Analytical Model of COVID-19 for lifting non-pharmaceutical interventions

Garry Jacyna, James R. Thompson*, Matt Koehler, and David M. Slater

The MITRE Corporation: 7515 Colshire Drive, McLean, VA 22102

Abstract

In the present work, we outline a set of coarse-grain analytical models that can be used by decision-makers to bound the potential impact of the COVID-19 pandemic on specific communities with known or estimated social contact structure and to assess the effects of various non-pharmaceutical interventions on slowing the progression of disease spread. This work provides a multi-dimensional view of the problem by examining steady-state and dynamic disease spread using a network-based approach. In addition, Bayesian-based estimation procedures are used to provide a realistic assessment of the severity of outbreaks based on estimates of the average and instantaneous basic reproduction number R_0.

Keywords: COVID-19, Epidemiology, Mathematical model, Contact Network, Intervention

1. Introduction

Network structure plays an important role in disease spread. Most network models assume full mixing where all individuals are equally likely to become infected. However, most real-world networks have vertex degree distributions that are highly nonuniform. In the sections that follow, we derive a mathematical framework for determining both the steady-state and dynamic disease spread on complex networks using the concept of bond percolation. The analysis borrows liberally from the work of Pastor-Satorras and Vespignani [1] and Newman [2, 3, 4, 5].

Percolation theory is easily explained through analogy. Consider an old fashion coffee percolator consisting of two glass flasks, the lower flask for heating the water and the upper flask

*corresponding author: jrthompson@mitre.org

for holding the coffee grounds. As water is heated in the lower flask, it begins to ascend through the tube connected to the upper flask. Initially, the water penetrates only a small portion of the coffee grounds. Upon further heating, the water penetrates more of the coffee grounds. At some point, there is an abrupt transition where all of the grounds are saturated and convective mixing occurs. Percolation implies the existence of a long path connecting points separated by a distance on the order of the network size (in this case, the layer of coffee grounds). For a pandemic, this long path connects individuals together into a large cluster. It turns out that percolation is a critical phenomenon; that is, the onset of percolation occurs rapidly.

Most networks of sufficient complexity undergo phase transitions, where small components (outbreaks) suddenly coalesce into a giant component (pandemic) that extends across the entire network when one or more critical parameters, such as disease transmissibility are exceeded. Mean-field theory, a branch of statistical mechanics used to analyze physical systems with multiple components, can be used to characterize these regions [6]. The main idea is to replace all interactions on a component with an average or effective interaction. Insights into the behavior of a system can, therefore, be obtained at relatively low computational cost.

The flow chart in Fig. 1 outlines two analytical models suitable for describing the steady-state and time-dependent (dynamic) properties of disease progression on a social network. We present a third model that is used for estimating critical parameters (such as the basic reproduction number R_0) from empirical outbreak case data. The steady-state model uses the theory of bond percolation to predict the outbreak size distribution prior to a pandemic, the size of the pandemic in terms of the proportion of affected individuals, and the risk of individual infection based on an individual's contact network. We consider two types of non-pharmaceutical Interventions (NPIs) – uniform social distancing where a randomly selected portion of the population is sequestered and directed social distancing where individuals with the largest number of contacts are sequestered (which effectively targets super-spreaders). The dynamic model uses a degree-based approximation to the Susceptible, Infectious, or Recovered (SIR) model based on an approach outlined by Barthélémy and Pastor-Satorras [7]. We generalized the model to examine the effects of uniform social distancing, testing, and contact tracing on the proportion of susceptible individuals, infected individuals, outbreak cases, and the basic

reproduction number (R_0) as a function of time. In addition, assessments of various NPIs, testing, and contact tracing procedures can be determined in near real-time for both analytical and empirical networks derived from census data.

The third model uses the degree-based SIR model and Bayesian estimation procedures to determine the average basic reproduction number (R_0) from tabulated outbreak cases at a state, county, and city-wide level. Additionally, we use a particle-based filter to determine the instantaneous R_0 over time. Together, these estimates can be used to predict the proportional number of infections and outbreak cases over time. The analytical models shown in Fig. 1 provide an overview of the mathematical framework derived in the present work. In the materials and methods section we first outline the steady-state disease spread on social contact networks as a function of the degree distribution of the community and the transmissibility of the disease. We then present the extension of the framework to capture the dynamic properties of disease spread and the implementation and lifting of NPIs. Next we introduce the Baysian-based estimation procedure for inferring network structure from empirical case data. In the results section we present examples for three different degree distributions and illustrate the different risk measures that can be derived from the mathematical framework. We also show how the Bayesian-based model can be used to predict cases and R_0 for New York state using cases data collected by Johns Hopkins University [8]. We close with a brief discussion of the results and potential application of the framework.

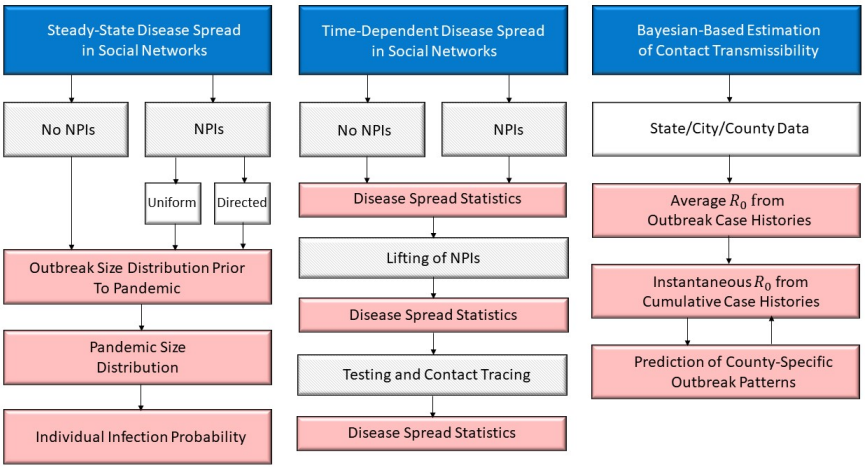

Figure 1. Process Flow Diagram Detailing the Three Analytical Models Used in this Study

2. Materials and methods

Throughout this section we use precise definitions of certain properties of epidemiological models based on Meyers et al., Bettencourt et al. and Chowell et al. [9, 10, 11].

- **Transmissibility** T is the average probability that an infectious individual will transmit the disease to a susceptible individual with whom they have contact
- **Critical Transmissibility** T_c is the minimum transmissibility required for an outbreak to become a pandemic. $T_c = \frac{\langle k \rangle}{(\langle k^2 \rangle - \langle k \rangle)}$ where $\langle k \rangle$ and $\langle k^2 \rangle$ are the mean and variance of the degree distribution of the contact network.
- **Basic Reproduction Number** R_0 is the expected number of cases directly generated by one case in the population of susceptible individuals. It can be shown to be equal to the ratio of transmissibility to the critical transmissibility $R_0 = T/T_c$.
- **The Instantaneous Reproduction Number** $R(t) = R_0 s(t)/N(t)$ where $s(t)$ is the number of susceptible individuals at time t and $N(t)$ is the total population.

Note that some studies refer to an effective reproductive number R_e. In the definitions above, R_0 depends on the degree distribution so there is no need to make this distinction. R_0 and R_e can be thought of as interchangeable terms.

2.1. Steady-State Disease Spread in Social Networks

Here we outline the procedure for characterizing steady-state disease spread on complex networks using bond percolation. The analysis is based on the work of Pastor-Satorras and Vespignani [1] and Newman [2, 3, 4, 5]. Generalizations to the theory are made to include both uniform and directed social distancing.

Given the transmission rate $r_{i,j}$ between node i and node j of a network graph and the infection time τ, the transmissibility is:

$$T = 1 - (1 - r_{i,j}\delta t)^{\tau/\delta t} \to 1 - e^{-r_{i,j}\tau}, \tag{1}$$

as $\delta t \to 0$. Typically, $r_{i,j} = r$, where r and τ are independent random variables. The average transmissibility is then:

$$T = 1 - \iint e^{-r\tau} P_r(r) P_\tau(\tau) dr\, d\tau, \tag{2}$$

where $P_r(r)$ and $P_\tau(\tau)$ are the respective probability density functions (pdfs). For simplicity, it is assumed that $P_r(r) = \delta(r - r_0)$ and $P_\tau(\tau) = \delta(\tau - \tau_0)$ so that $T = 1 - e^{-r_0\tau_0}$.

For a randomly chosen vertex, let p_k denote the probability that this vertex has k edges. Define $G_0(x)$ as the generating function for the degree distribution of this vertex:

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k. \tag{3}$$

This function is similar to the characteristic function; that is, given a sum of N independent and identically-distributed (i. i. d.) random variables, the generating function is $G_0^N(x)$.

Three types of social networks are examined. The Erdös-Renyi network has a Poisson degree distribution of the form:

$$p_k = \lambda^k e^{-\lambda}/k!, \tag{4}$$

where λ is the mean. The exponential network is used as a proxy for an urban network and has a degree distribution of the form:

$$p_k = (1 - e^{-\beta})e^{-\beta k}, \tag{5}$$

with parameter β. The power-law (Barabási-Albert) network has a majority of small degree links with a small minority of large degree links representing super-spreaders and a degree distribution of the form:

$$p_k = e^{-k/\kappa}/k^\alpha Li_\alpha(e^{-1/\kappa}), \tag{6}$$

where α and κ are parameters and $Li_\alpha(x) = \sum_{k=1}^{\infty} x^k/k^\alpha$ is the poly-logarithm function.

An important result by Feld is that the degree distribution of the first neighbor of a vertex is not the same as the degree distribution of vertices as a whole [12]. There is a greater chance that an edge will be connected to a vertex of high degree, in fact, in direct proportion to its degree. Let q_k denote the degree distribution of a vertex at the end of a randomly chosen edge. Then:

$$q_{k-1} = kp_k/z, \tag{7}$$

excluding the randomly-chosen edge, where $z = \sum_k kp_k$. The corresponding generating function for this distribution is:

$$G_1(x) = \sum_{k=0}^{\infty} q_k x^k = G_0'(x)/z, \tag{8}$$

where $G_0'(x)$ is the derivative of $G_0(x)$.

The transmissibility along each edge is taken into account by interpreting T in Eq. (2) as a probability ($0 \leq T \leq 1$). The probability that m out k edges is active is binomially distributed, so:

$$G_0(x;T) = \sum_{k=0}^{\infty} p_k(1-T+xT)^k = G_0(1-T+xT), \qquad (9)$$

where, similarly, $G_1(x;T) = G_1(1-T+xT)$.

2.1.1. Bond Percolation

Most complex networks experience phase transitions where small components suddenly coalesce into a giant component that extends across the entire network. The phase transitions of water as a function of temperature and pressure are examples. These regions can be described using mean-field theory and the generating functions $G_0(x)$ and $G_1(x)$. In order to apply mean field theory, it must be assumed that any finite component of connected vertices has no closed loops [5]. It can be shown that the probability of closed loops is on the order of $\mathcal{O}(1/n)$, where n is the network size. As $n \to \infty$, this means that all finite components have a tree-like (branching) structure.

Small outbreaks (percolation clusters) can be characterized as follows. Let $H_1(x)$ denote the generating function of the size distribution of the clusters at the end of the randomly chosen edge. Referring to the diagram in Fig. 2, the aggregate size of a cluster is the sum of all the clusters emanating from each vertex. This is a correct interpretation because there are no closed loops. For a vertex forming two clusters, the generating function of the size distribution is $H_1^2(x)$, since two i. i. d. clusters are summed together.

Similarly, for a vertex forming n clusters, the generating function of the size distribution is $H_1^n(x)$. Therefore, in the limit of large network size:

$$H_1(x) = x\sum_{k=0}^{\infty} q_k H_1^k(x) = xG_1(H_1(x)), \qquad (10)$$

where x pre-multiplying the above expression is the originating vertex and $G_1(x)$ is as defined in Eq. (8) [3]. Similarly, defining $H_0(x)$ as the generating function of the size distribution of clusters for a randomly chosen vertex, then:

$$H_0(x) = x\sum_{k=0}^{\infty} p_k H_1^k(x) = xG_0(H_1(x)). \qquad (11)$$

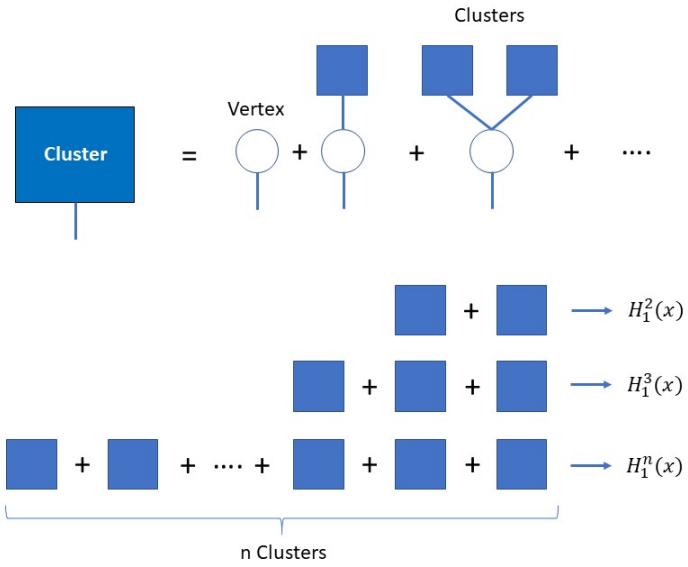

Figure 2. Computing the size of percolation clusters using mean-field theory

Modifying the above equations to include the transmissibility T:

$$\begin{aligned} H_1(x;T) &= xG_1(H_1(x;T);T) \\ H_0(x;T) &= xG_0(H_1(x;T);T). \end{aligned} \tag{12}$$

Equation (12) defines the bond percolation process and the analyses to follow.

2.1.2. Small Outbreaks

The critical transmissibility T_c leading to a pandemic (giant component) can be determined by computing the average size of small outbreaks (small components not associated with the pandemic). Since $H_0(x;T) = \sum_k p_k^s x^k$, where p_k^s is the cluster size probability, the average size of the cluster is $\langle c \rangle = \sum_k k p_k^s = H_0^{'}(1;T) = H_0^{'}(1)$. Newman shows that [2]:

$$\langle c \rangle = H_0^{'}(1) = 1 + TG_0^{'}(1)/(1 - TG_1^{'}(1)); \; 0 \leq T < T_c, \tag{13}$$

where $G_0(1) = G_1(1) = 1$. A phase transition occurs when $\langle c \rangle \to \infty$ or, from Eq. (13), when $T_c = 1/G_1^{'}(1)$ which implies:

$$T_c = \langle k \rangle \, / \left(\left\langle k^2 \right\rangle - \langle k \rangle \right). \tag{14}$$

The basic reproduction number R_0 can then be defined as $R_0 = T/T_c$, so $R_0 = 1$ when $T = T_c$.

In addition to the average outbreak size $\langle c \rangle$, the distribution of outbreak sizes can also be determined in a computationally efficient manner. Recalling that $H_0(x;T) = \sum_k p_k^s x^k$ and letting $x = e^{2\pi jm/M}$, the outbreak size distribution p_k is equivalent to the inverse discrete Fourier Transform (IDFT) of $H_0(e^{2pijm/M};T)$.

2.1.3. Pandemic Size

The size of the pandemic above the critical threshold T_c can also be determined. However, the giant component spans the entire network, so there have to be closed loops. This invalidates the tree-like assumption that led to Eq. (12). As Newman points out, the problem can be approached indirectly [5]. Recall that $H_0(1;T) = \sum_k p_k^s$ is the fraction of components *not* in the giant component. This implies that $P(T) = 1 - H_0(1;T)$ is the probability of a pandemic forming. Defining $v = H_1(1;T)$ and using Eq. (12):

$$\begin{aligned} v &= G_1(v;T) \\ P &= 1 - G_0(v;T). \end{aligned} \tag{15}$$

This is a fixed point problem. The average outbreak size not associated with the pandemic can also be determined. Repeating the derivation leading to Eq. (13) and noting that $H_0(1;T) = 1 - P(T)$, Newman shows that [3]:

$$\langle c \rangle = 1 + \frac{Tzv^2(T)}{(1-P(T))(1-TG_1'(v;T))}; \; T > T_c. \tag{16}$$

The risk of an individual infection is determined by noting that v in Eq. (15) is the probability that a node along a randomly chosen edge is not infected. For each edge, the probability of not getting infected is either v (contact not infected) or $(1-T)(1-v)$ (contact is infected but does not transmit the infection). Thus, $p(T) = v + (1-T)(1-v) = 1 - T - vT$. If the individual has k contacts, then the risk of infection is [2]:

$$\text{Risk}(k;T) = 1 - p^k(T). \tag{17}$$

2.1.4. Inclusion of NPIs

The known methodology outlined above is generalized to include social distancing. Specific applications including uniform and directed social distancing are then discussed. Let b_k denote the probability that a vertex of degree k is present. Define the generating function $F_0(x)$ as:

$$F_0(x) = \sum_{k=0}^{\infty} b_k p_k x^k. \tag{18}$$

Similarly, $F_1(x) = F_0^{'}(x)/z$. This is a generalization of the generating functions $G_0(x)$ and $G_1(x)$. Since $F_0(1) \neq 1$ and $F_1(1) \neq 1$, $1 - F_0(1)$ is the probability that a randomly chosen vertex has no active edges. From Eq. (12), it follow that [3]:

$$\begin{aligned} H_1(x;T) &= 1 - F_1(1) + xF_1(H_1(x;T);T) \\ H_0(x;T) &= 1 - F_0(1) + xF_0(H_1(x;T);T). \end{aligned} \tag{19}$$

The average outbreak size prior to a pandemic with social distancing is similar to Eq. (13):

$$\langle c \rangle = 1 + TF_0^{'}(1)/(1 - TF_1^{'}(1)). \tag{20}$$

A phase transition leading to a pandemic occurs at the critical transmissibility $T_c = 1/F_1^{'}(1)$, which follows directly from Eq. (20). Similarly, the outbreak size distribution follows directly from $H_0(x;T)$ by defining $x = e^{2\pi jm/M}$ and using the inverse Fourier Transform.

The fraction of the population affected by the pandemic is similar to Eq. (15). Since $H_0(1;T) = 1 - P(T)$ and using Eq. (19):

$$\begin{aligned} P &= F_0(1) - F_0(v;T) \\ v &= 1 - F_1(1) + F_1(v;T). \end{aligned} \tag{21}$$

The average outbreak size not associated with the pandemic is analogous to Eq. (16) with $G_0(x)$ and $G_1(x)$ replaced by $F_0(x)$ and $F_1(x)$, respectively. Noting that $F_0(v;T) = F_0(1) - P$, $F_0^{'}(v;T) = zF_1(v;T)$, and $F_1(v;T) = v - 1 + F_1(1)$:

$$\langle c \rangle = \frac{F_0(1) - P(T)}{1 - P(T)} + \frac{Tz(v - 1 + F_1(1))^2}{(1 - P(T))(1 - TF_0^{'}(v;T))}. \tag{22}$$

The risk of an infection during a pandemic is identical to Eq. (17) with v given by Eq. (21).

2.1.5. Uniform Social Distancing

Let $b_k = b; 0 \leq b \leq 1$, so the probability that a vertex of degree k is active is b. Basically, social distancing is applied to every individual regardless of the number of contacts an individual may have. For this case, $F_0(x;T) = bG_0(x;T)$ and $F_1(x;T) = bG_1(x;T)$. The average outbreak size prior to a pandemic is identical to Eq. (13) with T replaced by $T_{\text{eff}} = bT$. This implies that the critical threshold with uniform social distancing is $T_c^{'} = T_c/b$. Since the basic reproduction number $R_0 = T/T_c$,

the effective reproduction number is $R_0^{'} = T/T_c^{'} = bR_0$. This means that the effective reproduction number $R_0^{'}$ decreases with increased uniform social distancing. The outbreak size distribution similarly follows from Eq. (19).

The fraction of the population affected by the pandemic follows directly from Eq.(21). Letting $v^{'} = (v-1+b)/b$ and noting that $G_1(v;T) = G_1(1-bT+bTv^{'})$:

$$\begin{aligned} P &= b(1-G_0(v^{'};T)) = bP(T_{\text{eff}}) \\ v^{'} &= G_1(v^{'};T_{\text{eff}}), \end{aligned} \tag{23}$$

where $T_{\text{eff}} = bT$. This is identical to Eq. (15) with no social distancing except that the onset of the pandemic is shifted in accordance with T_{eff} and the affected population is reduced by a factor of b. Outbreaks not associated with the pandemic follow from Eq. (22). In particular:

$$\langle c \rangle = \frac{b(1-P(T_{\text{eff}}))}{1-bP(T_{\text{eff}})} + \frac{T_{\text{eff}} z b v^{'2}}{(1-bP(T_{\text{eff}}))(1-T_{\text{eff}}G_0^{'}(v^{'};T_{\text{eff}}))}, \tag{24}$$

where $v^{'} = (v-1+b)/b$. Additionally, the risk of individual infection is given by Eq. (17) with $p = 1 - T_{\text{eff}} + T_{\text{eff}}v^{'}$.

2.1.6. Directed Social Distancing

Here, it is assumed that:

$$b_k = \begin{cases} 1; & 0 \le k \le K_{max} \\ 0; & k > K_{max}, \end{cases} \tag{25}$$

where all individuals are distanced with contact degree greater than K_{max}. This implies that:

$$F_0(x) = \sum_{k=0}^{K_{max}} p_k x^k, \tag{26}$$

and $F_1(x) = F_0^{'}(x)/z$. For a Poisson network:

$$\begin{aligned} F_0(x) &= e^{-z} \sum_{k=0}^{K_{max}} (zx)^k/k! \\ F_1(x) &= e^{-z} \sum_{k=0}^{K_{max}} (zx)^k/k!. \end{aligned} \tag{27}$$

For an exponential network:

$$\begin{aligned} F_0(x) &= (1-e^{-\beta})(1-(xe^{-\beta})^{K_{max}+1})/(1-xe^{-\beta}) \\ F_1(x) &= (1-e^{-\beta})^2(K_{max}(xe^{-\beta})^{K_{max}+1} \\ &\quad +1-(K_{max}+1)(xe^{-\beta})^{K_{max}})/(1-xe^{-\beta})^2. \end{aligned} \tag{28}$$

For a power-law network:

$$\begin{aligned} F_0(x) &= Li_\alpha(K_{max}, xe^{-1/\kappa})/Li_\alpha(e^{-1/\kappa}) \\ F_1(x) &= Li_{\alpha-1}(K_{max}, xe^{-1/\kappa})/xLi_{\alpha-1}(e^{-1/\kappa}). \end{aligned} \tag{29}$$

The average outbreak size prior to a pandemic is given by Eq. (20), where $F_0^{'}$ and $F_1^{'}$ can be computed by taking respective derivatives of Eqs. (27), (28), and (29). Additionally, the fraction of the population affected by the pandemic is given by Eq. (21) together with Eqs. (27)-(29). The fractional number (f_c) of nodes removed due to directed social distancing is related to K_{max} by:

$$f_c = 1 - \sum_{k=0}^{K_{max}} p_k = 1 - F_0(1). \tag{30}$$

It is interesting to note that for the power-law network, if a small fraction of nodes f_c is removed, then the critical transmissibility $T_c^{'} \gg T_c$, indicating a lack of phase transition or pandemic onset. This was discussed by Callaway and Newman in another context [4]. Both the average outbreak size not associated with the pandemic and the individual risk of infection are given by Eqs. (22) and (17), respectively.

2.1.7. Calibration

The degree distribution p_k for each network is calibrated to have the same critical transmissibility T_c. Since $T_c = 1/G_1^{'}(1)$ from Eq. (14), it follows that:

$$T_c = \begin{cases} 1/z; & \text{Poisson} \\ (e^\beta - 1)/2; & \text{Exponential} \\ Li_{\alpha-1}(e^{-1/\kappa})/(Li_{\alpha-2}(e^{-1/\kappa}) - Li_{\alpha-1}(e^{-1/\kappa})); & \text{Power-Law.} \end{cases} \tag{31}$$

2.2. Time-dependent Disease Spread in Social Networks

In this section we expand the procedure to characterize time-dependent disease spread on complex networks using a stochastic SIR model. The analysis is based on the work of Barthélémy and Pastor-Satorras [7]. Generalizations to the theory are made to include social distancing, testing, and contact tracing.

2.2.1. Introduction

A stochastic treatment of the time-dependent properties of epidemic models involves determining the probabilities for vertices to be in specific disease states. This is typically a difficult problem because it involves higher-order moments of probabilities that can only be approximated using moment-closure

techniques, where moments are factored into pairwise moment products [5]. An alternative approach is degree-based approximation pioneered by Pastor-Satorras et al. This approximation assumes that all vertices of the same degree have the same probability of infection at any given time.

Consider the probability that vertex A becomes infected between times t and $t + dt$. To become infected, it must catch the disease from one of its neighbors, which requires that the neighbor be infected. The probability of a neighbor being infected is x_k, where k is the excess degree of the neighbor (recall that the excess degree distribution is q_k). So, the average probability of a neighbor being infected is $v(t) = \sum_{k=0}^{\infty} q_k x_k$. The total probability of transmission from a single neighbor is $\beta v(t)dt$, where β is the contact rate. The probability of transmission from any neighbor is $\beta k v(t)dt$, where k is the number of neighbors associated with vertex A. Thus, the rate of change in the probability that a vertex with degree k is susceptible (s_k) is simply:

$$ds_k/dt = -\beta k v(t) s_k. \tag{32}$$

Similarly, the probability that a vertex with degree k is infected (x_k) is:

$$dx_k/dt = \beta k v(t) s_k - \gamma x_k, \tag{33}$$

where $\gamma = 1/T_r$ is the recovery rate. Finally, the probability that a vertex of degree k has recovered (r_k) is:

$$dr_k/dt = \gamma x_k. \tag{34}$$

A general solution to Eqs. (32)-(34) is of the form [5]:

$$\begin{aligned} du/dt &= -\beta u v(t) \\ v(t) &= 1 + (\gamma/\beta)\log(u) - s_0 G_1(u), \end{aligned} \tag{35}$$

where $s_0 = s_k(0)$. Given $u(0) = 1$, Eq. (35) can be numerically solved for $u(t)$ and the average probability of infection $v(t)$ can be determined. In addition, the average susceptibility probability $s(t)$ can be computed from the degree distribution of a randomly chosen vertex:

$$s(t) = \sum_{k=0}^{\infty} p_k s_k = s_0 G_0(u), \tag{36}$$

where $G_0(x)$ is the moment generating function for p_k.

The function $u(t)$ in Eq. (35) has a particularly interesting interpretation. A fixed point occurs when $du/dt = 0$ or when

$1 + (\gamma/\beta)\log(u) - s_0 G_1(u) = 0$. If it is assumed that $u \approx 1$, then:

$$\begin{aligned} u &= 1 - T + TG_1(u) \\ P &= 1 - G_0(u), \end{aligned} \tag{37}$$

where the pandemic size $P = w(\infty)$ and $w(t) = \sum_{k=0}^{\infty} q_k r_k$ is the average recovery probability. Assuming $\beta/\gamma \ll 1$, the transmissibility is approximately $T \approx \beta/\gamma$. If $u = 1 - T + Tv$, then Eq. (15) and Eq. (37) are identical under the above assumptions. Therefore, the dynamic SIR model is consistent with bond percolation results at steady state.

2.2.2. Uniform Social Distancing

It is assumed that no NPIs are active over a period $0 \leq t \leq t_1$. Over a period $t_1 < t \leq t_2$, uniform social distancing is implemented where a fraction $(1 - b)$ of the population is sequestered. Over the period $t_2 < t \leq t_3$, the NPI is lifted. The first period corresponds to the analysis outlined in the previous section.

For the second period that includes social distancing, the stochastic SIR equations for the fraction of the population (b) not sequestered are:

$$\begin{aligned} ds_k/dt &= -\beta k v^{'}(t) s_k \\ dx_k/dt &= \beta k v^{'}(t) s_k - \gamma x_k \\ dr_k/dt &= \gamma x_k, \end{aligned} \tag{38}$$

where $v^{'}(t) = \sum_k q_k^{'} x_k$ and $q_k^{'} = b q_k$. Letting $w^{'}(t) = \sum_k q_k^{'} r_k$, it follows that $dw^{'}/dt = \gamma v^{'}(t)$. Substituting this expression into Eq. (38) and simplifying:

$$s_k(t) = s_{k1} u^k(t), \tag{39}$$

where $u(t) = \exp(-\beta(w^{'} - w_0^{'})/\gamma)$ and s_{k1} are the fractional number of susceptible individuals with k contacts remaining at the end of the first period ($t = t_1$). Since $x_k = 1 - s_k - r_k$, $v^{'}(t) = b - b\sum_k q_k s_k - b\sum_k q_k r_k$. Now, $s_{k1} = s_0 u_1^k$, where $u_1 = u(t_1)$, so $\sum_k q_k s_k = s_0 G_1(u_1 u)$. Also, $w^{'}(t) = b\sum_k q_k r_k$. Using the expression for $u(t)$, $w^{'}(t) = w_0^{'} - (\gamma/\beta)\log(u)$, where $w_0^{'} = -b(\gamma/\beta)\log(u_1)$. Combining results:

$$v^{'}(t) = b - b s_0 G_1(u_1 u) + b(\gamma/\beta)\log(u_1) + (\gamma/\beta)\log(u). \tag{40}$$

Finally, since $dw^{'}/dt = -(\gamma/\beta u)du/dt$ and $dw^{'}/dt = \gamma v^{'}(t)$:

$$du/dt = -\beta u v^{'}(t). \tag{41}$$

The total average infection probability $v_T(t)$ needs to include the sequestered fraction $(1-b)$ of the population. However, the average infection probability among the sequestered population is $v_s^0 = (1-b)v(t_1)$, where $v(t_1)$ is determined from Eq. (35). In particular:

$$v_s^0 = (1-b)\left(1 + (\gamma/\beta)\log(u_0) - s_0 G_1(u_1)\right), \tag{42}$$

where $u_1 = u(t_1)$. These infected individuals recover at a rate γ since, by definition, they are not in contact with each other. So:

$$v_T(t) = v^{'}(t) + v_s^0 e^{-\gamma(t-t_1)};\ t_1 < t \leq t_2. \tag{43}$$

Note that at $t = t_1$, $v_T(t_1) = v(t_1)$, so continuity is maintained. Finally, the fractional number of susceptibles during the NPI period is the sum of the sequestered and non-sequestered fraction of the population:

$$s(t) = bs_0 G_0(uu_1) + (1-b)s_0 G_0(u_1). \tag{44}$$

For the third period when social distancing is lifted, the stochastic SIR equations are identical to Eqs. (32)-(34) except that the initial conditions are different. Let n_k denote the fraction of the population with k contacts remaining. However, $n_k = b + (1-b)s_{k1}$, i.e., the original non-sequestered fraction (b) and the sequestered fraction $((1-b)s_{k1})$, where $s_{k1} = s_0 u_1^k$.

Let $v''(t) = \sum_k q_k x_k$ denote the average infection probability, where $x_k = n_k - s_k - r_k$. Now, s_k has the form:

$$s_k = \hat{s}_{k2} u^k, \tag{45}$$

where $\hat{s}_{k2} = bs_{k2} + (1-b)s_{k1}$, $s_{k2} = s_{k1}u_2^k$, and $u_2 = u(t_2)$. The various summations entering into the calculation of $v^{''}(t)$ are of the form:

$$\begin{aligned}
\sum_k q_k s_k &= bs_0 G_1(u_1 u_2 u) + (1-b)s_0 G_1(u_1 u) \\
\sum_k q_k n_k &= b + (1-b)s_0 G_1(u_1) \\
\sum_k q_k r_k &= w_0 - (\gamma/\beta)\log(u),
\end{aligned} \tag{46}$$

where $w_0 = -(\gamma/\beta)\log(u_2) - b(\gamma/\beta)\log(u_1)$. Combining results:

$$\begin{aligned}
v^{''}(t) &= b\left(1 - s_0 G_1(u_1 u_2 u)\right) + (1-b)s_0\left(G_1(u_1) - G_1(u_1 u)\right) \\
&\quad + (\gamma/\beta)\log(u_2 u) + b(\gamma/\beta)\log(u_1).
\end{aligned} \tag{47}$$

Note that $v^{''}(t_2) = v^{'}(t_2)$, as required. The fractional number of susceptibles is given by:

$$s(t) = bs_0 G_0(u_1 u_2 u) + (1-b)s_0 G_0(u_1 u). \tag{48}$$

2.2.3. Testing and Contact Tracing

For periods $t > t_3$, a fairly simple model is used to capture the relevant details of delayed testing and contact tracing based on a generalization of work by Young and Ruschel [13]. The modified stochastic SIR model is:

$$\begin{aligned} ds_k/dt &= -\beta k v^{'''}(t) s_k \\ dx_k/dt &= \beta k v^{'''}(t) s_k - \beta k p e^{-\gamma\eta} v^{'''}(t-\eta) s_k(t-\eta) - \gamma x_k \quad (49) \\ dr_k/dt &= \gamma x_k + \beta k p e^{-\gamma\eta} v^{'''}(t-\eta) s_k(t-\eta), \quad (50) \end{aligned}$$

where η is the delay in testing after becoming infectious and, as before, $v^{'''}(t) = \sum_k q_k x_k$. Rather than assuming that infectious individuals are tested, tracked and sequestered immediately with probability p, there is a delay η in their sequestration. Recovered individuals are not added back into the pool of susceptible individuals since it is assumed that they have developed immunity to the disease.

A general solution to Eq. (50) can be determined as follows. Integrating the first expression produces $s_k(t) = \hat{s}_{k3} u^k$, where:

$$u(t) = \exp\left(-\beta \int_{t_3}^{t} v^{'''}(s)\, ds\right), \quad (51)$$

$\hat{s}_{k3} = b s_{k3} + (1-b) s_{k2}$, $s_{k3} = s_{k2} u_3^k$, and $u_3 = u(t_3)$. From Eq. (51), it follows that:

$$du/dt = -\beta u v^{'''}(t). \quad (52)$$

Averaging the second expression in Eq. (50):

$$dv^{'''}/dt = \beta v^{'''}(t) \sum_k q_k k s_k(t) - \beta p e^{-\gamma\eta} v^{'''}(t-\eta) \sum_k q_k k s_k(t-\eta) - \gamma v^{'''}(t). \quad (53)$$

Noting that $G_1'(x) = \sum_k q_k k x^{k-1}$:

$$\sum_k q_k k s_k = u(t) \left(b G_1'(u_1 u_2 u_3 u) + (1-b) G_1'(u_1 u_2 u) \right). \quad (54)$$

Therefore, Eq. (53) can be rewritten as:

$$dv^{'''}/dt = -\gamma v^{'''}(t) + \beta s_0 v^{'''}(t) u(t) \zeta(t) - \beta s_0 p e^{-\gamma\eta} v^{'''}(t-\eta) u(t-\eta) \zeta(t-\eta), \quad (55)$$

where $\zeta(t) = b G_1'(u_1 u_2 u_3 u(t)) + (1-b) G_1'(u_1 u_2 u(t))$. Equations (52) and (55) describe a set of consistent equations for numerically computing $v^{'''}(t)$ and $s(t)$, where:

$$s(t) = s_0 b G_0(u_1 u_2 u_3 u(t)) + s_0 (1-b) G_0(u_1 u_2 u(t)). \quad (56)$$

Note that Eq. (55) is initialized by setting $v^{'''}(t_3) = v^{''}(t_3)$.

2.3. Bayesian-based Estimation of Contact Transmissibility

We now outline methods for estimating the contact transmissibility from outbreak case data using a stochastic SIR model and Bayesian estimation methods. The first approach estimates the average transmissibility T or basic reproduction number R_0 using either analytical or empirical networks where the degree distribution is known or estimated. The second approach estimates the instantaneous transmissibility $T(t)$ or basic reproduction number $R_0(t)$ using nonlinear tracking methods based on particle-based filtering where explicit network structure is not required.

2.3.1. Average Transmissibility

The probability that a vertex of degree k will experience an outbreak is the probability that it gets infected and then recovers. From Eq. (33):

$$dc_k/t = dx_k/dt + dr_k/dt = \beta k v(t) s_k. \tag{57}$$

The average outbreak case probability $C(t)$ is determined by averaging Eq. (57):

$$dC(t)/dt = \beta v(t) u(t) G_0^{'}\left(u(t)\right), \tag{58}$$

where $u(t)$ is given by Eq. (35) and $\sum_k p_k k s_0 u^k = s_0 u(t) G_0^{'}(u)$.

Define $\Delta C(t+\xi) = C(t+\xi) - C(t)$ as the fractional change in outbreak cases in the interval $[t, t+\xi]$. Then, from Eq. (58):

$$\Delta C(t+\xi) \approx \beta \xi s_0 u(t) G_0^{'}\left(u(t)\right) v(t+\xi). \tag{59}$$

Now, returning to Eq. (33), averaging, and integrating the results from t to $t+\xi$, assuming that $\xi \ll 1$:

$$v(t+\xi) \approx v(t) b_\xi(t;T), \tag{60}$$

where $b_\xi(t;T) = \exp\left(\gamma\xi(s_0 T u(t) G_1^{'}(u(t)) - 1)\right)$ and $T = \beta/\gamma$. Finally, substituting this expression into Eq. (59):

$$\Delta C(t+\xi) = \Delta C(t) b_\xi(t;T). \tag{61}$$

This implies a linear relationship between the fractional change in outbreak cases between time steps ξ.

The change in the number of outbreak cases $\Delta N(t)$ is $\Delta N(t) = NC(t)$, where N is the population size. Assume that $\Delta N(t)$ is integer-valued and let the measurements comprise the collection

$\{\Delta N_1, \Delta N_2, \ldots, \Delta N_J\}$, where $\Delta N_i = \Delta N(t_i)$. The resulting likelihood function can be factored as:

$$p(\Delta N_1, \Delta N_2, \ldots, \Delta N_J|T) = \prod_{n=1}^{J} p(\Delta N_n|\Delta \mathbf{N}^{(n-1)}; T), \quad (62)$$

where $\Delta \mathbf{N}^n = [\Delta N_1, \Delta N_2, \ldots, \Delta N_n]$. This equation can be simplified by noting that $\Delta N_{i+1} = b_\xi(t_i; T)\Delta N_i$, so ΔN_i defines a Markov process. Therefore:

$$p(\Delta N_1, \Delta N_2, \ldots, \Delta N_J|T) = \prod_{n=1}^{J} p(\Delta N_n|\Delta N_{n-1}; T), \quad (63)$$

where $p(\Delta N_1|\Delta N_0; T) = p(\Delta N_1)$.

In general, $p(\Delta N_n|\Delta N_{n-1}; T)$ is not known. Bettencourt points out that the maximum entropy density is the preferred density when only the mean is known [10]. This turns out to be the Poisson density:

$$p(\Delta N_n|\Delta N_{n-1}; T) = \lambda_{n-1}^{\Delta N_n} e^{-\lambda_{n-1}}/(\Delta N_n)!, \quad (64)$$

where $\lambda_n = b_\xi(t_n; T)\Delta N_n$. Taking the logarithm of Eq. (63), using Eq. (64), and eliminating terms independent of T, $-\log p(\Delta N_n|\Delta N_{n-1}; T)$ is proportional to:

$$\Phi(\Delta N_1, \Delta N_2, \ldots, \Delta N_J|T) = \sum_{n=2}^{J} \lambda_{n-1}(T) - \sum_{n=2}^{J} \Delta N_n \log\left(\lambda_{n-1}(T)\right). \quad (65)$$

A maximum likelihood estimate of the transmissibility T or basic reproduction number R_0 is equivalent to finding a T such that $\Phi(\Delta N_1, \Delta N_2, \ldots, \Delta N_J|T)$ is minimized.

Although a Poisson density is used in deriving the conditional likelihood function in Eq. (64), it has been observed that the number of differential outbreak cases $\Delta N(t_i)$ is over-dispersive, i.e., the variances are larger than expected. It can be shown that Consul's generalized Poisson distribution is a better match to the data. This density has the form [14]:

$$p(n) = (1-\omega)\lambda\left((1-\omega)\lambda + \omega n\right)^{n-1} \exp\left(-(1-\omega)\lambda - \omega n\right)/n!, \quad (66)$$

where $0 \leq \omega < 1$ is the dispersion parameter and λ is the mean. Euler's difference formula can be used to show that $p(n)$ is a valid density ($\sum_n p(n) = 1$).

2.3.2. Estimating Instantaneous Transmissibility

Define the instantaneous basic reproduction number $R_0(t)$ as $R_0(t) = Tu(t)G_1'(u(t))$, which follows from Eq. (60), so that $b_\xi(t;T) = \exp(\gamma\xi(R(t)-1))$. This equation has the advantage of being network independent. Now, consider the vector of estimates: $\mathbf{R}^J = [R_(t_1), R(t_2), \ldots, R(t_J)]$ and the vector of measurements $\Delta\mathbf{N}^J$, where $\Delta\mathbf{N}^J = [\Delta N_1, \Delta N_2, \ldots, \Delta N_J]$. The goal is to estimate the posterior density $p(\mathbf{R}^J|\Delta\mathbf{N}^J)$. Because the likelihood function is a nonlinear function of R_0, a linear Kalman filter cannot be used. In fact, an extended Kalman filter is not robust enough to handle the rapid fluctuations in the differential outbreak case histories. Therefore, a Bayesian approach is used where the posterior probability density of the state is constructed from the data. However, this density may be difficult to evaluate using kernel or grid-based estimation procedures. Therefore, the density is approximated using sampling procedures.

In order to illustrate the procedure, the following simple problem is considered. Suppose one is required to evaluate the N^{th} moment of $p(x|z)$:

$$\left\langle x^N|z\right\rangle = \int x^N p(x|z)\,dx. \tag{67}$$

Assume a proposal density $q(x|z)$ that is relatively easy to sample from. These samples are denoted by $x^i \sim q(x|z)$ such that $q(x|z) = \sum_{i=1}^{M}\delta(x-x^{(i)})/M$, where $\delta(x)$ is the Dirac delta function. Equation (67) can be rewritten as:

$$\left\langle x^N|z\right\rangle = \sum_{i=1}^{N} x^{(i)}\tilde{w}_i, \tag{68}$$

where $\tilde{w}_i = w_i/\sum_i w_i$ and $w_i = p(x^{(i)}|z)/q(x^{(i)}|z)$. Therefore, the N^{th} moment can be approximated by weighting samples $x^{(i)}$ from a proposal density $q(x|z)$ by a set of importance weights $\tilde{w}_i$.

The same procedure can be used to estimate the posterior density. In this case $w_i = p(\mathbf{R}^{J(i)}|\Delta\mathrm{N}^J)/q(\mathbf{R}^{J(i)}|\Delta\mathrm{N}^J)$ for a suitably chosen proposal density $q(\mathbf{R}^J|\Delta\mathbf{N}^J)$. Now, using Bayes theorem:

$$\begin{aligned} p(\mathbf{R}^J|\Delta\mathbf{N}^J) &= p(\Delta\mathbf{N}^J|\mathbf{R}^J)p(\mathbf{R}^J)/p(\Delta\mathbf{N}^J) \\ &= p(\Delta N_J|\Delta\mathbf{N}^{J-1},\mathbf{R}^J)p(\Delta\mathbf{N}^{J-1}|\mathbf{R}^J)p(\mathbf{R}^J)/p(\Delta\mathbf{N}^J) \\ &= p(\Delta N_J|\Delta N_{J-1},R_J)p(\mathbf{R}^J|\Delta\mathbf{N}^{J-1})/p(\Delta N_J|\Delta\mathbf{N}^{J-1}), \end{aligned} \tag{69}$$

where it is assumed that $p(\Delta N_J|\Delta\mathbf{N}^{J-1},\mathbf{R}^J) = p(\Delta N_J|\Delta N_{J-1},R_J)$. In addition, assume that R_J is Markov so that $p(\mathbf{R}^J|\Delta\mathbf{N}^{J-1}) = p(R_J|R_{J-1})p(\mathbf{R}^{J-1}|\Delta\mathbf{N}^{J-1})$. From Eq. (69):

$$p(\mathbf{R}^J|\Delta\mathbf{N}^J) = \frac{P(\Delta N_J|\Delta N_{J-1},R_J)p(R_J|R_{J-1})p(\mathbf{R}^{J-1}|\Delta\mathbf{N}^{J-1})}{p(\Delta N_J|\Delta\mathbf{N}^{J-1})}. \tag{70}$$

The proposal density can be factored into the following form using the product of conditional densities:

$$q(\mathbf{R}^J|\Delta\mathbf{N}^J) = q(R_J|\mathbf{R}^{J-1},\Delta\mathbf{N}^J)q(\mathbf{R}^{J-1}|\Delta\mathbf{N}^J).$$

Suppose that $q(R_J|\mathbf{R}^{J-1},\Delta\mathbf{N}^J) = q(R_J|R_{J-1})$ and $q(\mathbf{R}^{J-1}|\Delta\mathbf{N}^J) = q(\mathbf{R}^{J-1}|\Delta\mathbf{N}^{J-1})$. Using the definition of w_i above:

$$w_i = \frac{p(\Delta N_J|\Delta N_{J-1},R_J^{(i)})p(R_J^{(i)}|R_{J-1}^{(i)})}{p(\Delta N_J|\Delta N^{J-1})q(R_J^{(i)}|R_{J-1}^{(i)})}w_{i-1}. \tag{71}$$

For a simple particle filter examined here, $q(R_J^{(i)}|R_{J-1}^{(i)}) = p(R_J^{(i)}|R_{J-1}^{(i)})$, so that:

$$p(R_J|\Delta N_J,\Delta N_{J-1}) = \sum_i \tilde{w}_i\delta(R_J - R_J^{(i)}), \tag{72}$$

where $R_J^{(i)} \sim p(R_J|R_{J-1})$, $w_i = p(\Delta N_J|\Delta N_{J-1},R_J^{(i)})w_{i-1}$, and $\tilde{w}_i = w_i/\sum_i w_i$. In order to prevent particle degeneracy (collapse to a few particles), the nonuniform measure in Eq. (72) is replaced by a uniform measure by computing the cumulative distribution function (cdf) of $\tilde{w}_i$ and uniformly sampling to produce $\tilde{R}_J^{(i)}$ such that $p(R_J|\Delta N_J,\Delta N_{J-1}) = \sum_i \delta(R_J - \tilde{R}_J^{(i)})$ [15].

The state transition (sampling) density $p(R_J|R_{J-1})$ can be determined as follows. Augment the state R_J such that $\mathbf{R}_J = [R_J,\dot{R}_J)]^{\mathrm{T}}$ and assume a linear projection of the form:

$$\mathbf{R}_J = \mathbf{F}\mathbf{R}_{J-1} + \mathbf{w}_J, \tag{73}$$

where $\dot{R} = dR/dt$ and:

$$\mathbf{F} = \begin{pmatrix} 1 & \xi \\ 0 & 1 \end{pmatrix} \tag{74}$$

and $\mathbf{w}_J$ is a zero-mean Gaussian noise vector. Assume that $\ddot{R} = w(t)$, where $w(t)$ is a zero-mean white Gaussian noise process and $\ddot{R} = d^2w/dt^2$. This constitutes a random "acceleration" model. It can be shown that the corresponding covariance matrix associated with the state $\mathbf{R}_J$ in Eq. (73) is:

$$\mathrm{Cov}(R,\dot{R}) = q\begin{pmatrix} \xi^3/3 & \xi^2/2 \\ \xi^2/2 & \xi \end{pmatrix}, \tag{75}$$

where q is a variance-like term. It then follows that $p(R_J|R_{J-1})$ is a zero-mean Gaussian density with covariance $\mathrm{Cov}(R, \dot{R})$. The likelihood function is computed from Eq. (64).

3. Results and Discussion

3.1. Steady-State Disease Spread

Some illustrative results using the steady-state disease spread model are described below for three types of contact networks: Poisson, Exponential, and Power-law. For these examples, the critical transmissibility is $T_c = 0.049$. This means that if the observed transmissibility $T = T_c$, then $R_0 = 1$, i.e., T_c is the critical value above which the pandemic is self-sustaining. For comparative purposes, each contact network has the same critical threshold T_c.

Fig. 3 depicts the outbreak distribution size prior to a pandemic when $R_0 = 0.8$ with no NPIs. This is the predicted number of people infected by a small outbreak. Both the Poisson and exponential (urban) contact networks have similar size distributions, where large outbreak sizes (> 20) are unlikely. The power-law contact network is highly peaked for small outbreak sizes as expected. Although diminishingly small for larger outbreak sizes, it is not zero because there are a minority of super-spreaders with large contact degree. Fig. 4 depicts the fraction of the population affected by a pandemic as a function of the basic reproduction number R_0 when $R_0 > 1$ with no NPIs. Communities have diverse experiences based on their contact patterns. The Poisson contact network has the greatest fraction of the population affected by a pandemic because individuals in a group are equally likely to become infected and to infect others. For an exponential (urban) contact network, there is a 50% reduction in the number of infected individuals compared to a well-mixed population (40% versus 80%) for $R_0 = 2$. For a power-law network, only 5% of the population is affected by the pandemic for $R_0 = 2$. Therefore, outbreaks are consistently less likely to reach pandemic proportions.

There are outbreaks that can occur outside the main pandemic cluster, although they are relatively small. This is not predicted from SIR or Susceptible, Exposed, Infected, or Recovered (SEIR) models. Fig. 5 shows the average number of people infected by outbreaks outside the main pandemic cluster. This average outbreak size can be determined as a function of R_0 for the three contact networks. For $R_0 > 2$ all three contact networks have low average outbreak sizes (less than 2)

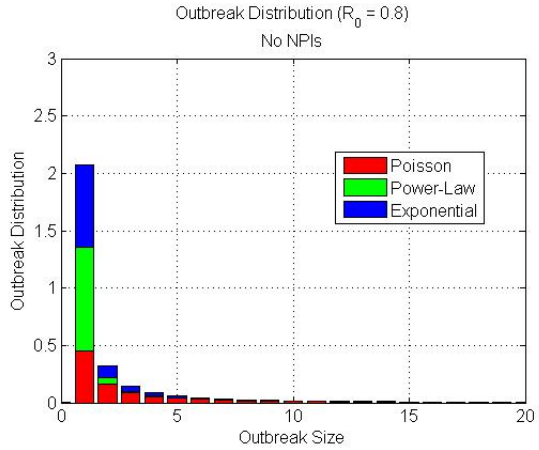

Figure 3. Outbreak size distribution prior to a pandemic ($R_0 = 0.8$) for three types of contact networks

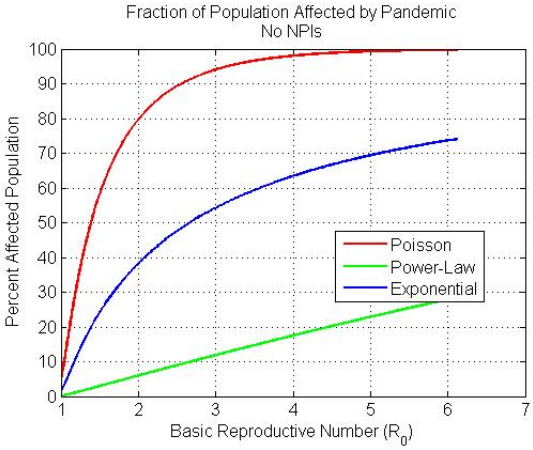

Figure 4. Pandemic size (or fraction of individuals affected) as a function of R_0 for three types of contact networks

outside the main pandemic cluster. As expected, there are more outbreaks near $R_0 = 1$ because less of the population is affected by the main pandemic and there is more opportunity for infections to spread outside the main pandemic cluster. Note that the power-law network has a larger outbreak size for larger R_0 because there are a minority of super-spreaders. Fig. 6 illustrates the individual risk of infection when $R_0 = 2$ based on the number of social contacts. The Poisson contact network shows the most risk of individual infection. For ten contacts, the risk of infection is approximately 55%. The exponential (urban) network shows an individual infection risk of approximately 45% for ten contacts. The power-law contact network shows the smallest individual risk of infection, roughly 20% for ten contacts.

For uniform social distancing, it is assumed that the prob-

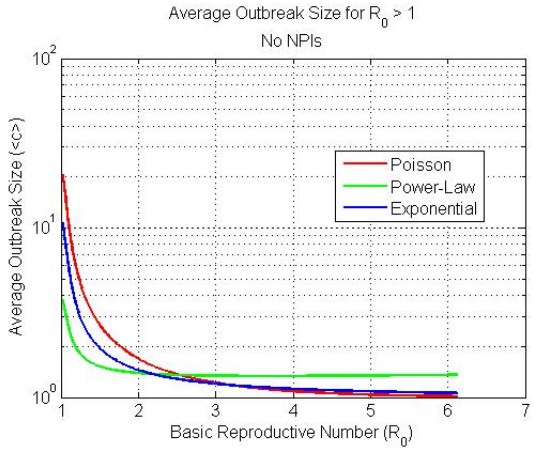

Figure 5. Average outbreak size removed from the main pandemic cluster for three types of contact networks

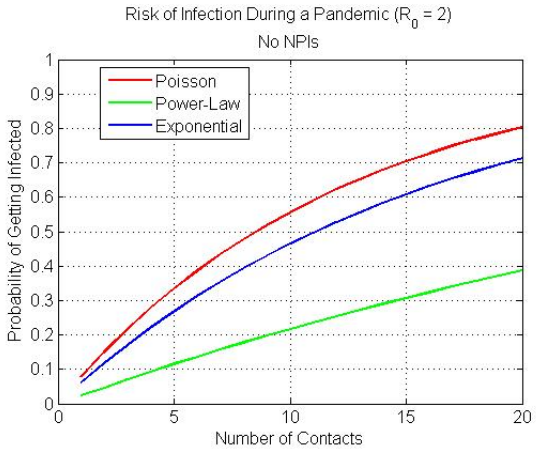

Figure 6. Individual risk of an infection during a pandemic based on the number of contacts for three types of contact networks

ability that a vertex of degree k is active is $b_k = b$, where $0 \leq b \leq 1$, so $p_k^{'} = b_k p_k$. Basically, social distancing is applied to every individual regardless of the number of contacts an individual may have. Fig. 7 depicts the fraction of the population affected by a pandemic as a function of R_0 for different degrees of uniform social distancing for three types of networks. Uniform social distancing has the effect of shifting the onset of a pandemic to larger effective R_0 values. This is because the effective critical transmissibility threshold T_{ceff} is now greater than the baseline critical threshold T_c before intervention is imposed. Asymptotic results for the Poisson network are consistent with SIR/SEIR compartmental model results. In each case, the fraction of the affected population is reduced exactly by the amount of imposed social distancing. The exponential network shows

similar trends although its asymptotic values are smaller than the Poisson network because the population is not well-mixed. The power-law network shows little variation in its pandemic onset or its peak (at $R_0 = 6$) for a 20% increase in social distancing from the baseline. Even at 40% social distancing, the percent affected population is only somewhat reduced, although its pandemic onset is shifted to a larger R_0 value. As explained previously, this is due to a minority of super-spreaders that add a degree of robustness to the network

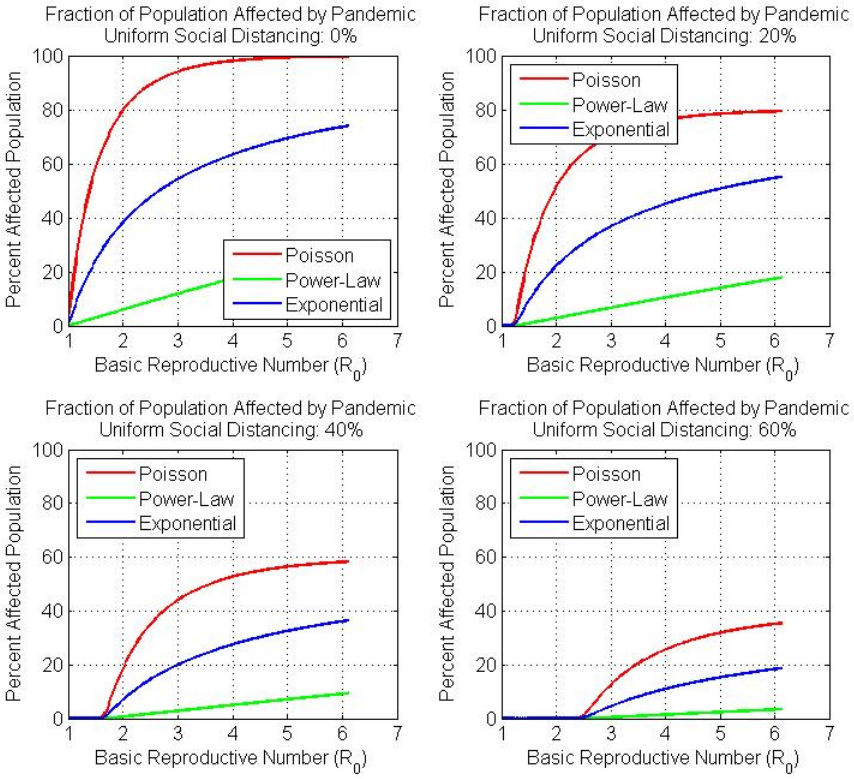

Figure 7. Pandemic size as a function of R_0 for three types of contact networks assuming uniform social distancing

For directed social distancing, it is assumed that $b_k = 1$ for $0 \leq k \leq K_{\max}$ and $b_k = 0$, otherwise. Basically, all individuals with contact degree greater than $K_{\max}$ are distanced. Note that $K_{\max}$ is related to the fractional number f_c of nodes removed, or $f_c = 1 - \sum_{k=0}^{K_{\max}} p_k$. Fig. 8 illustrates the fraction of the population affected by a pandemic as a function of R_0 for different degrees of directed social distancing for three types of networks. Directed social distancing has a large effect on the onset of a pandemic for both the exponential and power-law contact networks. For a 10% reduction in social distancing for

high social contact individuals, the affected population is reduced from 75% to 50% for an exponential network and 30% to zero for a power-law network when $R_0 = 6$. The affected population for a Poisson network is reduced by only 10%. Although not depicted in this figure, it can be shown that for a 1.5% reduction in directed social distancing, there is a 25% reduction in the affected population for a power-law network when $R_0 = 6$. For a 2% reduction in directed social distancing, none of the population is affected for $R_0 \leq 6$. However, the exponential network shows only a 5% reduction and the Poisson network shows only a 2% reduction in the affected population, respectively. The power-law network is the least robust to the removal of high contact nodes. This is an important result that could impact the way contact tracing is performed.

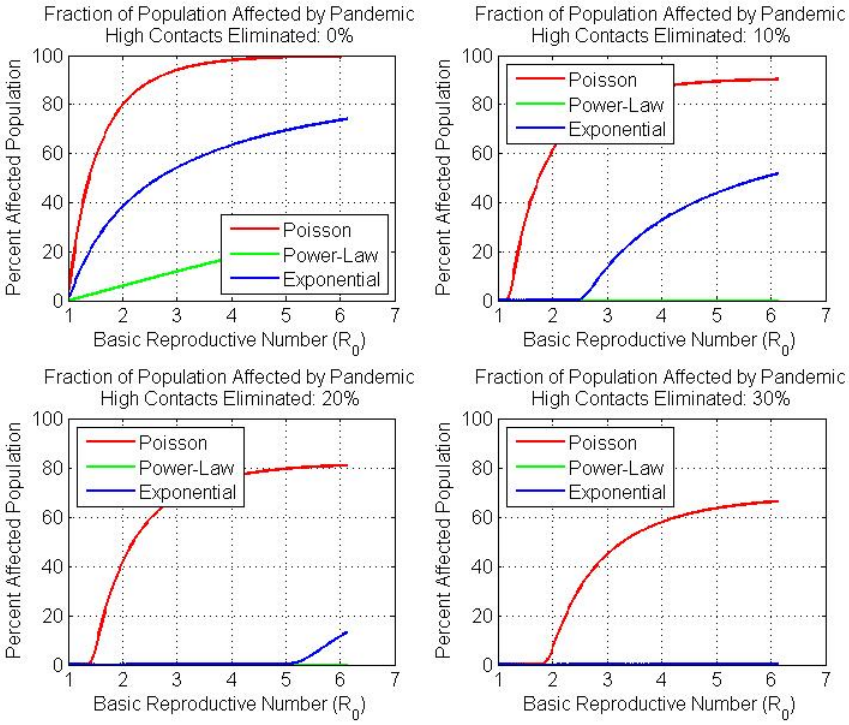

Figure 8. Pandemic size as a function of R_0 for three types of contact networks assuming directed social distancing

3.2. Time-Dependent Disease Spread in Social Networks

Some illustrative results using the time-dependent disease spread model are described below. Here, it is assumed that the critical transmissibility is $T_c = 0.049$, the observed transmissibility is $T = 0.098$ (implying $R_0 = 2$ initially), the infectious

period T_r is 7 days, the inception and release of (uniform) social distancing starts at week 6 and ends at week 14, respectively, and testing and contact tracing begins at week 16. Additionally, 80% of individuals are social distanced, the percent isolated per week is 20%, and the percent traced per week is 30% per week.

Figures 9 and 10 depict the susceptible, infectious, and outbreak case load probabilities for a Poisson and exponential network, respectively, for the example parameters outlined above. The pandemic begins to build after week 2 and is arrested starting at week 6 as a result of social distancing. Social distancing is relaxed at week 14 with a consequent buildup in infections until testing and contact tracing are initiated at week 16. There is a large decline in the susceptible probability due to contact tracing around week 20. Since testing and contract tracing is assumed to continue over the model run (50 weeks), the outbreak case load reaches a steady state. Contrasting the differences between the Poisson and exponential networks, it is apparent that the number of outbreak cases is smaller for the exponential network and remediation strategies such as social distancing, testing, and contact tracing are more effective for the exponential network. Recall that this network does not assume a well-mixed population where all individuals are equally likely to become infected. The power-law network (not shown) has an even smaller number of outbreak cases.

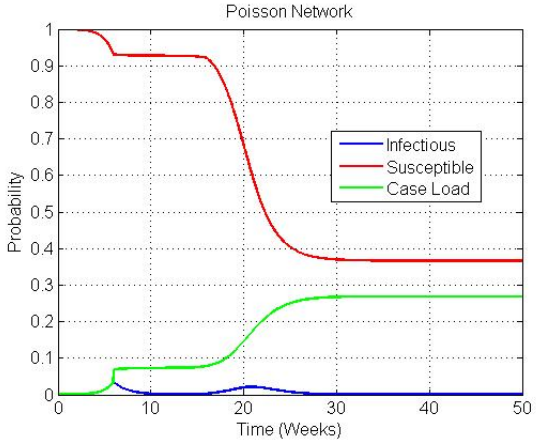

Figure 9. Susceptible, infectious, and outbreak case load probabilities as a function of time for a Poisson network based on example parameters

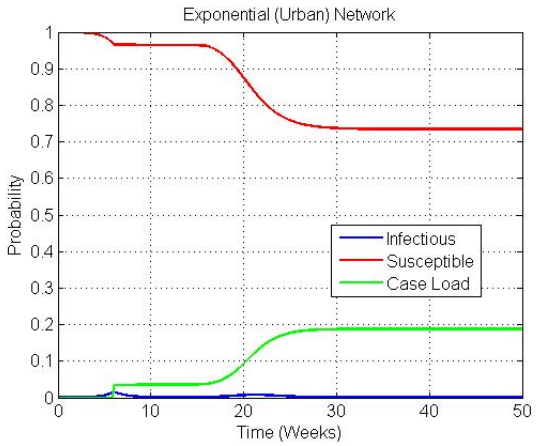

Figure 10. Susceptible, infectious, and outbreak case load probabilities as a function of time for an exponential network based on example parameters

Fig. 11 depicts the instantaneous basic reproduction number R_0 for three contact networks as a function of time for the NPI, testing, and contact tracing example outlined above. The inception and lifting of social distancing protocols are clearly evident, where R_0 is sizeably reduced over a period of 2 months. Note that for all three networks, $0.3 \leq R_0 \leq 0.4$ over this interval of time. When the NPI is lifted at week 14, sheltered individuals are added back into the reservoir with a subsequent increase in R_0. This pool slowly attrits from infections until week 16 when testing and contact tracing is initiated. A combination of testing and contact tracing lowers R_0 to below unity near week 20, at which point, the disease ceases to spread. It is interesting to note that the power-law network has consistently lower R_0 values except during testing and contact tracing. This is because the rather low rates of testing and tracking miss a number of super-spreaders that are more prevalent in a power-law network.

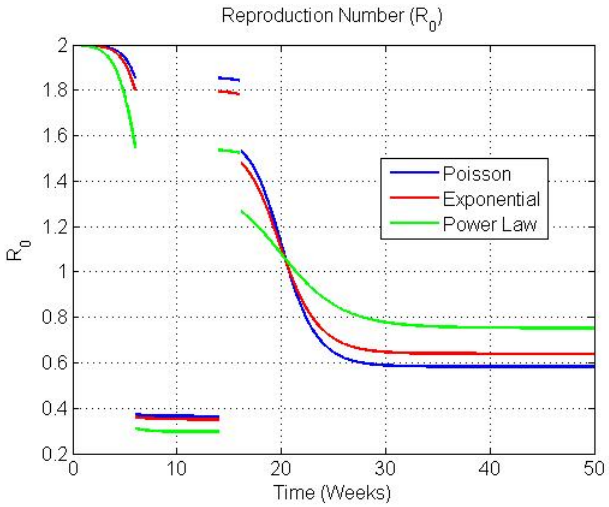

Figure 11. Instantaneous basic reproduction number R_0 as a function time for three contact networks based on example parameters

3.3. Bayesian-based Estimation of Contact Transmissibility

Some illustrative results using the Bayesian-based estimation procedure for contact transmissibility are discussed below. Figures 12 and 13 illustrate particular examples of the approach for Massachusetts and New York given collected outbreak case histories from each state over a 2-3 month period starting in February 2020. It is assumed that the critical transmissibility $T_c = 0.049$ and the infection period $T_r = 7$ days. The models that best fit the data are used for each state – a Poisson

network for Massachusetts and an exponential network for New York. The estimated transmissibility for Massachusetts is 0.089, which translates into a R_0 value of 1.8. Similarly, the estimated transmissibility for New York is 0.185, which translates into a R_0 value of 3.8. Both networks provide reasonable fits to their respective observed differential outbreak case histories $\Delta C(t)$ and reinforces the notion that the rapidity of disease spread in New York was much more severe.

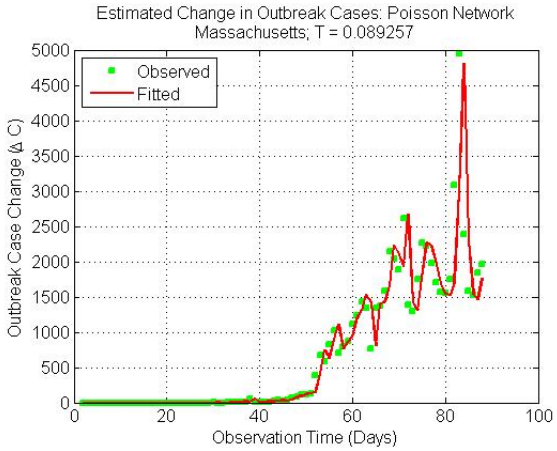

Figure 12. Estimated and predicted change in outbreak cases $\Delta C(t)$ for a Poisson network based on Massachusetts State outbreak data

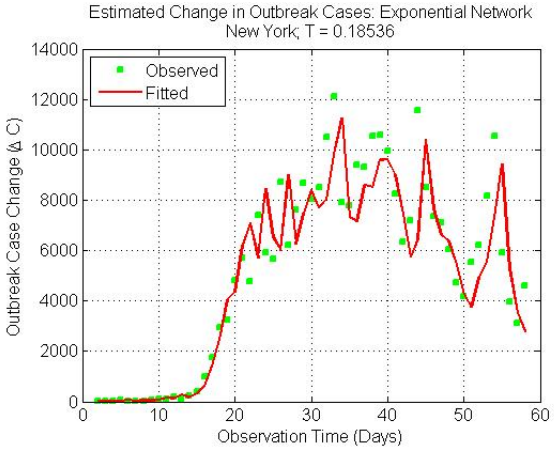

Figure 13. Estimated and predicted change in outbreak cases $\Delta C(t)$ for an exponential network based on New York State outbreak data

The instantaneous basic reproduction number $R_0(t)$ can also be estimated by tracking the differential outbreak cases over time using a particle filter. Figures 14 and 15 illustrate the estimation technique for New York given the collected outbreak

history. The fit to the observed differential outbreak case data $\Delta C(t)$ is quite good. In addition, the instantaneous $R_0(t)$ values are consistent with the average R_0 value using an exponential network model depicted in Fig. 13. However, Fig. 15 is more illustrative because it allows one to examine the trend in R_0 over the progression of the pandemic. In this case, there is a downward trend after reaching a peak of approximately 3.5. There is also an up tick in R_0 at later times probability due to an increase in testing. Note that estimation errors may result in negative $R(t)$ values that are not realistic.

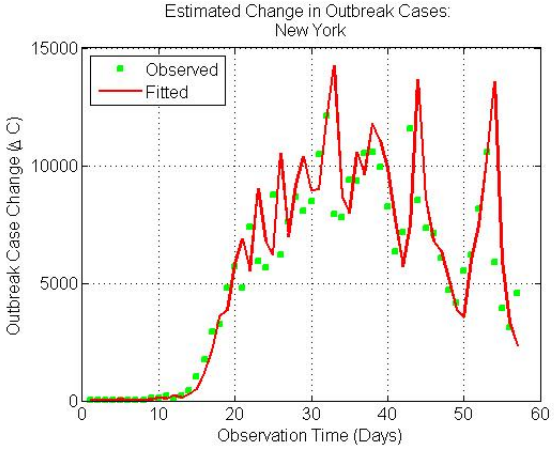

Figure 14. Estimated and predicted change in outbreak cases $\Delta C(t)$ based on New York State outbreak data using a particle filter

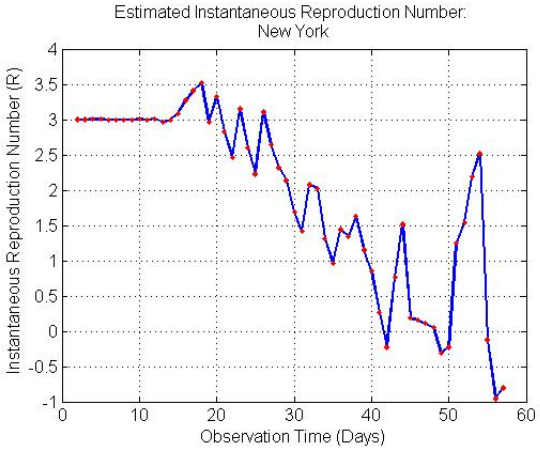

Figure 15. Estimated instantaneous basic reproduction number R_0 based on New York State outbreak data using a particle filter

4. Conclusion

In the present work, we developed and illustrated a coarse-grain analytic modeling procedure that can be used to assess alternative strategies of implementing and subsequently lifting non-pharmaceutical interventions in response to the COVID-19 pandemic. This work was based on developing a multi-dimensional view of the problem by examining steady-state and dynamic disease spread using a network-based approach. The steady-state models, based on percolation theory, highlighted the previously known result that social contact structure is a key factor in the size of an outbreak. In addition, it was shown that the social contact structure influences the types of social distancing protocols that are deemed most effective. The dynamic models, based on a stochastic reformulation of the SIR equations, further extended the work to include the effects of lifting non-pharmaceutical interventions and the importance of testing and contact tracing in reducing the overall infection rate. Providing a realistic assessment of the basic reproduction number R_0 is also important in gauging the severity of the outbreak within a specific geographic area. A Bayesian-based estimation procedure was developed to estimate both the average and instantaneous basic reproduction number from outbreak case histories at a state and county-wide level. These estimates can be used to seed other models or analysis procedures.

No single model is a panacea. Therefore, we advocate an ensemble modeling approach based on a combination of analytic and fine-grain agent-based models (ABMs). As outlined in a companion paper, this approach has the potential to provide valuable insights into disease spread and the effectiveness of non-pharmaceutical interventions. It could prove to be a valuable tool for decision-makers in conjunction with empirical analysis.

References

[1] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Physical review letters 86 (14) (2001) 3200.

[2] M. E. Newman, Spread of epidemic disease on networks, Physical review E 66 (1) (2002) 016128.

[3] M. E. Newman, The structure and function of complex networks, SIAM review 45 (2) (2003) 167–256.

[4] D. S. Callaway, M. E. Newman, S. H. Strogatz, D. J. Watts, Network robustness and fragility: Percolation on random graphs, Physical review letters 85 (25) (2000) 5468.

[5] M. Newman, Networks: An Introduction, Oxford, England, 2010.

[6] P. M. Chaikin, T. C. Lubensky, T. A. Witten, Principles of condensed matter physics, Vol. 10, Cambridge university press Cambridge, 1995.

[7] M. Barthélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, Journal of theoretical biology 235 (2) (2005) 275–288.

[8] J. H. University, Coronavirus Resource Center, `https://coronavirus.jhu.edu` (2020).

[9] L. A. Meyers, B. Pourbohloul, M. E. Newman, D. M. Skowronski, R. C. Brunham, Network theory and sars: predicting outbreak diversity, Journal of theoretical biology 232 (1) (2005) 71–81.

[10] L. M. Bettencourt, R. M. Ribeiro, Real time bayesian estimation of the epidemic potential of emerging infectious diseases, PLoS One 3 (5) (2008).

[11] G. Chowell, C. Castillo-Chavez, P. W. Fenimore, C. M. Kribs-Zaleta, L. Arriola, J. M. Hyman, Model parameters and outbreak control for sars, Emerging Infectious Diseases 10 (7) (2004) 1258.

[12] S. L. Feld, Why your friends have more friends than you do, American Journal of Sociology 96 (6) (1991) 1464–1477.

[13] L.-S. Young, S. Ruschel, S. Yanchuk, T. Pereira, Consequences of delays and imperfect implementation of isolation in epidemic control, Scientific reports 9 (1) (2019) 1–9.

[14] P. Consul, F. Famoye, Generalized poisson regression model, Communications in Statistics-Theory and Methods 21 (1) (1992) 89–109.

[15] N. Gordon, B. Ristic, S. Arulampalam, Beyond the kalman filter: Particle filters for tracking applications, Artech House, London 830 (5) (2004) 1–4.