Journal of Theoretical Biology 399 (2016) 13–21

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Discrete-time moment closure models for epidemic spreading in populations of interacting individuals

Mattia Frasca [a,*], Kieran J. Sharkey [b]

[a] DIEEI, Università degli Studi di Catania, Viale A. Doria 6, 95125 Catania, Italy
[b] Department of Mathematical Sciences, University of Liverpool, Peach Street, Liverpool L69 7ZL, United Kingdom

HIGHLIGHTS

- Derivation of a novel set of 'discrete-time moment equations' at the level of individual nodes and pairs of nodes.
- Introduction of appropriate approximations of the joint probabilities appearing in the 'discrete-time moment equations' to close them.
- Formulation of two types of model: one assuming statistical independence at the level of individuals and one at the level of pairs.
- Derivation of a model at the level of the population which captures the behavior of epidemics on homogeneous random networks.
- Validation of the proposed models through numerical simulation over different network topologies.

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form
7 March 2016
Accepted 17 March 2016
Available online 30 March 2016

Keywords:
Epidemics
Mathematical models
SIR processes

ABSTRACT

Understanding the dynamics of spread of infectious diseases between individuals is essential for forecasting the evolution of an epidemic outbreak or for defining intervention policies. The problem is addressed by many approaches including stochastic and deterministic models formulated at diverse scales (individuals, populations) and different levels of detail. Here we consider discrete-time SIR (susceptible–infectious–removed) dynamics propagated on contact networks. We derive a novel set of 'discrete-time moment equations' for the probability of the system states at the level of individual nodes and pairs of nodes. These equations form a set which we close by introducing appropriate approximations of the joint probabilities appearing in them. For the example case of SIR processes, we formulate two types of model, one assuming statistical independence at the level of individuals and one at the level of pairs. From the pair-based model we then derive a model at the level of the population which captures the behavior of epidemics on homogeneous random networks. With respect to their continuous-time counterparts, the models include a larger number of possible transitions from one state to another and joint probabilities with a larger number of individuals. The approach is validated through numerical simulation over different network topologies.

1. Introduction

Epidemic modeling is a continuously evolving field that is increasingly important for understanding the spread of infectious diseases, investigating outbreak scenarios, and identifying prevention and control policies (Pastor-Satorras et al., 2015; Fu et al., 2013). For example, mathematical models of the 2014 West Africa Ebola outbreak have provided valuable quantitative analysis for assessing the risk of international virus diffusion, the impact of travel restrictions, and the effectiveness of intervention strategies (Gomes et al., 2014; Poletto et al., 2014; Merler et al., 2015).

Early models of the spread of infectious diseases were based on deterministic ordinary differential equations (Anderson et al., 1992; Heesterbeek, 2000) using the assumption of homogeneous mixing between individuals in the population (that is, any two individuals are equally likely to interact at any time, Pastor-Satorras et al., 2015), and provided a description of the epidemics at the level of the population. With the introduction of complex networks (Newman, 2003; Boccaletti et al., 2006) into epidemics models, the hypothesis of homogeneous mixing was removed by explicitly incorporating the heterogeneity of the interaction pattern among individuals (Pastor-Satorras and Vespignani, 2001; Lloyd and May, 2001). These models and the related theoretical approaches to understanding their critical properties (Moreno et al., 2002; Newman, 2002; Barthélemy et al., 2004) have been widely studied. The investigation of network-based

* Corresponding author.
E-mail addresses: mfrasca@diees.unict.it (M. Frasca), kjs@liv.ac.uk (K.J. Sharkey).

http://dx.doi.org/10.1016/j.jtbi.2016.03.024

approaches has led to the development of microsimulation models, able to track billions of individuals, and used to perform stochastic simulations of entire populations at the scale of single individuals, by explicitly taking into account the spatial structures and individual heterogeneity that can be inferred from the analysis of available datasets on the structure of human interactions, their mobility and contact patterns (Merler et al., 2011; Gomes et al., 2014).

In parallel to stochastic simulation methods, deterministic representations of epidemic dynamics on networks have also been developed. In scenarios where infection is treated as spreading from individual to individual via a network of contacts, pair approximation methods have proved to be a valuable extension to the classic mean-field methods. These methods have been investigated both at the population level (Matsuda et al., 1992; Keeling, 1999) and at the individual node level (Sharkey, 2008, 2011). Insights into the relationship between microscopic stochastic dynamics and mean-field descriptions are gained through the analysis of these models (Sharkey, 2008).

Although most of the mean-field or pair approximation models are based on continuous time, a few works have dealt with their discrete-time counterpart (Wang et al., 2003; Gómez et al., 2010; Valdano et al., 2015). On the other hand, stochastic discrete-time epidemic models are widely used (Pastor-Satorras et al., 2015; Frasca et al., 2006; Buscarino et al., 2008, 2014), especially in data-driven approaches where information is available at discrete sample times. Consequently, developing deterministic versions of discrete-time models may offer a relevant complementary approach. Furthermore, the dynamics of discrete-time models is typically far richer in behavior than their analogous continuous-time counterparts.

In previous works on deterministic discrete-time models, Markov chains have been employed to model SI (susceptible–infectious) or SIS (susceptible–infectious–susceptible) processes on static contact networks (Wang et al., 2003; Gómez et al., 2010), and then extended to deal with the case of temporal networks in Valdano et al. (2015). However, these works investigate epidemic processes under the assumption of stationarity and also assume the absence of correlations between individual infection probabilities. In our work, we derive the 'discrete-time moment equations' for the probability of the states of an SIR process. A novel feature of these discrete-time equations is that, unlike a standard continuous-time BBGKY-type hierarchy of moment equations, they are expressed in terms of joint probabilities whose degree is governed by the network structure itself via the degrees of individual nodes. To close these equations we introduce appropriate approximations of the joint probabilities that appear. We do this with two different assumptions on statistical independence: one at the level of individuals and one at the level of pairs. We then derive models at the level of the population from the individual-based and pair-based ones by making appropriate homogeneity assumptions (Sharkey, 2008). We validate the approach through numerical evaluation and compare the results with stochastic simulation.

2. Individual-based model

We consider an undirected network $\mathcal{G}=(\mathcal{N},\mathcal{L})$ with N nodes and L edges. Two nodes $i,j\in\mathcal{N}$ are connected only if $(i,j)\in\mathcal{L}$, and self-loops are not allowed, that is, $(i,i)\notin\mathcal{L}$.

Each node of the network represents an individual and a link is a contact between two individuals. In this framework, we consider the terms 'individuals' and 'nodes' to be synonymous. The network is also represented by its adjacency matrix G, where $G_{ij}=1$ if there is contact from individual j to individual i, and $G_{ij}=0$ otherwise.

We focus on the discrete-time SIR model, where each node/individual may assume one of the three possible states denoted as S, I, and R. A susceptible individual may become infected if contacted by an infectious individual with a probability given by T_{ij}. In the case where the transmission rate across each link per unit time is the same and equal to τ, then $T_{ij}=\tau G_{ij}$. An infected individual recovers from the disease with probability γ per unit time.

We denote the probability that the i-th individual is in the susceptible, infectious or recovered state at time t by $\langle S_i\rangle_t$, $\langle I_i\rangle_t$ and $\langle R_i\rangle_t$, respectively. Additionally, for convenience we also introduce the uninfected state U as a state in which the agent is either susceptible or recovered, and denote the probability that the i-th individual is in this state by $\langle U_i\rangle_t$ (by definition $\langle U_i\rangle_t=\langle S_i\rangle_t+\langle R_i\rangle_t$). The discrete-time equations governing the evolution of the state probabilities are:

$$\begin{aligned}\langle S_i\rangle_{t+1}&=\langle S_i\rangle_t-\Pi_{S_i\to I_i}\\ \langle I_i\rangle_{t+1}&=\langle I_i\rangle_t+\Pi_{S_i\to I_i}-\Pi_{I_i\to R_i}\end{aligned}\tag{1}$$

where $\Pi_{S_i\to I_i}$ represents the probability that the i-th individual in the state S becomes infectious and $\Pi_{I_i\to R_i}$ the probability that the i-th individual in the state I recovers from the disease.

To develop expressions for these terms, we need to introduce the subset $\mathcal{N}_i\subseteq\mathcal{N}$ containing the node i and all its first neighbors, and the subset $\mathcal{L}_i\subseteq\mathcal{L}$ containing all of the arcs connecting i to one of its first neighbors. Let us assume that the cardinality of $\mathcal{N}_i$ is m, so in addition to i there are another $m-1$ elements in $\mathcal{N}_i$. To keep the notation simple, let us define a new labelling of the nodes in $\mathcal{N}_i$ such that $J_1=i$ and the other nodes are $J_2,J_3,J_4,\ldots,J_m$. With these definitions, the probability $\Pi_{S_i\to I_i}$ reads:

$$\begin{aligned}\Pi_{S_i\to I_i}=&\langle S_iI_{J_2}U_{J_3}U_{J_4}\ldots U_{J_m}\rangle_t[1-(1-T_{iJ_2})]\\ &+\langle S_iU_{J_2}I_{J_3}U_{J_4}\ldots U_{J_m}\rangle_t[1-(1-T_{iJ_3})]+\cdots\\ &+\langle S_iU_{J_2}U_{J_3}U_{J_4}\ldots I_{J_m}\rangle_t[1-(1-T_{iJ_m})]\\ &+\langle S_iI_{J_2}I_{J_3}U_{J_4}\ldots U_{J_m}\rangle_t[1-(1-T_{iJ_2})(1-T_{iJ_3})]\\ &+\langle S_iI_{J_2}U_{J_3}I_{J_4}\ldots U_{J_m}\rangle_t[1-(1-T_{iJ_2})(1-T_{iJ_4})]+\cdots\\ &+\langle S_iI_{J_2}I_{J_3}\ldots I_{J_m}\rangle_t\left[1-\prod_{h=2}^{m}(1-T_{iJ_h})\right]\end{aligned}\tag{2}$$

where $i\in\{1,2,\ldots,N\}$. We note that $\Pi_{S_i\to I_i}$ is a function of the probabilities of the different possible states of the nodes of $\mathcal{N}_i\backslash\{i\}$ given that node i itself is susceptible. Each term on the right-hand side of (2) expresses the joint probability of the states of m individuals multiplied by the probability that, given that state, i gets infected over the next time step. For example, the term $\langle S_iI_{J_2}U_{J_3}\ldots U_{J_m}\rangle_t$ represents the probability that individual i is susceptible, J_2 is infected, and all the others are uninfected. Under this condition, i can be infected only through contact with J_2. In fact, the term $(1-T_{iJ_2})$ represents the probability that i does not get infected through the link with J_2 and $[1-(1-T_{iJ_2})]$ the probability that it does. Similarly, when contacts with more than one infected individuals are possible, for instance, if $\langle S_iI_{J_2}I_{J_3}U_{J_4}\ldots U_{J_m}\rangle_t\neq 0$, then $(1-T_{iJ_2})(1-T_{iJ_3})$ is the probability that i does not get infected through the link with J_2 or through the link with J_3, and $[1-(1-T_{iJ_2})(1-T_{iJ_3})]$ is the probability that it does.

By contrast the recovery probability $\Pi_{I_i\to R_i}$ does not depend on the state of the neighbors, and is expressed by:

$$\Pi_{I_i\to R_i}=\gamma\langle I_i\rangle_t.\tag{3}$$

Eq. (1) is exact, but not closed. We propose to close it either at the level of individuals or at the level of pairs. The first case is dealt with in this section, while the second one is discussed in Section 3. In the first case, we assume statistical independence at the level of the individual probabilities; that is, we approximate the m-node state (or m-state) probability as:

$$\langle A_iB_{J_2}C_{J_3}D_{J_4}\ldots M_{J_m}\rangle\approx\langle A_i\rangle\langle B_{J_2}\rangle\langle C_{J_3}\rangle\langle D_{J_4}\rangle\ldots\langle M_{J_m}\rangle\tag{4}$$