Physica A 389 (2010) 561–576

Contents lists available at ScienceDirect

Physica A

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

Stochastic epidemics and rumours on finite random networks

Valerie Isham [a], Simon Harden [a,*], Maziar Nekovee [b,c]

[a] Department of Statistical Science, University College London, London, UK

[b] Centre for Computational Science, University College London, Gordon Street, London WC1H 0AJ, UK

[c] Mobility Research Centre, BT, Polaris 134, Adastral Park, Martlesham, Suffolk IP5 3RE, UK

ARTICLE INFO

Article history:
Received 19 May 2009
Received in revised form 24 September 2009
Available online 7 October 2009

Keywords:
Epidemic models
Rumour models
Random networks
Stochastic models

ABSTRACT

In this paper, we investigate the stochastic spread of epidemics and rumours on networks. We focus on the general stochastic (SIR) epidemic model and a recently proposed rumour model on networks in Nekovee et al. (2007) [3], and on networks with different random structures, taking into account the structure of the underlying network at the level of the degree–degree correlation function. Using embedded Markov chain techniques and ignoring density correlations between neighbouring nodes, we derive a set of equations for the final size of the epidemic/rumour on a homogeneous network that can be solved numerically, and compare the resulting distribution with the solution of the corresponding mean-field deterministic model. The final size distribution is found to switch from unimodal to bimodal form (indicating the possibility of substantial spread of the epidemic/rumour) at a threshold value that is higher than that for the deterministic model. However, the difference between the two thresholds decreases with the network size, n, following a $n^{-1/3}$ behaviour. We then compare results (obtained by Monte Carlo simulation) for the full stochastic model on a homogeneous network, including density correlations at neighbouring nodes, with those for the approximating stochastic model and show that the latter reproduces the exact simulation results with great accuracy. Finally, further Monte Carlo simulations of the full stochastic model are used to explore the effects on the final size distribution of network size and structure (using homogeneous networks, simple random graphs and the Barabasi–Albert scale-free networks).

1. Introduction

In the 21st century, there are many infectious diseases that pose substantial threats to human and animal populations. There are concerns over the spread of HIV/AIDS, the emergence of new infections such as SARS, the threat of a new human influenza pandemic arising from avian influenza, possible outbreaks of smallpox resulting from terrorist action, and many others. Mathematical models have an important role to play in controlling the spread of such infections. In addition, the infectious agents and the infected hosts need not be biological systems. Individual personal computers and servers are regularly targeted by viruses spread across computer networks, while the effective transmission of information over the internet and the behaviour of social interaction networks are topical research interests. In many ways, the spread of information resembles that of infection, and the models that have been developed have many features in common.

In each case, individuals are classified as being of one of three types: susceptible, infected or removed for infections; ignorant, spreader or stifler for information or rumours. The simplest models assume homogeneous mixing of hosts, so that susceptibles become infected at a rate proportional to the current numbers of susceptibles and infectives, and similarly for

* Corresponding author. Tel.: +44 2076791872; fax: +44 2073834703.
E-mail address: simon@stats.ucl.ac.uk (S. Harden).

0378-4371/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2009.10.001

the rate of spread of information from spreaders to ignorants. For the infection process, the only other transition is from infective to removed (e.g. through recovery, quarantine or death) which is generally assumed to take place at a constant rate per infective. For the information process, similar transitions (from spreader to stifler) at a constant rate per spreader may also be allowed, but in addition such transitions may occur if a spreader contacts another spreader or stifler (who is already in possession of the information or rumour and so discourages the spreader). Thus, the 'stifling' rate is nonlinear and this leads to interesting differences between some properties of epidemics and rumours.

Mathematical models have many uses. They may be used for careful exposition of issues and gaining theoretical understanding, for example, of the role played in the spread of infection or information by heterogeneity of hosts with regard to their behaviour or susceptibility. They can be used to answer specific questions of interest, such as the effect of a particular control strategy, or to reach qualitative conclusions that distinguish those factors and sources of heterogeneity largely responsible for driving the dynamics from those that have little influence. In turn, these may help in formulating what data are needed before parameters can be determined and the models used for prediction. Simple models give robust approximations, while more complex models require more detailed assumptions (and tend to have more parameters). In real time, models may sometimes be used simply as a means of summarising large quantities of information about the dynamics of an outbreak.

Understanding the mechanisms by which an infection spreads brings possibilities for its control. Questions include how to detect a new outbreak of an existing infection or the emergence of a new one, what action is needed to prevent its spread, and what contingency plans are needed for its practical implementation? For a recurrent infection, what/when/how should a routine strategy (*e.g.* vaccination) be implemented? On the other hand, rather than trying to limit the spread of infection, the aim may be to find the most efficient way to spread information or a rumour quickly over a computer network. For instance, large organisations might wish to distribute a critical anti-virus software upgrade over corporate networks or a viral marketing campaign may want to spread a video quickly over the Internet. Rumour-like mechanisms form the basis of an important class of data dissemination algorithms in distributed computer systems. These algorithms are generally known as *gossip* protocols [1]. The principle underlying these algorithms mimics the spread of rumour among humans or epidemics in populations. A process that wishes to disseminate a new piece of information to the system does not send it to a server, or a cluster of servers, in charge of forwarding, but rather to a (randomly chosen) set of other peer processes that it *knows*. In turn each of these processes does the same, forwarding the information to some of its peers.

The most fundamental question for the spread of infection is, if a single infectious individual is introduced into a closed and susceptible population, under what conditions (summarised as a *threshold theorem*) will the infection spread to infect a large proportion of the population, as opposed to dying out without having done so? If it does spread, then what is the distribution of the *final size* of the outbreak, that is, the total number infected? Similar questions are equally fundamental to the spread of information or a rumour. Analytic results are usually asymptotic, applying in the limit as the population size becomes arbitrarily large, and so the question arises as to how these thresholds behave for finite populations. In addition, threshold theorems are most easily obtained for a homogeneously mixing population in which contacts between every pair of individuals are equally likely, and there is considerable interest in extending these results to structured populations, either to metapopulations consisting of subgroups where there is homogeneous mixing within and between subgroups, or to populations with a network structure. Related issues arise in other fields and there is a long history of research on such questions in the statistical physics literature; see Ref. [2] for an extensive review.

With the growing interest in complex networks, the question of the behaviour of epidemic and rumour models on such structures has become the subject of much recent research. In particular, a number of recent studies have shown that introducing the complex topology of the social networks along which the spreading takes place can greatly impact on both the threshold behaviour and the dynamics of these models [3–6]. The majority of previous studies have their focus on the *infinite population* limit for which statistical fluctuations were expected to become negligible, hence justifying a deterministic treatment which focuses on obtaining the mean values of the quantities of interest.

However, even with large populations fluctuations do not always average out to result in a small overall effect. Furthermore, even if deterministic models give the mean behaviour of the corresponding stochastic system, in application to finite systems it is important to take the variability of individual realisations into account and, for example, to have information on the probability distribution of the number of individuals within a society who hear a certain rumour after its spreading, rather than only the mean number. Finally, the threshold concept that has been used so far in both epidemic [4,5,7] and rumour dynamics [3] is based on bifurcations caused by nonlinearities in the deterministic mean-field models. A suitable extension of the threshold concept to stochastic epidemics is currently missing.

The focus of the current paper is, therefore, to investigate in detail such stochastic aspects of epidemic and rumour models on networks, paying particular attention to the relationship between the threshold concept and the final size of the epidemic/rumour in the stochastic versus deterministic models.

In Section 2 we first briefly review the necessary background on stochastic epidemic and rumour models for homogeneously mixing populations, and discuss some simple models for random networks. Then, in Section 3, we consider a stochastic model for a rumour spreading on the nodes of a random network; the corresponding epidemic special case is obtained by setting one parameter to zero. An approximation is proposed in which the correlation structure of the network is retained but dependence between the infection status of neighbouring nodes is omitted. The resulting model is simple to analyse mathematically and its properties, including thresholds and final size distributions, can be explored numerically and related to those of a deterministic approximation discussed in an earlier paper, Ref. [3]. In particular, we introduce the