SIR model*, Math. Biosci. 208 (2007), pp. 205–215.