Physica A 389 (2010) 2663–2674

Contents lists available at ScienceDirect

Physica A

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

Disease dynamics in a dynamic social network

Claire Christensen [a], István Albert [b], Bryan Grenfell [c,d,1], Réka Albert [a,b,c,*]

[a] Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA
[b] The Huck Institutes for the Life Sciences, The Pennsylvania State University, University Park, PA 16802, USA
[c] Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, PA 16802, USA
[d] Fogarty International Center, National Institutes of Health, Bethesda, MD 20892-2220, USA

ARTICLE INFO

Article history:
Received 5 August 2009
Received in revised form 11 December 2009
Available online 26 February 2010

Keywords:
Social network
Contact network
Network dynamics
Disease spreading
Disease dynamics
Measles

ABSTRACT

We develop a framework for simulating a realistic, evolving social network (a city) into which a disease is introduced. We compare our results to prevaccine era measles data for England and Wales, and find that they capture the quantitative and qualitative features of epidemics in populations spanning two orders of magnitude. Our results provide unique insight into how and why the social topology of the contact network influences the propagation of the disease through the population. We argue that network simulation is suitable for concurrently probing contact network dynamics and disease dynamics in ways that prior modeling approaches cannot and it can be extended to the study of less well-documented diseases.

1. Introduction

A question of fundamental importance to epidemiology is determining which characteristics of a population are most salient in dictating the manner in which a disease will spread through that population. That is, what is it about the demographics and connectivity of an underlying contact network that creates the most prominent features of the landscape in which the disease travels, and moreover how do *changes* to that landscape affect the dynamic behavior of the disease? An extensive body of work has explored this question, demonstrating, for example, that a global demographic such as population size can have profound effects on epidemic occurrence, leading to large, regular epidemics with few fadeouts (periods during which no one is infected) in populations above a certain threshold size, while populations whose size is below this threshold experience only small, chaotic epidemics [1]. The large amount of work on childhood diseases such as measles has shown that dynamic trends in the underlying social structure – for example, the change in aggregation among schoolchildren with the onset and end of the school term – create dynamic trends in epidemic profiles (epidemics tend to occur when children aggregate [1–3]). In addition, an increase in birthrate can cause a shift in epidemic periodicity from biennial to annual in populations above a certain threshold [1,2].

To date much of the mathematical modeling explaining epidemiological data has employed fully mixed compartmental models [1,2,4,5]. In the simplest of these models – the basic SIR model – the population is understood to consist of individuals who are either susceptible (S), infected/infectious (I), or recovered/immune (R). The population is assumed to mix fully, and interactions are governed by coupled differential equations such that the rate of change in the number of susceptible individuals is proportional to $-\beta SI$, where β is the contact rate, and individuals recover at a rate γ per unit time, such that the rate of change in the number of infected individuals is proportional to $\beta SI-\gamma I$ (see, for example, the supplement in

* Corresponding author at: Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, PA 16802, USA.
E-mail addresses: ralbert@phys.psu.edu, rza1@psu.edu (R. Albert).
[1] Current address: Department of Ecology and Evolutionary Biology and Woodrow Wilson School, Princeton University, Princeton, NJ 08544, USA.

0378-4371/$ – see front matter
doi:10.1016/j.physa.2010.02.034

Ref. [3]). This basic SIR model is quite successful in reproducing and explaining the real-world behaviors of large and even some intermediate-sized populations for which the assumption of full mixing is a good approximation. In fact, a handful of the model's more complex variants capture some of the nuances of more complicated populations for which full mixing is *not* a fair assumption [6,7]. However, as useful as fully mixed compartmental models are for large populations, their utility tends to be drastically reduced as population size decreases and as population structure (e.g. heterogeneity in connectivity) increases, since the assumption of full mixing fails to hold and the mathematics needed to describe the heterogeneities in mixing becomes intractable. Compartmental models also tend to do poorly at describing the initial or final stages of an outbreak, when few individuals are involved in transmission and stochastic person-to-person effects play an important role.

New avenues in epidemic modeling, involving individual-based *in-silico* simulation of the propagation of disease, address some of the shortcomings of compartmental modeling [8,9]. For example, stochastic, spatially-structured, individual-based simulations were used to model highly-contagious, aerosol-transmitted diseases, such as influenza [10,11]. Transmission in these models is based on co-location of individuals in schools, in workplaces, on public transportation etc. Generally these models define mixing groups such as households, schools, workplaces, neighborhoods and communities, and each mixing group is assumed to be (close to) well-mixed. Network-based models, on the other hand, simulate disease propagation in person-to-person contact networks, without adopting a well-mixed assumption, and therefore enabling an analysis of the relationships between the topology of contact networks and disease dynamics at multiple scales. The work of Newman on epidemic spreading in random networks [12], for example, revealed that the probability of a major epidemic depends on the average degree (connectivity) of the network. Pastor-Satorras and Vespignani showed that epidemics are always possible in populations whose interpersonal contacts are power law-distributed [13]. Realistic and highly-structured contact networks formed from real-world statistics for population composition were constructed to model SARS transmission in Vancouver, Canada [8] and to capture the movement of individuals between locations in a city [14].

Most individual-based and contact network models (including those mentioned previously) are demographically *static* representations; i.e. the implicit assumption is that demographic changes such as births, deaths and marriages will not affect the structure of the mixing groups or contact network, or will affect it at a rate much slower than the rate of disease spread in the network. The question therefore arises whether the dynamics of a disease on a demographically *dynamic* contact network will differ from the dynamics of the disease on such a network's static counterpart. We explore answers to this question by simulating realistic and detailed underlying topologies that are built from relevant statistical data and that are allowed *to evolve* according to rates collected from statistical (demographic) data for real societies. With this framework, we can then observe *directly*, for any population size, how the dynamics of the topology is influenced by demographic changes, and how, in turn, demographic dynamics affect the dynamics of disease.

Here, we simulate a contact/social network that is a conglomerate of family networks, work networks, school and preschool networks, and individuals, and that grows and changes according to salient real-world statistical rates. We track both the dynamics (demographic and topological) of the population and the dynamics of a disease propagating through this population. Simulations have been generated for measles, and we present a comparison of our findings to data found in Refs. [1–3]. We demonstrate agreement both in long-term epidemic profiles as well as in a multitude of epidemiological measures for a range of population sizes from 10,000 to ∼250,000 individuals. Furthermore, our results provide unique insight into how and why the social topology of the contact network influences the propagation of the disease through the population.

2. Simulating social networks and disease dynamics

In our simulations individuals are represented by vertices (nodes) and their most salient social interactions (familial, working, and (pre)school) are indicated by edges. Each node is characterized by the age and gender of the corresponding individual, and has edges to family members, classmates (if the individual is of school age) or work colleagues (if the individual is an adult). The social network grows and changes over time due to births, marriages, deaths, immigration, and to individuals joining and leaving schools and workplaces. The rates of these events are estimated from statistical data such as age distributions, birth rates, marriage rates, immigration and unemployment rates, etc. The social network algorithms are interlinked with the disease algorithm, and two dominant timescales are adopted in the simulations: a yearly timescale for "slow" social processes – i.e. marriages, formation of work groups and (pre)school groups – and a weekly timescale for "fast" or distributed social processes – i.e. births, deaths, immigration – and for (most) disease updates. While statistical data is abundant for *node-related* quantities in social networks, it is almost non-existent in regard to social *edges* in large social networks. It is therefore necessary to establish logical rules that are based on observation and "reverse-engineering" of the social underpinnings of social institutions (e.g. families, workplaces, (pre)schools) to account for how and why people are connected (in terms of having "social edges" between them) in a population. In the following subsections we briefly describe the major social network algorithms we employed in our population model as well as the disease algorithms (tailored for childhood diseases, such as measles) included in our simulations.

2.1. Dominant social processes

Basic demographics and social processes: births, deaths, age distribution and immigration: For a simulated population of size N, an initial age distribution is adapted from the vital statistics of London [15] and New York City [16], which have been