Network-based models of transmission of infectious diseases: a brief overview*

Winfried Just† Hannah Callender‡ M. Drew LaMar§

May 2, 2015

The main purpose of this document is to give a brief but mathematically rigorous description of the network-based models of transmission of infectious diseases that are studied on this web site[1]. Readers will be able to find a much more detailed development of this material in our book chapters [5, 6]. We also briefly describe how the network-based models that are defined here are related to compartment-level models that are used in most on the literature on mathematical epidemiology.

1 Network-based models and their parameters

1.1 Introduction

Epidemiology studies the spread of diseases caused by *pathogens,* such as viruses or bacteria, in *populations* of *hosts,* which can be humans, animals, or plants. The goal is to *predict* the time course of an *outbreak* of a given disease in a population and the effect of conceivable *control measures,* such as vaccination, quarantine, culling, or behavior modification on the severity of the outbreak. Of particular importance is the question of how such control measures would affect the *final size* of the outbreak, that is, the proportion of hosts in the population who will eventually experience infection.

Mathematical models are greatly simplified representations of reality. Based on implementation of such models in computer code, epidemiologists can *simulate* hypothetical outbreaks under various assumptions about control that might possibly be implemented and derive predictions about their likely effectiveness. We designed a software tool IONTW that allows for studying this type of questions.

The models embodied in IONTW are *agent-based network models* for diseases that are transmitted by *direct contact* between two hosts. This leaves out vector-born infectious

†Department of Mathematics, Ohio University, Athens, OH 45701 E-mail: mathjust@gmail.com

‡University of Portland E-mail: callende@up.edu

§The College of William and Mary E-mail: drew.lamar@gmail.com

[1] `http://www.ohio.edu/people/just/IONTW`

diseases such as malaria and diseases such as cholera that are transmitted through shared environmental resources.

When making inferences from mathematical models to the spread of infections in populations of actual hosts, one needs to carefully examine how the assumptions of the models influence their predictions. IONTW allows us to study how the simplifying assumptions about the contact pattern that are embodied in the *contact network* influence the outcomes of simulations. It also can be used for exploring the *sensitivity* of these outcomes with respect to the other model parameters.

1.2 Basics of disease transmission modeling

A standing assumption of our modeling is that we investigate the spread of one given infectious disease. We ignore *demographics* (births, deaths from causes that are unrelated to the disease, immigration, and emigration) and assume a fixed population of hosts that are numbered from 1 to N. At any given time t, host i can be in one of the following states:

- S: *Susceptible* to infection.
- E: *Exposed* to the pathogens, but not yet infectious.
- I: *Infectious,* that is, able to infect other hosts through direct contact.
- R: *Removed,* that is, not infectious and immune to subsequent infection (through vaccination or recovery from the disease). Hosts who have died from the disease or been subjected to culling or quarantine also are in this state.

The sets of hosts that are in states S, E, I, R will be referred to as the ***S**-compartment,* the ***E**-compartment,* the ***I**-compartment,* and the ***R**-compartment,* respectively. Membership in the compartments changes over time, and one can conceptualize the time course of an outbreak as *movement of hosts between compartments.*

We use the following notation:

- At time T_i^E host i moves from the **S**-compartment into the **E**-compartment. We refer to T_i^E as the *time of exposure* of host i.
- At time T_i^I host i moves into the **I**-compartment. We refer to T_i^I as the *onset of infectiousness* of host i.
- At time T_i^R host i moves out of the **I**-compartment. We refer to T_i^R as the *cessation of infectiousness* of host i.
- The time interval $[T_i^E, T_i^I)$ is called the *latent period* of host i. Its length $T_i^I - T_i^E$ is denoted by τ_i^E and called the *duration of latency of host i.*
- The time interval $[T_i^I, T_i^R)$ is called the *period of infectiousness* of host i. Its length $T_i^R - T_i^I$ is denoted by τ_i^I and called the *duration of infectiousness of host i.*

Expressions like $T_i^E = t$ specify *transition events* as they tell us *what* happened *when* to *which host.* One standing assumption is that for every transition event $T_i^E = t_0$ there is a transition event $T_i^I = t_1$ with $t_0 < t_1$. In other words, we assume that each exposed host *will* eventually become infectious. In all models that can be implemented in version 1.0 of IONTW it is assumed that removal is permanent so that hosts never can leave the **R**-compartment.

Further restrictions on the transition times and transition events are imposed by the particular *model type.* In *SEIR-models* we assume that for every transition event $T_i^I = t_1$ there is a transition event $T_i^R = t_2$ with $t_1 < t_2 < \infty$ so that all hosts who become infectious will subsequently be removed. Figure 1 schematically depicts this type of models.

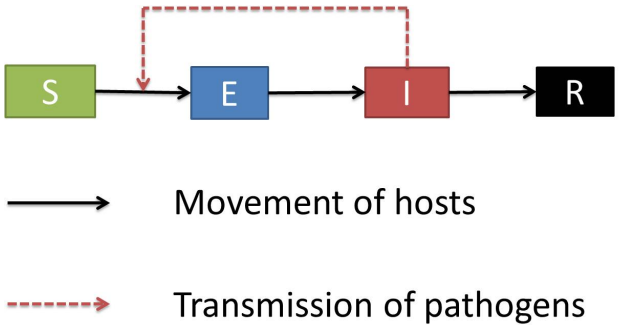

Figure 1: Schematic representation of *SEIR*-models.

The same assumption is made in *SIR-models* as schematically depicted by Figure 2. However, in *SIR*-models we ignore the latent period and assume that at time T_i^I host i directly moves from the **S**-compartment to the **I**-compartment. Both *SEIR*- and *SIR*-models are suitable for the study of *immunizing infections,* such as measles, chicken pox, or the seasonal flu.

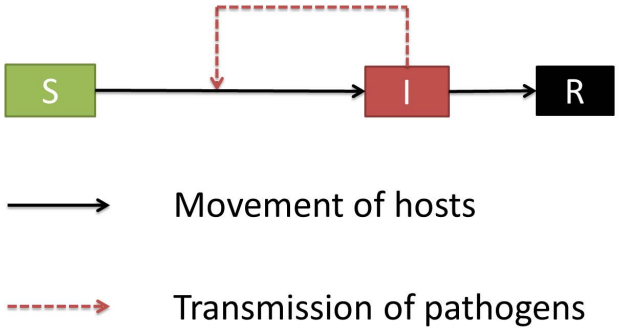

Figure 2: Schematic representation of *SIR*-models.

For diseases such as HIV-infection from which hosts never recover one might use *SI-models* as schematically depicted by Figure 3. In these models there is no **R**-compartment, which could be mathematically expressed by setting $T_i^R = \infty$. Addition of an **E**-compartment to an SI-model will give an SEI-model.

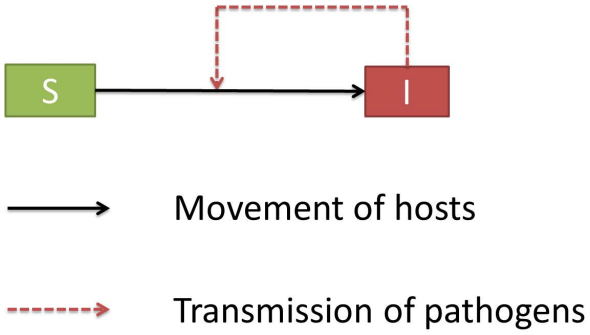

Figure 3: Schematic representation of SIR-models.

For diseases such as Gonorrhea that do not confer immunity upon recovery one might use SIS-*models* as schematically depicted by Figure 4. In these models, at time T_i^R host i would move back to the **S**-compartment. Note that while in $SEIR$-, SIR-, and SI-models there can be at most one time $T_i^I = t_1$ and at most one time $T_i^R = t_2$ for each host i, in SIS-models there could be multiple such times. Our notation allows for this ambiguity if we treat $T_i^I = t_1$ and $T_i^R = t_2$ as symbolic expressions rather than equations that uniquely specify a time. Addition of an **E**-compartment to an SIS-model will give an $SEIS$-model.

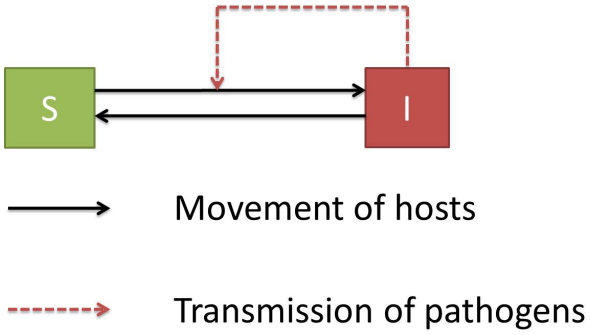

Figure 4: Schematic representation of SIS-models.

Time in our models could be either continuous or discrete. In the former case, we can assume that at most one state transition event will happen at any given time, while in the latter case multiple such events may occur simultaneously.

Transition of host i from the **S**-compartment into the **E**-compartment at time $t = T_i^E$ (in an $SEIR$-, SEI-, or $SEIS$-model) or from the **S**-compartment into the **I**-compartment

at time $t = T_i^I$ (in an SIR-, SI-, or SIS-model) requires a *successful* (from the point of view of the pathogen) contact with another host j at time t, which is defined as a contact during which a sufficient number of pathogens will be transferred from host j to host i. A successful contact at time t requires that host i is susceptible and host j infectious at time t. Thus the probability of a successful contact will depend on the state of the population and will change over time.

An *effective* contact between hosts i and j is a contact that *would* be successful *if* host i were susceptible and host j infectious at the time of the contact. We assume that for any given pair (i, j) of hosts the probability of at least one effective contact over a time interval of length Δt remains fixed, and hosts make contacts independently. In discrete-time models this probability is denoted by $b_{i,j}$, and Δt is taken as the physical time that corresponds to one time step in the model. In continuous-time models we assume that hosts i and j make contacts at a fixed rate $\beta_{i,j}$.

Similarly, in discrete-time models we assume that the probability that a host who is infectious at time t will cease to be infectious at time $t+1$ (Δt units of physical time later) with probability a, which is fixed and the same for all hosts. In continuous-time models we assume that hosts will leave the **I**-compartment at a fixed rate α. Similarly, in discrete-time models with **E**-compartments we assume that the probability that a host who is in state E at time t will be infectious at time $t+1$ (Δt units of physical time later) with probability el, which is fixed and the same for all hosts. In continuous-time models with **E**-compartments we assume that hosts will leave the **E**-compartment at a fixed rate γ.

The mean duration of infectiousness $\langle \tau^I \rangle$ and the mean duration of latency $\langle \tau^E \rangle$ can be calculated from these parameters as follows:

$$\begin{aligned}
\langle \tau^I \rangle &= \frac{\Delta t}{a} && \text{in discrete-time models,} \\
\langle \tau^I \rangle &= \frac{1}{\alpha} && \text{in continuous-time models,} \\
\langle \tau^E \rangle &= \frac{\Delta t}{el} && \text{in discrete-time models,} \\
\langle \tau^E \rangle &= \frac{1}{\gamma} && \text{in continuous-time models.}
\end{aligned} \tag{1}$$

For a given continuous-time model one can build a *discrete-time approximation* by choosing an appropriate unit Δt of physical time that corresponds to one step of the discrete-time model and using the following conversions of rates into probabilities:

$$\begin{aligned}
b_{i,j} &= 1 - e^{-\beta_{i,j}\Delta t}, \\
a &= 1 - e^{-\alpha \Delta t}, \\
el &= 1 - e^{-\gamma \Delta t}.
\end{aligned} \tag{2}$$

For actual diseases, the durations of infectiousness for individual hosts tend to be normally distributed instead of exponentially distributed as they are in our continuous-time

models. One might get more realistic discrete-time SIR- or SIR-models by using $\Delta t = \langle \tau^I \rangle$ and setting $a = 1$ instead of using the conversion (2). We call such discrete-time models where hosts stay infectious for exactly one time step *next-generation models.* One can treat them as approximations to the spread of the infection in time, or, alternatively, as models for the evolution of so-called *generations of the infection.* Generation $Gen(0)$ comprises all initially infected hosts, and generation $Gen(n+1)$ comprises all hosts that became infected by a host in generation $Gen(n)$. In real outbreaks, the order which hosts become infectious is positively correlated with, but usually not identical with the numbering of the generations to which they belong. Thus the variable n does not represent physical time when we consider generations of the infection.

In most of the modules and exercises at this web site, we will consider outbreaks that start at time 0 with exactly one infectious host j^* (the so-called *index case*), so that $Gen(0) = \{j^*\}$. If all other hosts are susceptible at time 0, we refer to such an initial state as *introduction of one index case into an otherwise susceptible population.*

1.3 Network-based models

We assume that we are given a simple graph G with $V(G) = \{1, \ldots, N\}$ that represents the *contact network* of the hosts in the population. Effective contacts can occur *only* between hosts i, j for which $\{i, j\} \in E(G)$, that is, between hosts that are represented by *adjacent* nodes in the graph G. If such an edge is present, then we assume that the frequency of contact does not depend on the particular hosts i, j that are connected by the edge. In the context of a discrete-time model, this means that

$$b_{i,j} = \begin{cases} b & \text{if } \{i,j\} \in E(G), \\ 0 & \text{if } \{i,j\} \notin E(G), \end{cases} \tag{3}$$

where b is a fixed parameter of the model. In the context of continuous-time models we get

$$\beta_{i,j} = \begin{cases} \beta & \text{if } \{i,j\} \in E(G), \\ 0 & \text{if } \{i,j\} \notin E(G), \end{cases} \tag{4}$$

where β is a fixed parameter that represents the rate at which adjacent hosts make contact. Since the graph G was assumed simple, $\{i, i\} \notin E(G)$ and $b_{i,i} = \beta_{i,i} = 0$ for each host i.

Thus our network-based models are defined in terms of the following parameters:

- The population size N,
- the contact network G, and
  - the transition probabilities a, b, el and the length Δt of the time step (for discrete-time models) or
  - the transition rates α, β, γ (for continuous-time models).

The Reference Guide to IONTW gives details about how these parameters are controlled by the interface of the software and which types of networks can be explored with IONTW.

1.4 Modeling disease transmission on random contact networks

For transmission of most diseases in large real communities it is usually impossible to construct the actual relevant contact network. However, it is often feasible to estimate some parameters of the network, such as the mean degree $\langle k \rangle$ by collecting partial data on the network. One can then choose the graphs G for our network-based models *randomly* so that with high probability it will have parameters similar to the ones that have been empirically observed for contact networks of interest.

If we choose several such graphs, we might reasonably expect that they form a *representative sample* of actual contact networks. Essentially this is the same trick that we use whenever we run simulations. During the run of multiple simulations, the computer produces in effect a representative sample of state transition sequences that could occur in a given model, and we can use the outcome of simulations to form reasonably reliable hypotheses about what should happen in real outbreaks. Considering "random" objects (graphs or state transition sequences) is a mathematician's way of dealing with uncertainty. The only novelty is that we are now considering two distinct sources of uncertainty: Uncertainty about the actual contact network, and uncertainty about the actual sequence of state transitions that will happen in a model based on a given network. In order to draw valid conclusions, we will need to take into account both the variability that we see between multiple runs of simulations of a model that is based on a fixed random graph (variability *within* a fixed model) as well as the variability that we observe when we average the results of multiple runs performed for fixed models and compare these averages for multiple draws of random graphs (variability *between* models).

Several of the network options of IONTW give instances of random networks that are drawn from certain probability distributions.

2 Compartment-level models and R_0

2.1 Compartment-level models

Many mathematical models of disease transmission are based on the assumptions of *homogeneity of hosts* and *uniform mixing.* In our terminology, the first of these assumptions boils down to assuming that the parameters el, a or γ, α that govern the transition from the **E**-compartment into the **I**-compartment or from the **I**-compartment into the **R**-compartment are identical for each host. We make this assumption in all our models.

In contrast, the uniform mixing assumption boils down to assuming that the parameters $b_{i,j}$ or $\beta_{i,j}$ that give the probabilities or rates of making effective contact are identical for each pair (i, j) of hosts with $i \neq j$. In our network-based models, this assumption is satisfied if, and only if, the contact network is the complete graph K_N with N nodes that contains all possible edges $\{i, j\}$. Thus our software IONTW allows us to investigate models that use the uniform mixing assumption. We can then explore to what extent predictions of models that were constructed under the this simplifying assumption remain valid for more realistic contact networks.

When both assumptions of homogeneity of hosts and uniform mixing are made, hosts lose all individual characteristics, and the state of the model at time t can be conceptualized as the vector $(|\mathbf{S}(t)|, |\mathbf{E}(t)|, |\mathbf{I}(t)|, |\mathbf{R}(t)|)$ of numbers[2] of hosts in the compartments. The distribution of future states of the model is then entirely determined by the current state. Models that operate entirely on the level of these counts or their expected values are usually called *compartment-based models.* However, we prefer the phrase *compartment-level models.* All models of disease transmission use compartments, but the models that we defined in the previous section distinguish between individual hosts, while models based on summary information for the compartments do not.

2.2 The basic reproductive ratio R_0

Consider an initial state that corresponds to introduction of an index case j^* into an otherwise susceptible population. The number of *secondary infections* caused by the index case j^* is a r.v. Its mean value is the most important parameter in disease modeling.

Definition 1 *The* basic reproductive ratio *or* basic reproductive number R_0 *is the mean number of secondary infections caused by an average index case in a large entirely susceptible population.*

In continuous-time network-based models as defined in Section 1, the value of R_0 can be approximated as follows:

$$R_0 \approx \frac{\beta}{\alpha + \beta}\langle k \rangle. \tag{5}$$

In discrete-time network-based models we could use the approximation:

$$R_0 \approx \frac{b}{a + b - ab}\langle k \rangle. \tag{6}$$

In both (5) and (6), the symbol $\langle k \rangle$ stands for the mean degree of the contact network. IONTW uses these approximations for calculating R_0 with its **Metrics** function. The derivation of (5) and (6) can be found in Module 5 of the online appendix to [5].

For large network sizes the mean degree of the complete graph K_N is $\langle k \rangle = N - 1 \approx N$. Moreover, in this case realistic models of actual diseases would have parameters $\beta \ll \alpha$ or $b \ll a$. For small values of Δt, in view of (1) and (2), this leads to the following approximation (5) and (6) that is often used in the literature on compartment-level models.

$$R_0 \approx \frac{\beta N}{\alpha} = \beta \langle \tau^I \rangle N \approx \frac{(1 - e^{-\beta \Delta t})\langle \tau^I \rangle N}{\Delta t} \approx \frac{b \langle \tau^I \rangle N}{\Delta t}. \tag{7}$$

where the last two approximations are valid for sufficiently small Δt.

[2]The symbol $|A|$ denotes the size of a set A here.

2.3 Some predictions of compartment-level models

In compartment-level models, the value of R_0 entirely determines whether there is a positive probability of *major outbreaks* that will affect a significant fraction of the population. More precisely, if $R_0 \leq 1$, then for every given probability p there exists a finite bound $B(p)$ that does not depend on the population size N, such that with probability at least p, no more than $B(p)$ hosts will experience infection during and outbreak that is caused by introduction of an index case j^* into an otherwise susceptible population. We say that in this case all outbreaks will be *minor*. Note, however, that in a relatively small population the fraction $\frac{B(p)}{N}$ may still represent a significant proportion of the population.

In contrast, if $R_0 > 1$ and we are working with an SIR- or $SEIR$-model, then there are numbers $r(\infty)$ and z_∞ that satisfy the inequalities $0 < r(\infty), z_\infty < 1$ such that as long as the population size is large, with probability very close to $1 - z_\infty$, introduction of a single index case into an otherwise susceptible population will result in a major outbreak with final size close to $r(\infty)$.

The number z_∞ represents the probability of a minor outbreak for very large population sizes. The notation can be understood by thinking of z_∞ as the probability that a near zero proportion of hosts in a near infinite population will experience infection during the outbreak. One important piece of information here is that $z_\infty < 1$, so that the probability $1 - z_\infty$ of a *major* outbreak is positive.

The number $r(\infty)$ represents the predicted final size for large populations. For SIR- or $SEIR$-models, this is the proportion $\frac{|\mathbf{R}(t)|}{N}$ for a very large t, after the outbreak has run its course. Mathematicians are fond of thinking of very large times as $t = \infty$, which explains the notation.

Note that $r(\infty)$ gives the approximate proportion only if the observed outbreak was in fact a major one. It is very interesting that $r(\infty)$ is always predicted to be less than 1. This means that even if R_0 is very large, a proportion $1 - r(\infty) > 0$ of hosts is predicted to escape infection.

For a given modeling approach, the values of z_∞ and r_∞ do not depend on the actual population size N, so that if we run many simulations for a variety of very large population sizes N, we should observe major outbreaks with approximately the same probability $1-z_\infty$. These values may be significantly different though for different values of R_0; the larger R_0, the less likely it is that the outbreak will only be a minor one, and major outbreaks will affect a larger proportion $r(\infty)$ of the population.

Another important prediction of compartment-level models is that vaccination of a proportion of at least HIT randomly chosen hosts will guarantee that all outbreaks are restricted to minor ones. The critical proportion HIT is called the *herd immunity threshold* and is given by

$$HIT = \frac{K}{N} = 1 - \frac{1}{R_0}. \tag{8}$$

3 Some pointers to the literature

The previous subsection only gave a couple of highlights from the vast literature on compartment-level models. In some modules at this web site, we will refer to these results for comparison with the prediction of network-based models. Students who want to learn more about these models have a broad choice of good introductory texts, such as [1, 4, 7, 8]. We particularly recommend the textbook [2] that gives the most comprehensive and treatment. It also uses the same notation as in Subsection 2.3 and contains many excellently structured exercises with hints and worked-out solutions.

Network-based models are covered to a limited extent in some of the introductions that we listed above. A mathematically more advanced treatment can be found in [3]. We recommend our chapter [5] as a detailed elementary introduction into this type of models.

References

[1] Fred Brauer and Carlos Castillo-Chavez. *Mathematical models for communicable diseases*, volume 84. SIAM, 2012.

[2] Odo Diekmann, Hans Heesterbeek, and Tom Britton. *Mathematical tools for understanding infectious disease dynamics*. Princeton University Press, 2012.

[3] Moez Draief and Laurent Massoulié. *Epidemics and Rumors in Complex Networks*, volume 369 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 2010.

[4] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

[5] Winfried Just, Hannah Callender, and M Drew LaMar. Disease transmission dynamics on networks: Network structure *vs.* disease dynamics. In Raina Robeva, editor, *Algebraic and Discrete Mathematical Methods for Modern Biology*. Academic Press, 2015.

[6] Winfried Just, Hannah Callender, M Drew LaMar, and Natalia Toporikova. Transmission of infectious diseases: Data, models, and simulations. In Raina Robeva, editor, *Algebraic and Discrete Mathematical Methods for Modern Biology*. Academic Press, 2015.

[7] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2008.

[8] Emilia Vynnycky and Richard White. *An introduction to infectious disease modelling*. Oxford University Press, 2010.