BOUNDING THE SIZE AND PROBABILITY OF EPIDEMICS ON NETWORKS

JOEL C. MILLER,* Center for Nonlinear Studies and Mathematical Modeling & Analysis, Los Alamos National Laboratory

Abstract

We consider the spread of infectious disease on a network. The population has heterogeneous infectiousness and/or susceptibility. These may be quantified by the out-transmissibility (the marginal probability of infecting a neighbor given a node’s infectiousness) or the in-transmissibility (the marginal probability of being infected by a neighbor given a node’s susceptibility). For given distributions of in-transmissibility, we find distributions of out-transmissibility which give upper and lower bounds on both the size and probability of an epidemic under weak assumptions. We similarly find distributions of in-transmissibility giving upper and lower bounds for a fixed out-transmissibility distribution. In the special case of networks with high girth, we are able to prove stronger results. In general, the probability and size of epidemics are maximal when the population is homogeneous and minimal when the variance of in- or out-transmissibility is maximal.

Keywords: Epidemiology, Networks

2000 Mathematics Subject Classification: Primary 92D30
Secondary 60K35

1. Introduction

Mathematical theories modeling the spread of infectious diseases have been developed in a number of fields. The interaction between these different fields has been relatively sparse, leading to repeated discoveries of some results as well as a lack of cohesion in the topics studied. In this paper, we investigate the effect of heterogeneity in the population, attempting to fuse some of the different lines of study [2, 3, 25, 17, 14, 23, 29] and to extend them.

The earliest mathematical epidemic models were developed by Kermack and McKendrick [15] and Reed and Frost (unpublished) as described by [1]. The Kermack–McKendrick model has been most widely used and extended. It uses deterministic ordinary differential equations (ODEs) to calculate the dynamics of the outbreak. It assumes that the population may be divided into three compartments: Susceptible (S), Infected (I), and Recovered (R), evolving according to

\[ \dot{S} = -\beta IS \]
\[ \dot{I} = \beta IS - \gamma I \]
\[ \dot{R} = \gamma I \]

Because of the constant recovery rate \( \gamma \), the model implicitly assumes that the duration of infection is exponentially distributed.

*Postal address: MS B284, Los Alamos National Laboratory, Los Alamos, NM 87544, USA
This is a simple example of an SIR model. This general class of models has been expanded to study a number of effects such as age stratification or spatial structure. Most of these studies have used ODEs and divided the $S$ and $I$ compartments into subcompartments with different parameters for each. One such investigation of particular interest is [19] which considered a wide class of variations in infectiousness, finding that only the average, rather than the details of the distribution, has an effect on the final size.

ODE models have a number of weaknesses which can be overcome to varying degrees: the infectious period is exponentially distributed, the models are deterministic and so cannot determine the probability an epidemic occurs, and the models assume random mixing between groups. Although these may be addressed to some extent (usually by increasing the number of compartments), the methods are awkward and do not necessarily have clear epidemiological meaning.

An alternative way to eliminate these weaknesses is to introduce stochastic or network models. By considering the disease progression at the individual level, we have complete control over the infectious period duration as well as the interaction of individuals. Although stochastic and network models are generally considered separately, under weak assumptions, stochastic models may be placed into a network framework (see [13, 25] and appendix A).

A further method of modeling epidemics without ODEs uses explicit agent-based simulation [7, 4, 6, 8, 9]. In these simulations, individuals move through populations, having contacts which may lead to disease transmission. Although there is time data available from the simulation, we gain considerable insight by considering the network of contacts as static and investigating disease spread on that network. For the remainder of this paper, we use the framework of network-based epidemic models.

In a typical network epidemic model, the outbreak is initialized with a single infection. The disease may then spread to neighboring nodes. The outbreak may be self-limiting and die out before infecting many nodes, or it may become an epidemic, spreading until its size is limited by the population size.

Network (and stochastic) epidemic models have primarily been studied by the statistics community [30, 17, 18, 2, 3, 24] and the statistical physics community [26, 22, 21, 28, 10, 27, 20]. Some work has also been done directly by epidemiologists [11, 12]. In general the statistics community has produced more rigorous results, but has considered more restricted classes of networks. The physics community has considered a wider range of networks, but the results are less rigorous.

Before discussing earlier results, we introduce some terminology. The transmissibility $T_{uv}$ is the probability that an infection of node $u$ would result in direct infection of the neighbor $v$. The in-transmissibility $T_{in}(v)$ is the marginal probability that a neighbor of $v$ would infect $v$ given the characteristics of $v$ and the out-transmissibility $T_{out}(u)$ is the marginal probability that $u$ would infect a neighbor given the characteristics of $u$. Both the in- and out-transmissibility necessarily have the same average transmissibility $\langle T \rangle$. These definitions will be made more precise later.

Most network-based epidemic models assume that $T_{uv} = \langle T \rangle$ for all pairs of nodes. Those models that allow heterogeneities generally show heterogeneities reduce the probability or size of epidemics [2, 17, 29, 23, 14]. We single out [17], for which $T_{in}(v) = \langle T \rangle$ for all $v$, but $T_{out}(u)$ is heterogeneous (induced by a heterogeneous infection period). It was shown that in a general network an epidemic is most likely
and largest if $T_{\text{out}}(u) = \langle T \rangle$ for all $u$. It was noted by [29] that the argument proving this also proves that the epidemic is least likely and smallest if $T_{\text{out}}(u) = 1$ for a proportion $(T)$ of the vertices, and $T_{\text{out}}(u) = 0$ for the remainder. This is similar to recent work [23, 14] where the same cases give upper and lower bounds on epidemic probability, but the size of epidemics was unaffected. The fact that epidemic size is unaffected in these more recent models (with constant $T_{\text{in}}$) is a result of the fact that [17] considered general networks, while [23, 14] considered networks with no short cycles. The work of [23] also found that the upper bound for the size of epidemics occurs when $T_{\text{in}}$ is homogeneous and the lower bound occurs when it takes only the values 0 or 1.

In this paper we investigate the spread of epidemics in which both infectiousness and susceptibility vary and generalize the results of [17, 23]. We consider arbitrary networks.

In general, at early stages in an outbreak, the data on transmission is limited. It is possible that policy-makers will be faced with a decision given only information on the distribution of $T_{\text{in}}$ or the distribution of $T_{\text{out}}$. Consequently, it is important to understand how an epidemic may progress given just knowledge of one of these distributions. For a fixed distribution of $T_{\text{in}}$ we find the distributions of $T_{\text{out}}$ giving upper and lower bounds on epidemic size and probability. Similarly for a fixed distribution of $T_{\text{out}}$ we find distributions of $T_{\text{in}}$ giving upper and lower bounds.

This paper is structured as follows: in section 2 we introduce the model and clarify definitions. In section 3 we consider epidemics spreading on general networks. In section 4 we find stronger results for networks with no short cycles. Finally in section 5 we discuss extensions and implications of our results. Although our results are stated for families of networks, this is a convenience which allows us to consider the network size growing to infinity. In fact, our results are proven by first considering individual networks, and so the upper and lower bounds found apply to individual networks, not just families of networks.

2. The Model

We consider the spread of disease on a network $G$. An outbreak begins when a single node (the index case) chosen uniformly from the population is infected. The disease spreads from an infected node $u$ to a susceptible node $v$ with a probability equal to the transmissibility $T_{uv}$. Each infected node attempts to infect each of its susceptible neighbors and then recovers (and is no longer infected or susceptible). The outbreak ends when no infected nodes remain.

Following [23], we assume that the factors influencing infectiousness of $u$ may be summarized in a (possibly vector-valued) quantity $I_u$ and the factors influencing the susceptibility of $v$ may be summarized in a (possibly vector-valued) quantity $S_v$. Then $T_{uv}$ may be expressed as

$$T_{uv} = T(I_u, S_v)$$

We assume that $I$ and $S$ are assigned independently. We use $P(I)$ or $P(S)$ to denote the probability density function for $u$ to have $I$ or $S$ respectively. It is understood that $P(I)$ and $P(S)$ are not the same function.

The probability that $u$ would infect a randomly chosen neighbor is given by the
out-transmissibility of $u$

\[ T_{out}(u) = \int T(I_u, S) P(S) dS \]

and the probability that $v$ would be infected by a randomly chosen neighbor is given by the in-transmissibility of $v$

\[ T_{in}(v) = \int T(I, S_v) P(I) dI \]

At times it will be convenient to use $T_{out}(I)$ and $T_{in}(S)$ [rather than $T_{out}(u)$ and $T_{in}(v)$] by which we mean the out- and in-transmissibility of nodes with $I$ or $S$ respectively. In general $T_{out}(I)$ and $T_{in}(S)$ may not be invertible functions.

From $P(S)$ and $P(I)$, we may find the distribution of $T_{in}$ and $T_{out}$. We use $Q_{in}(T_{in})$ to be the probability density function for the in-transmissibility $T_{in}$ and $Q_{out}(T_{out})$ to be the probability density function for the out-transmissibility $T_{out}$. The averages $\int T_{in} Q_{in}(T_{in}) dT_{in}$ and $\int T_{out} Q_{out}(T_{out}) dT_{out}$ are both equal to $\langle T \rangle$.

Given distributions of $I$ and $S$ and the form of $T$, there is always a $Q_{in}$ and $Q_{out}$ pair that result. Also, given a $Q_{in}$ or $Q_{out}$ it is always possible to find $P(I)$, $P(S)$ and $T$ that are consistent. For example, given any $Q_{in}$ we may take $P(S) = Q_{in}(S)$ with constant $I$. Then $T(I, S) = S$ is consistent and yields $Q_{out}(T_{out}) = \delta(T_{out} - \langle T \rangle)$. That is, for any in-transmissibility distribution, homogeneous out-transmissibility is consistent. However, given a pair $Q_{in}$ and $Q_{out}$ it is not always possible to find $P(I)$, $P(S)$, and $T$. So $Q_{in}$ need not be consistent with $Q_{out}$ even if both have the same $\langle T \rangle$.

Although in principle $I$ and $S$ may be vector-valued, they frequently are assumed to be scalars with

\[ T(I_u, S_v) = 1 - \exp(-\alpha I_u S_v) \]  

(1)

A number of interpretations can yield this form. For example: Let $\alpha$ be the rate at which virus from an infected person reaches a susceptible person. Let $I_u$ be the infectious period of $u$. Let $S_v$ be the probability that a virus reaching $v$ causes infection. Then the probability $p$ that $v$ has not become infected satisfies

\[ \dot{p} = -\alpha S_v \]

Integrating this over the infectious period $I_u$ of $u$ yields (1). It is frequently assumed that $I$ is exponentially distributed since this corresponds to the constant recovery rate in ODE models.

2.1. Epidemic Percolation Networks

From the original network $G$ we define a new directed network $\mathcal{E}$. The nodes of $\mathcal{E}$ are taken to be the nodes of $G$. For each edge $(u, v)$ of $G$ we place directed edges $(u, v)$ and $(v, u)$ into $\mathcal{E}$ with probability $T_{uv}$ and $T_{vu}$ respectively. The resulting network $\mathcal{E}$ is called an epidemic percolation network (EPN) [13]. For a given node $u$, the distribution of out-components over possible realizations of $\mathcal{E}$ is identical to the distribution of outbreaks beginning with $u$ in the dynamic process above. Thus we may think of an outbreak as the set of nodes infected by defining an EPN, choosing an index case, and infecting all nodes in the out-component of the index case. Under this interpretation, the probability that a node becomes infected in an outbreak is the average fraction of
Epidemic Size and Probability

Figure 1: G is an Erdős–Rényi graph of 50 nodes with average degree 3. The EPN E was created using (1) with $\alpha = 1.6$, $I$ chosen from an exponential distribution with parameter 1, and $S$ chosen uniformly from $(0,1)$. In $E$, $H_{scc}$ is denoted by $\circ$. The nodes in $H_{in}$ are the $\circ$ and $\Box$ nodes. The nodes in $H_{out}$ are the $\circ$ and $\Diamond$ nodes. Infection of any node in $H_{in}$ results in infection of all nodes in $H_{out}$.

Figure 1: (a) $G$ and (b) An EPN $E$ from $G$.


2.2. Epidemic size and probability

We take nonempty families of finite connected networks $G_n$ with $n \to \infty$

$$G_1 \supset G_2 \supset G_3 \supset \cdots$$

We assume that any sequence of networks $G_n$ such that $G_n \in G_n$ satisfies $|G_n| \to \infty$ as $n \to \infty$.

In the following we will frequently choose a node $u$ from the family $G_n$. There are several ways to do this, and the method used can affect resulting statistics. We assume that a network $G \in G_n$ is chosen uniformly and then a node $u \in G$ is chosen uniformly.

We use $S_d$ to denote a subgraph induced by taking all nodes within a distance $d$ of a central node. The structure of a particular $S_d$ may be repeated many times in a given network or family of networks. We define $P_{G_d}(S_d)$ to be the probability that a randomly chosen $u$ in $G_d$ yields $S_d$. 

We assume that $P_{G_n}(S_d) = P_{G_d}(S_d)$ for all $n \geq d$. That is we assume that small-scale structure is constant for large enough $n$. Thus the probability that a node begins or ends a path of length $d$ in an EPN is the same for all $n \geq d$. For a given EPN, we define $H_{in}(d)$ and $H_{out}(d)$ to be the set of nodes from which a length $d$ path starts or at which a length $d$ path ends respectively.

For a given network $G$, we define $\mathcal{P}_d(G) = \mathbb{E}[|H_{in}(d)|]/|G|$ and $\mathcal{A}_d(G) = \mathbb{E}[|H_{out}(d)|]/|G|$. This is the probability that a randomly chosen node will be in $H_{in}(d)$ or $H_{out}(d)$ respectively. We extend these definitions to families of networks: for $G_n$ we define $\mathcal{P}_d(G_n)$ and $\mathcal{A}_d(G_n)$. By our assumptions, if $n \geq d$, $\mathcal{P}_d(G_n) = \mathcal{P}_d(G_d)$ and $\mathcal{A}_d(G_n) = \mathcal{A}_d(G_d)$. We finally define

$$\mathcal{P} = \lim_{d \to \infty} \mathcal{P}_d(G_d)$$
$$\mathcal{A} = \lim_{d \to \infty} \mathcal{A}_d(G_d)$$

$\mathcal{P}$ measures the probability of an epidemic and $\mathcal{A}$ measures the attack rate in the limit $n \to \infty$.

We restrict our attention to families of graphs for which the above limits exist. For applications, we want these limits to make sense for individual networks, rather than families of networks. So we are primarily interested in families for which choosing a sequence of networks $G_n$ such that $G_n \in \mathcal{G}_n$ gives

$$\lim_{d \to \infty} \lim_{n \to \infty} \mathcal{A}_d(G_n) = \mathcal{A}$$

and

$$\lim_{n \to \infty} \lim_{d \to \infty} \mathcal{P}_d(G_n) = \mathcal{P}$$

This occurs when $P_{G_n}(S_d)$ converges to $P_{G_d}(S_d)$ as $n \to \infty$. Thus, in a given large network for which the small-scale structure is well-behaved in this sense, the probability and size of epidemics is close to that predicted by $\mathcal{P}$ and $\mathcal{A}$.

Having defined the probability $\mathcal{P}$ and attack rate $\mathcal{A}$, we move on to some other necessary quantities. Let a set of nodes $V = \{v_1, v_2, \ldots, v_n\}$ be chosen and assign susceptibilities to each $v \in V$ so that $\vec{S} = (S_{v_1}, S_{v_2}, \ldots, S_{v_n})$. Define

$$\phi_{in}(V, \vec{S}) = \int \prod_{v \in V} [1 - T(I, S_v)]P(I)dI$$

(3)

This is the \textit{a posteriori} probability that a randomly chosen neighbor of all nodes $v \in V$ will not infect any $v$ given their susceptibilities. We may similarly define

$$\psi_{in}(V) = \int \phi_{in}(V, \vec{S})P(\vec{S})d\vec{S}$$
$$= \int (1 - T_{out})^{|V|}Q_{out}(T_{out})dT_{out}$$

(4)

This is the \textit{a priori} probability that a randomly chosen neighbor will not infect any $v \in V$ prior to assigning the susceptibilities of $V$. Note that if $|V| = 1$, then $\psi_{in}(V) = 1 - \langle T \rangle$. 

We may equivalently define

\[ \phi_{\text{out}}(V, \mathbf{I}) = \int \prod_{v \in V} [1 - T(I_v, S)] P(S) dS \]

and

\[ \psi_{\text{out}}(V) = \int (1 - T_{\text{in}})^{|V|} Q_{\text{in}}(T_{\text{in}}) dT_{\text{in}} \]

where we instead measure the probability of a node not to be infected by a node in \( V \).

We have now completed the definitions needed for the remainder of the paper.

### 3. Bounds in general networks

We begin by considering epidemics spreading in general networks. We make a simplifying assumption and show that for a given \( Q_{\text{in}} \), both \( P \) and \( A \) are maximal when \( T_{\text{out}} \) is homogeneous. They are minimal when the variance of \( T_{\text{out}} \) is maximal.

We are able to prove that the global upper bound on \( P \) and \( A \) occurs when \( T_{uv} = \langle T \rangle \) for all edges. We are not able to prove a global lower bound, but we give a conjecture.

We begin by making an assumption that infectiousness and susceptibility are ordered.

**Assumption 1.** (Ordering Assumption.) We assume that if \( T(I_u, S_v) > T(I_{u'}, S_v) \) for any \( v \), then for all \( v' \), \( T(I_u, S_{v'}) \geq T(I_{u'}, S_{v'}) \).

Further, if \( T(I_u, S_v) > T(I_{u'}, S_v) \) for any \( v \), then the set of \( S \) for which \( T(I_u, S) > T(I_{u'}, S) \) has positive measure.

We also make the same assumptions on \( S \).

The ordering assumption holds for equation (1), but there are many scenarios where it does not hold.

We will drop the ordering assumption in section 4 when we consider networks with no short loops. Many of the results of this section hold in the absence of the ordering assumption, but the proofs are cleaner with the assumption. We only need it for Corollaries 2, 4, and 5.

The ordering assumption implies that \( T_{\text{out}}(I) \) is an invertible mapping. It also allows us to assume that \( I \) is a scalar quantity ordered such that

\[ I_u \geq I_{u'} \]

\[ \Leftrightarrow T_{\text{out}}(u) \geq T_{\text{out}}(u') \]

\[ \Leftrightarrow T(I_u, S) \geq T(I_{u'}, S) \quad \forall S \]

and further \( I_u > I_{u'} \Leftrightarrow T_{\text{out}}(u) > T_{\text{out}}(u') \). We may assume that \( I \) is a scalar. There is more than one way to represent \( I \) as a scalar, and given one representation, any strictly increasing function of \( I \) with appropriate modification to \( T \) will yield another acceptable representation. We may make similar statements about \( S \). In particular, it will frequently (but not always) be convenient to identify \( I \) with \( T_{\text{out}}(I) \) or \( S \) with \( T_{\text{in}}(S) \).

We now assume that the distribution \( Q_{\text{in}}(T_{\text{in}}) \) is given. Using the ordering assumption we may take \( S = T_{\text{in}} \) without loss of generality. We seek distributions of \( T_{\text{out}} \) which give the maximum and minimum probability and size of epidemics.
To make the notation cleaner in our first theorem, we identify $S$ with $T_{in}$ and so we may use $T(I, T_{in})$ in place of $T(I, S)$. We prove this theorem using the ordering assumption, but we will show later that it is not needed.

**Theorem 1.** Assume a given sequence of network families $G_n$ and a given distribution of susceptibilities $Q_{in}(T_{in})$. Assume that the ordering assumption holds and consider a distribution of infectiousness $P_1(I)$ with transmissibility given by $T_1(I, T_{in})$ which is consistent with $Q_{in}(T_{in})$. Let $\phi_{in,1}(V, \tilde{S})$ be as in equation (3). Let $A_1$ and $P_1$ be the corresponding attack rate and epidemic probability. Similarly take another $P_2(I)$, $T_2(I, T_{in})$, $\phi_{in,2}(V, \tilde{S})$ also consistent with $Q_{in}(T_{in})$ with corresponding $A_2$ and $P_2$.

Assume that $\phi_{in,1}(V, \tilde{S}) \leq \phi_{in,2}(V, \tilde{S})$ for all $V$ and $\tilde{S}$. Then $A_1 \geq A_2$ and $P_1 \geq P_2$.

**Proof.** Let $d \geq 0$ be given.

We choose a network $G$ from $G_n$, $n \geq d$. We will show that a node in an EPN created from $G$ using the first distribution is more likely to be in $H_{out}(d)$ than a node in an EPN created using the second distribution.

Choose any node $u$ from $G$. Partition the nodes of $G$ into disjoint sets $\{u\}$, $U_1$, and $U_2$. To the nodes in $U_1$ we assign $I$ from $P_1(I)$ and to the nodes in $U_2$ we assign $I$ from $P_2(I)$. We assign $T_{in}$ to all nodes from $Q_{in}(T_{in})$. We will consider the effects of adding $u$ to $U_1$ versus adding it to $U_2$.

Consider a partial EPN $\mathcal{E}$ created by assigning edges $(w, v)$ from all $w \neq u$, using $T_1(I_w, T_1(v))$ if $w \in U_1$ and $T_2(I_w, T_1(v))$ if $w \in U_2$. Now consider an arbitrary node $u'$ which may be $u$ which is not already in $H_{in}(d)$, but which could join $H_{in}(d)$ if the appropriate edges were added from $u$. Consider the set of directed paths $R$ from $u'$ to $u$. Let $V$ be the set of neighbors $v$ of $u$ for which adding the edge $(u, v)$ would allow a path in $R$ to be extended to a path of length $d$.

We consider extensions of $\mathcal{E}$ formed by placing $u$ into $U_1$ or $U_2$. The probability that $u'$ would be in $H_{in}(d)$ in the extended EPN is equal to the probability that $u$ has at least one edge to some node in $V$. By our assumption $\phi_{in,1}(V, \tilde{S}) \leq \phi_{in,2}(V, \tilde{S})$, this probability is at least as high if $u \in U_1$ as if $u \in U_2$. Consequently the probability of $u'$ to be in $H_{in}(d)$ is maximal if $u \in U_1$. Repeating the process for all nodes shows that $P_d(G)$ is largest if all nodes are in $U_1$.

We now show that $u \in U_1$ increases $A_d$ compared with $u \in U_2$. We can prove that placing $u$ in $U_1$ versus $U_2$ can only increase the probability of a node to be at the end of a length $d$ path. The proof proceeds as above, except that path directions are reversed.

Thus placing $u$ in $U_1$ results in larger values of $P_d(G)$ and $A_d(G)$ as compared to placing it in $U_2$. By induction it follows that if $U_1$ contains all nodes of $G$, then $P_d(G)$ and $A_d(G)$ are at least as large as if $U_2$ contains all nodes of $G$. It follows then that $P_d(G_n)$ and $A_d(G_n)$ are increased if all nodes are given $I$ from $P_1(I)$. Consequently the inequalities must hold in the limit $d \rightarrow \infty$.

**Lemma 1.** Given $T_{uv} = T(I_u, S_v)$, if we interchange the roles of infectiousness and susceptibility so that $T_{uv} = T(I_v, S_u)$ then $P$ and $A$ interchange values.

**Proof.** If we replace $T_{uv} = T(I_u, S_v)$ with $\bar{T}_{uv} = T(I_v, S_u)$ then the new EPNs correspond to reversing the direction of edges in the original EPNs. Since reversing the direction of edges in an EPN interchanges $H_{in}(d)$ and $H_{out}(d)$ this interchanges $P$ and $A$. 
Corollary 1. Under the same assumptions as Theorem 1, but with $Q_{out}(T_{out})$ fixed and $P(S)$ varying, if $\phi_{out}$ replaces $\phi_{in}$ then $P_1 \geq P_2$ and $A_1 \geq A_2$.

Proof. This follows immediately from Lemma 1.

Corollary 2. Let $Q_{in}(T_{in})$ be given. Assume that the ordering assumption holds. $P$ and $A$ are maximized when $T(I,T_{in}) = T_{in}$.

Proof. From Chebyshev’s “other” inequality [16], if $h_1$ and $h_2$ are decreasing functions,

$$E[h_1(x)h_2(x)] \leq E[h_1(x)]E[h_2(x)]$$

By induction $E[\prod h_j(x)] \geq \prod E[h_j(x)]$.

If we set $h_j(I) = 1 - T(I,T_{in}(v_j))$ then $h_j$ is a decreasing function of $I$. Consequently for any distribution we have

$$\phi_{in}(V,\bar{S}) = E\left[\prod_{v \in V} h_j(I)\right] \geq \prod_{v \in V} 1 - T_{in}(v)$$

with equality if $T(I,T_{in}) = T_{in}$. Thus by Theorem 1, $A$ and $P$ are maximal.

Corollary 3. Let $Q_{in}(T_{in})$ be given. Assume that the ordering assumption holds. Take $I$ to be chosen uniformly from $[0,1]$. Take

$$T(I,T_{in}) = \begin{cases} 0 & T_{in} < I \\ 1 & T_{in} > I \end{cases}$$

(5)

This minimizes $P$ and $A$.

Proof. Given this $T$, we have $\phi_{in}(V,\bar{S}) = \min_{v \in V} \{1 - T_{in}(v)\}$.

We need to show for arbitrary $T$ that $\phi_{in}(V,\bar{S}) \leq \min_{v \in V} \{1 - T_{in}(v)\}$. To do this, let $T$ be given, $T_{in}$ assigned to $v_1, \ldots, v_n$ and assume $v_1, \ldots, v_n$ are ordered such that $T_{in}(v_1) \geq T_{in}(v_2) \geq \cdots \geq T_{in}(v_n)$.

Then

$$\phi_{in}(V,\bar{S}) = \int \prod_{j=1}^n (1 - T(I,T_{in}(v_j)))P(I)dI$$

$$\leq \int [1 - T(I,T_{in}(v_1))]P(I)dI$$

$$\leq 1 - T_{in}(v_1)$$

This shows that for any $T$, $\phi_{in}(V,\bar{S})$ is at most the value it takes for (5). Thus (5) gives a lower bound for $P$ and $A$.

In this case $T_{out}(I) = 1 - \int_0^I Q_{in}(T_{in})dT_{in}$ is the complementary cumulative distribution function of $Q_{in}(T_{in})$.

Corollary 4. Let $Q_{out}(T_{out})$ be given. Assume that the ordering assumption holds.
• If 

\[ T(T_{\text{out}}, S) = T_{\text{out}} \]

Then \( P \) and \( A \) are maximized.

• If \( S \) is chosen uniformly in \([0, 1]\) and

\[ T(T_{\text{out}}, S) = \begin{cases} 0 & T_{\text{out}} < S, \\ 1 & T_{\text{out}} > S \end{cases} \]

Then \( P \) and \( A \) are minimized.

Proof. This follows immediately from Lemma 1 with Corollaries 2 and 3.

**Corollary 5.** Let \( (T) \) be given. Under the ordering assumption, the maximum of \( P \) and \( A \) occur when \( T_{uv} = (T) \) for all pairs.

Proof. Consider any \( Q_{\text{in}} \) and \( Q_{\text{out}} \) which are consistent. Replacing \( Q_{\text{in}} \) with \( \delta(T_{\text{in}} - (T)) \) can only increase \( P \) and \( A \). A further replacement of \( Q_{\text{out}} \) with \( \delta(T_{\text{in}} - (T)) \) can again only increase \( P \) and \( A \), but then \( T_{uv} = (T) \) for all pairs \( u \) and \( v \). Thus homogeneous transmissibility gives an upper bound on \( P \) and \( A \).

We now finish with a conjecture on global lower bounds.

**Conjecture 1.** Under the ordering assumption, the minimum of \( P \) occurs when \( Q_{\text{in}}(T_{\text{in}}) = \delta(T_{\text{in}} - (T)) \) and \( Q_{\text{out}}(T_{\text{out}}) = (T) \delta(T_{\text{out}} - 1) + (1 - (T))\delta(T_{\text{out}}) \).

The minimum of \( A \) occurs when \( Q_{\text{out}}(T_{\text{out}}) = \delta(T_{\text{out}} - (T)) \) and \( Q_{\text{in}}(T_{\text{in}}) = (T) \delta(T_{\text{in}} - 1) + (1 - (T))\delta(T_{\text{in}}) \).

### 4. Bounds in unclustered networks

When we study networks with no short cycles we are able to prove stronger results. We find that \( P \) depends on the network and \( Q_{\text{out}}(T_{\text{out}}) \) only, while \( A \) depends on the network and \( Q_{\text{in}}(T_{\text{in}}) \) only. We can prove global upper and (unlike in the general case) lower bounds on \( P \) and \( A \).

**Assumption 2.** (Unclustered Assumption.) Given a sequence of network families as in equation (2), we assume that any \( G \in \mathcal{G}_n \) has girth greater than \( 2n \).

This assumption means that \( S_d \) chosen from any \( \mathcal{G}_n \) with \( n \geq d \) must be cycle-free. Consequently between any two nodes there is at most one path in \( S_d \). In particular, there is no alternate path between a node and a neighbor. By contrast, in a network with short paths, it is possible (if the ordering assumption fails) that the non-existence of edge \((u, v)\) in the EPN is positively correlated with the existence of an alternate short path from \( u \) to \( v \). It is because of this that we were forced to use the ordering assumption earlier. This cannot happen under the unclustered assumption, so we can abandon the ordering assumption. When we drop the ordering assumption, we can no longer assume that \( T_{\text{in}} \) and \( T_{\text{out}} \) are invertible mappings. The unclustered assumption also allows us to use \( \psi_{\text{in}}(V) \) rather than \( \phi_{\text{in}}(V, S) \).

**Theorem 2.** Let the sequence of \( \mathcal{G}_n \) satisfy the unclustered assumption. Take \( P_1(\mathcal{I}), P_1(S), \) and \( T_1(\mathcal{I}, S) \). Let \( \psi_{\text{in},1}(V) \) be as in equation (4). Similarly take \( P_2(\mathcal{I}), P_2(S), T_2(\mathcal{I}, S), \) and \( \psi_{\text{in},2}(V) \). If \( \psi_{\text{in},1}(V) \leq \psi_{\text{in},2}(V) \), then \( P_1 \geq P_2 \).
Note that we do not make any assumptions on $Q_{in}(T_{in})$.

**Proof.** The proof is similar to that of Theorem 1.

Let $d \geq 0$ be given. Take $G \in \mathcal{G}_n$, $n \geq d$. Choose a node $u$ from $G$ and partition the nodes of $G$ into $\{u\}, U_1$, and $U_2$. To the nodes in $U_1$ we assign $\mathcal{I}$ from $P_1(\mathcal{I})$ and to the nodes of $U_2$ we assign $\mathcal{I}$ from $P_2(\mathcal{I})$. To each node $w$ (including $u$) we assign $S_{w,1}$ and $S_{w,2}$ such that $S_{w,1}$ comes from $P_1(S)$ and $S_{w,2}$ from $P_2(S)$.

We create a partial EPN $\mathcal{E}$ as follows. For each $v \in U_1$ we assign edges $(v, w)$ using $T_1(\mathcal{I}_v, S_{w,1})$ and for $v \in U_2$ we assign them using $T_2(\mathcal{I}_v, S_{w,2})$. Consider any $u'$ not in $H_{in}(d)$ with a path $r$ of length $d' < d$ to $u$ in $\mathcal{E}$ [there is at most one such path]. Consider the set of nodes $V$ which are neighbors of $u$ for which adding an edge from $u$ to $v \in V$ would extend $r$ to a path of length $d$. By assumption $\psi_{in,1}(V) \leq \psi_{in,2}(V)$ and so the probability is greatest if $\mathcal{I}_u$ is chosen from $P_1(\mathcal{I})$. It follows that $\mathcal{A}_1 \geq \mathcal{A}_2$.

**Corollary 6.** If the assumptions of Theorem 2 hold except that $\psi_{out}$ replaces $\psi_{in}$, then $\mathcal{A}_1 \geq \mathcal{A}_2$.

**Proof.** This follows immediately from Lemma 1.

**Theorem 3.** Let the sequence of $\mathcal{G}_n$ satisfy the unclustered assumption. Let $Q_{in}(T_{in})$ be fixed. Then $\mathcal{A}$ is fixed.

**Proof.** We follow the technique used to prove $\mathcal{A}$ is larger for one distribution than the other in Theorem 1. However, in following that proof, because of the lack of clustering, $|V| = 1$. Since for any distribution, $\psi(V) = 1 - \langle T \rangle$ when $|V| = 1$, all distributions must give the same $\mathcal{A}$.

**Corollary 7.** If the assumptions of Theorem 3 hold except that $Q_{out}$ is fixed rather than $Q_{in}$, then $\mathcal{A}_1 = \mathcal{A}_2$.

**Proof.** This follows immediately from Lemma 1.

**Corollary 8.** Let $Q_{in}$ be given. $\mathcal{P}$ is maximized when $T(\mathcal{I}, \mathcal{S}) = T_{in}(\mathcal{S})$.

The proof for this is different from Corollary 2 because we no longer have ordering.

**Proof.** We have

$$\psi_{in}(V) = \int (1 - T_{out})^{|V|} Q_{out}(T_{out}) dT_{out}$$

$(1 - T_{out})^{|V|}$ is a convex function and by Jensen’s inequality, $\psi_{in}$ takes its minimum value if $T_{out} = \langle T \rangle$.

**Corollary 9.** Let $Q_{in}$ be given. $\mathcal{P}$ is minimized when $\mathcal{I}$ is chosen uniformly from $[0, 1]$ and

$$T(\mathcal{I}, \mathcal{S}) = \begin{cases} 0 & \mathcal{I} > T_{in}(\mathcal{S}) \\ 1 & \mathcal{I} < T_{in}(\mathcal{S}) \end{cases}$$

**Proof.** We have

$$\psi_{in}(V) = \int \phi_{in}(V, \mathcal{S}) P(\mathcal{S}) d\mathcal{S}$$
Following the proof of Corollary 3 we may show that $\phi_{in}$ takes its maximum value (subject to $Q_{in}$) exactly when these assumptions hold. Thus $\psi_{in}$ takes its maximum value exactly when these assumptions hold.

**Corollary 10.** Let $Q_{out}$ be given.

- $A$ is maximized when $T(I, S) = T_{out}(I)$.
- $A$ is minimized when $S$ is chosen uniformly from $[0, 1]$ and

$$T(I, S) = \begin{cases} 0 & S > T_{out}(I) \\ 1 & S < T_{out}(I) \end{cases}$$

**Proof.** Identical to the proof of Corollary 4.

We finally end with results on global upper and lower bounds.

**Lemma 2.** Let $f(x)$ be a convex function on $[0, 1]$ and let $\rho(x)$ be a probability density function on $[0, 1]$, with expected value $\rho_0$. Then

$$\int f(x) \rho(x) dx \leq (1 - \rho_0) f(0) + \rho_0 f(1)$$

**Proof.** Jensen’s inequality shows that for a given $x$,

$$f(x) \rho(x) \leq (1 - \rho(x)) f(0) + \rho(x) f(1)$$

Thus

$$\int f(x) \rho(x) dx \leq \int (1 - \rho(x)) f(0) + \rho(x) f(1) dx \leq (1 - \rho_0) f(0) + \rho_0 f(1)$$

**Corollary 11.** Let $\langle T \rangle$ be given.

- The upper bound for both $P$ and $A$ occurs when $T_{uv} = \langle T \rangle$ for all pairs.
- The lower bound for $P$ occurs when $Q_{out}(T_{out}) = \langle T \rangle \delta(T_{out} - 1) + (1 - \langle T \rangle) \delta(T_{out})$.
- The lower bound for $A$ occurs when $Q_{in}(T_{in}) = \langle T \rangle \delta(T_{in} - 1) + (1 - \langle T \rangle) \delta(T_{in})$.

**Proof.** The proof of the upper bound is identical to that of Corollary 5. We prove the lower bound for $P$. The lower bound for $A$ will follow from Lemma 1. From Corollary 9, we know that the lower bound for $P$ given $Q_{out}$ occurs when

$$T(I, S) = \begin{cases} 0 & I > T_{in}(S) \\ 1 & I < T_{in}(S) \end{cases}$$

Then

$$\psi_{in}(V) = \int (1 - T_{out})^{|V|} Q_{out}(T_{out}) dT_{out}$$

We now seek to find $Q_{out}$ which maximizes $\psi_{in}$.

Since $(1 - T_{out})^{|V|}$ is a convex function, we may apply Lemma 2 with $Q_{out}$ playing the role of $\rho$. The maximum occurs when $Q_{out}(T_{out}) = \langle T \rangle \delta(T_{out} - 1) + (1 - \langle T \rangle) \delta(T_{out})$ and so Theorem 2 finishes the proof.
The results of this section generalize those of [23] where short cycles were also neglected and there were no correlations between degrees of neighbors.

5. Discussion

Although we have proven these results for families of networks, the proofs relied on showing the corresponding results in individual networks. Thus for a given network, our general results hold.

With slight modification, the proof of Theorem 2 will hold on a network with short cycles, without needing the ordering assumption. We define

$$\hat{\phi}_{\text{in}}(V, T_{\text{in}}) = \int_{S} \phi(V, S)P(S|T_{\text{in}})dS$$

Assigning susceptibilities such that the value of $T_{\text{in}}(v)$ measured for each distribution is the same for each $v$, the proof of Theorem 2 applies, with $\hat{\phi}_{\text{in}}$ replacing $\psi_{\text{in}}$. This shows that Theorem 1 applies without needing the ordering assumption. We did not present that proof because it is more technical without offering additional insight. Many of the corollaries also follow without needing the ordering assumption, but the ordering assumption is vital to the proof of Corollary 2 where it allows us to use Chebyshev’s “other” inequality. This corollary almost certainly does not require a condition as restrictive as the ordering assumption, and so we expect these results to apply more widely.

In general we have shown that increasing the variance of the population in- or out-transmissibility reduces $P$ and $A$. This is consistent with earlier results [2, 23]. We have shown under weak assumptions that for given $Q_{\text{in}}$, maximizing the variance of $Q_{\text{out}}$ gives the smallest $P$ and $A$ while minimizing the variance gives the largest $P$ and $A$. This leads us to observe that the global maximum of $P$ and $A$ occurs when transmissibility is homogeneous. However, in a network with no short cycles, $A$ depends only on $Q_{\text{in}}$, so for given $Q_{\text{in}}$, only $P$ can vary. Similarly $P$ depends only on $Q_{\text{out}}$. This leads us to the observation that for unclustered networks the global minimum of $P$ occurs if the variance of $T_{\text{out}}$ is maximal (in which case $T_{\text{in}}$ is homogeneous and $A$ is maximal), while the global minimum of $A$ occurs if the variance of $T_{\text{in}}$ is maximal (in which case $T_{\text{out}}$ is homogeneous and $P$ is maximal). We believe that these global minima hold in general for all networks.

These results provide a general rule for epidemic interventions. Given two strategies which have the same average impact on transmissibility, the one which is most heterogeneous is preferable. In particular, if the goal is to reduce the probability of an epidemic, the one which has the most heterogeneous impact on out-transmissibility is preferable, while if the goal is to reduce the size, the one which has the most heterogeneous impact on in-transmissibility is preferable.

Appendix A. Stochastic epidemic models

We give a brief argument showing that stochastic epidemic models may be studied in the context of random graph epidemic models. More complete arguments are found in [13, 25].

We consider a stochastic model which proceeds as follows: an infected individual $u$ makes contacts with others in a Poisson process. If the contacted individual $v$ is
susceptible, then $v$ becomes infected. After some period of time $u$ recovers.

This simple model may be extended in a number of ways. Let us assume that there are several groups $C_1, \ldots, C_n$ in the population, and $u \in C_i$ makes contacts with nodes in $C_j$ at some rate $\lambda_{ij}$ [normalized such that the probability $u \in C_i$ contacts $v \in C_j$ is the same as the probability $v$ contacts $u$]. Once contact from $u$ to $v$ occurs, if $v$ is susceptible then infection happens probability $T_{uv}$.

We may instead generate networks along the lines of Erdős–Rényi networks. We place each node into some $C_i$. For each $u \in C_i$ and $v \in C_j$, we place a $\{u, v\}$ edge with probability equal to the probability that $u$ and $v$ contact one another. We may then generate the EPN by placing a directed edge from $u$ to $v$ with probability $T_{uv}$ and a directed edge from $v$ to $u$ with probability $T_{vu}$.

As long as a new network is chosen for each simulation, this process will generate the same outbreak distributions as the stochastic process. Consequently, a wide class of stochastic models are equivalent to a wide class of network models.

References


