



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Physica A 325 (2003) 33–39

PHYSICA A

[www.elsevier.com/locate/physa](http://www.elsevier.com/locate/physa)

# Modeling of contact tracing in social networks

Lev S. Tsimring<sup>a,\*</sup>, Ramón Huerta<sup>b,1</sup>

<sup>a</sup>*Institute for Nonlinear Science, University of California, San Diego, 9500 Gilman Dr.,  
La Jolla, CA 92093-0402, USA*

<sup>b</sup>*GNB, E.T.S. de Ingeniería Informática, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

Received 31 October 2002

---

## Abstract

Spreading of certain infections in complex networks is effectively suppressed by using intelligent strategies for epidemic control. One such standard epidemiological strategy consists in tracing contacts of infected individuals. In this paper, we use a recently introduced generalization of the standard susceptible-infectious-removed stochastic model for epidemics in sparse random networks which incorporates an additional (traced) state. We describe a deterministic mean-field description which yields quantitative agreement with stochastic simulations on random graphs. We also discuss the role of contact tracing in epidemics control in small-world and scale-free networks. Effectiveness of contact tracing grows as the rewiring probability is reduced.

© 2003 Elsevier Science B.V. All rights reserved.

PACS: 87.23.Ge; 87.10.+e

Keywords: Epidemics; SIR model; Networks

---

## 1. Introduction

Last several years have seen a tremendous growth of interest in properties and dynamics of complex networks [1]. It has been recognized that in many cases, the structure of the network is intimately related to the mechanisms of network growth, and can differ significantly from random graphs or regular lattices. The most interesting examples of complex networks are so-called small-world (SW) networks and scale-free (SF) networks. It is believed that many social networks has features reminiscent of SW

---

\* Corresponding author. Fax: +1-858-534-7664.

E-mail addresses: [ltsimring@ucsd.edu](mailto:ltsimring@ucsd.edu) (L.S. Tsimring), [rhuerta@ucsd.edu](mailto:rhuerta@ucsd.edu) (R. Huerta).

<sup>1</sup> Present address: Institute for Nonlinear Science, University of California, San Diego, La Jolla, CA 92093-0402, USA.

and SF networks. The structure of contacts in populations has important implications for propagation of infection (or information). Until recently, theoretical epidemiologists usually ignored this issue and studied so-called mean-field models of epidemics in which it is assumed that at any time the probability to get infected is the same for all individuals (see, for example Ref. [2]). In other works, spreading of an infection on relatively simple lattices of individuals has been studied within so-called “forest-fire” models [3]. In the last few years, the studies which elucidate the role of underlying network structure in the infection spreading began to appear in the literature [4–7].

In this paper, we analyze the influence of the network structure on the elimination of epidemics through tracing of contacts of infected individuals. This strategy is widely used for treatment of slowly transmitting and not self-curing diseases such as tuberculosis, AIDS and other sexually transmitted diseases. Similar strategies are employed by law-enforcement organizations trying to uncover criminal or terrorists networks, as well as for suppressing computer virus propagation, etc. Evidently, contact tracing only makes sense when there is a well-defined network structure underlying population, and so the structure of this network is crucial for understanding the dynamics of epidemics.

In the recent paper [8], a model of contact tracing has been proposed which was based on the assumption that infection is a slow branching process, while the contact tracing occurs at a much shorter time scale. This leads to a standard susceptible-infectious-removed (SIR) model with rescaled parameters and similar dynamics. In our recent paper [9], we proposed another model in which infection and contact tracing possess comparable time scales, and so their interplay determines the dynamics of the system. In this paper, we continue the analysis of the SITR model in social networks.

## 2. Stochastic model

Let us consider a population of  $N$  individuals whose connections to each other form a fixed graph. The degree  $k$  of a node  $n$  is a number of links between  $n$  and other individuals. In regular lattices, the degree of each node is the same, and in random graphs the degree distribution is Poissonian with a certain mean degree  $K = \langle k(n) \rangle$ .

We divide all individuals in four categories: susceptible (S), infected (I), traced (T), and recovered (R).

*Infection  $S \rightarrow I$ :* Before epidemics, the whole population is assumed to be susceptible to infection. An outbreak is initiated by infecting one or a few hosts. The probability of subsequent host infection depends on the state of its nearest neighbors. We assume that the infection proceeds as a simple contact process: if a susceptible node  $n$  has  $k_i(n)$  infectious neighbors, during a small  $\Delta t$  time interval it can become infected with the probability  $\alpha k_i(n) \Delta t$ .

*Tracing  $I \rightarrow T$ :* For simplicity we shall assume that there is no spontaneous recovery, an infectious individual can only be disinfected externally through screening. Immediately upon disinfecting, the individual becomes *traced* (T) for a certain period of time during which its neighbors are checked for possible infection. After that time, the individual spontaneously becomes removed, and its neighbors are no longer traced.

Accordingly, individuals are checked with a probability  $\beta$  which depends on the state of its neighbors. We assume that if an infectious host is checked, it is immediately cured, eliminated or at least isolated so it cannot infect other hosts. We introduce two non-exclusive strategies of checking for infectious hosts: random checking and contact tracing. Random checking means choosing an arbitrary host with probability  $\beta_r \Delta t$ , while contact tracing of host  $n$  is done with probability  $\beta_i k_i(n) \Delta t$  where  $k_i(n)$  is the number of neighbors of  $n$  which are in the *traced* state  $T$ . The random checking process is equivalent to the removal process of standard epidemics [2].

*Removal*  $T \rightarrow R$ : With certain probability  $\gamma \Delta t$ , traced hosts are transformed into the *removed* state, in which they also cannot be infected, but they are no longer under observation, so they do not initiate contact tracing.

### 3. Mean-field theory

It is easy to see that a mean-field approach cannot be applied to the contact tracing, since it does not take into account the non-uniform distribution of infection in the population. However, in Ref. [9] we introduced a more sophisticated mean-field model which operates not only with the mean densities of states, but also with the densities of links connecting nodes with different states. This model is based on the multi-site (or cluster) mean-field theory devised in non-equilibrium kinetics (see e.g. Ref. [10]) and used in mathematical epidemiology by Rand [6] (see also Ref. [7]). While in general this approach yields an infinite hierarchy of equations for multi-site correlation functions, in case of sparsely connected populations, this hierarchy can be closed by expressing the density of triplets vial local densities and pair densities. Without going into details of derivation (see Ref. [9]), let us present the resulting mean-field equations

$$\dot{s} = -\alpha s \hat{i}, \quad (1)$$

$$\dot{i} = \alpha s \hat{i} - \beta_r i - \beta_i i \hat{\tau}, \quad (2)$$

$$\dot{\hat{i}} = (\alpha K s - \alpha - \beta_r) \hat{i} - \beta_i \hat{i} \hat{\tau}, \quad (3)$$

$$\dot{\hat{\tau}} = \alpha \frac{s}{i} \hat{i} \hat{w} + \left( \beta_i \hat{\eta} - \beta_i - \gamma - \alpha \frac{s}{i} \hat{i} \right) \hat{\tau} + \beta_r \hat{\eta}, \quad (4)$$

$$\dot{\hat{w}} = \beta_r \hat{i} + \beta_i \hat{i} \hat{\tau} - \gamma \hat{w}, \quad (5)$$

$$\dot{\hat{\eta}} = \alpha \frac{s}{i} (2\hat{i} + 2 - \hat{\eta}) \hat{i} - (\beta_r + \beta_i \hat{\tau}) \hat{\eta}. \quad (6)$$

Here  $s, i$  are the densities of susceptible and infected nodes,  $\hat{i} = [is]/s$  is the mean number of infectious neighbors per susceptible node,  $\hat{\tau} = [i\tau]/i$  is the mean number of traced neighbors per infectious node,  $\hat{w} = [st]/s$  is the mean number of traced neighbors per susceptible node, and  $\hat{\eta} = [ii]/i$  is the mean number of infectious neighbors of an infectious node.

In the standard SIR model ( $\beta_i = 0$ ), the outcome of the epidemics is determined by the basic reproduction number  $R = K\alpha(\alpha + \beta_r)^{-1}$ : if  $R > 1$ , the epidemics engulfs a

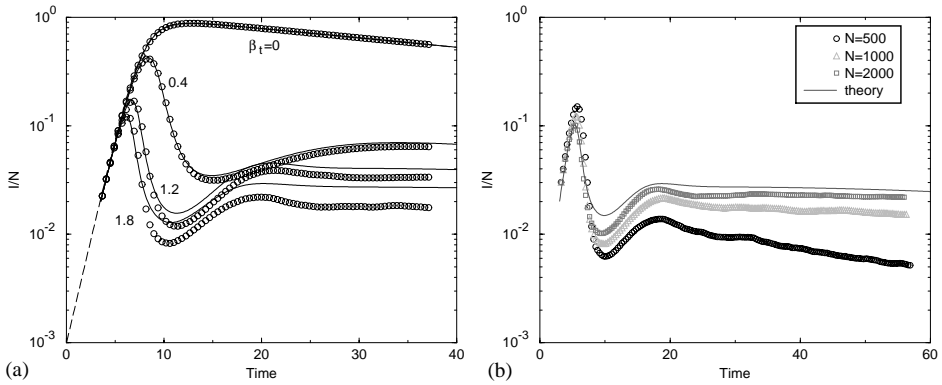


Fig. 1. (a) Evolution of the infection prevalence for different  $\beta_t$ , mean-field model (lines) and stochastic simulations (symbols),  $\alpha = 0.1, K = 10, \beta_r = 0.02, \gamma = 0.5$ . (b) Same for a fixed  $\beta_t = \dots$  at different system sizes  $N = 500, 1000, 2000$  compared with the mean-field prediction (line).

finite portion of the total population and if  $R < 1$ , the population of infected individuals exponentially shrinks. With contact tracing, this is no longer the case. In fact, at the initial phase of infection the percentage of traced nodes is vanishingly small, and the number of infected nodes grows exponentially if  $R > 1$ , however contact tracing can arrest this growth before  $i$  becomes large. Indeed, it can be shown that the fraction of traced nodes per infected node rapidly reaches a finite value  $\hat{\tau}_0$ , and that value re-normalizes the basic reproduction number  $R_\tau = K\alpha[\alpha + \beta_r(1 + \hat{\tau}_0)]^{-1}$ . At  $\beta_t$  greater than

$$\beta_{cr} = \frac{(\alpha(K - 1) + \gamma)(\alpha(K - 1) - \beta_r)}{\beta_r}, \tag{7}$$

$R_\tau < 1$ , and the initial exponential growth of infection is suppressed before the epidemics becomes finite, so the total number of infected nodes is independent of the population size, and depends only on the initial size of infection  $i(0)$ .

For  $R_\tau > 1$ , a finite epidemic outbreak occurs. More interestingly, the prevalence of infected nodes exhibit a smaller secondary peak. The nature of this secondary epidemics can be attributed to the presence of a slow time scale  $\gamma^{-1}$  of removal of traced nodes. Once the fraction of traced nodes falls below a certain threshold value at which  $R_\tau = 1$ , a second epidemic outbreak occurs. This prediction is confirmed by direct Monte-Carlo simulations of the original stochastic model (for details of simulation procedure see Ref. [9]). Fig. 1a shows the fraction of infected nodes  $i$  calculated from Eqs. (1)–(6) for different values of  $\beta_t$ . As seen from the figure, the maximum number of infectious nodes is drastically reduced with increase of  $\beta_t$ . In the same figure, we show the results of direct stochastic simulations for several  $\beta_t$ . Both numerical simulations and the model exhibit the emergence of the secondary epidemic outbreak after the first one is nearly suppressed.

While the agreement between the theory and simulations is very good at small  $\beta_t$ , the theoretical prediction of the secondary epidemics underestimates the results of

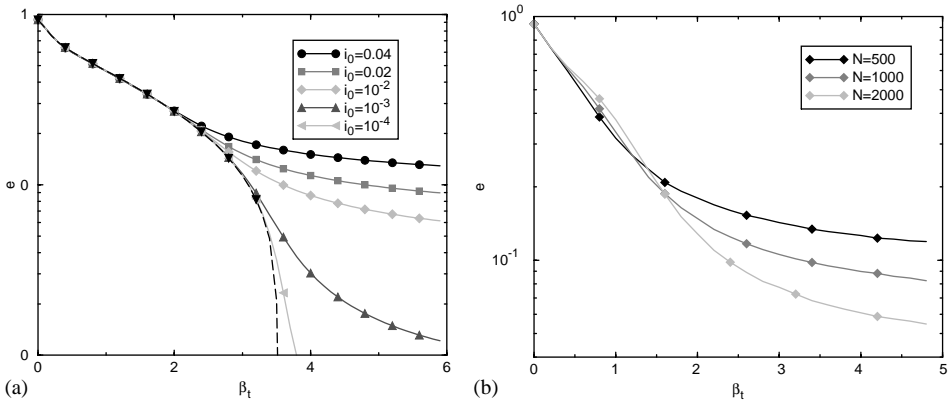


Fig. 2. Epidemic size  $e$  as a function of  $\beta_t$  for  $\alpha = 0.1, K = 10, \beta_r = 0.25, \gamma = 0.5$ : (a) theoretical curves; (b) numerical simulations for three different system sizes  $N = 500, 1000, 2000$ .

numerical simulations for large  $\beta_t$ . This systematic deviation is caused by the finite-size effects: for large  $\beta_t$  the number of infected nodes after the primary epidemics becomes very small, and fluctuations become essential. An implicit confirmation of this can be seen in Fig. 1b, in which the prevalence of infected nodes is shown as a function of time for three different system sizes. As seen from the figure, the agreement between theory and simulations steadily improves as the system size is increased.

The suppression of the major epidemics by contact tracing can also be seen in the dependence of the epidemic size  $e$  (total number of infected nodes normalized by the total number of network nodes  $N$ ) on  $\beta_t$  for different relative sizes of the initial infection  $i(0)$ . Theoretical prediction is shown in Fig. 2a. For chosen parameter values  $\beta_{cr} = 3.58$ . For  $\beta_t < \beta_{cr}$ , the epidemic size should become independent of  $i(0)$  at small  $i(0)$ . For  $\beta_t < \beta_{cr}$ ,  $e$  should approach zero as  $i(0) \rightarrow 0$ . Numerical results shown in Fig. 2b were obtained in calculations with three different network sizes  $N = 500, 1000, 2000$  and the same number of initially infected nodes (twenty) which correspond to  $i(0) = 0.04, 0.02, 0.01$ . At small  $\beta_t$ ,  $e$  is independent on the system size in agreement with the theory. At larger  $\beta_t$ , smaller networks demonstrate smaller  $e$  because of the finite size effects (small epidemics tend to die out early). However, at even larger  $\beta_t > \beta_{cr}$ , smaller networks demonstrate larger  $e$ , because below the epidemic threshold, the epidemic size is determined by the ratio of the initial infection to the total population size.

#### 4. Contact tracing in SW and SF networks

The theory presented in the previous sections was developed for random graphs which are characterized by small clustering and narrow degree distribution. However, for many social networks these conditions are violated, and they have wide degree distribution (such as SF networks [11]) or large clustering (SW networks [12]), or

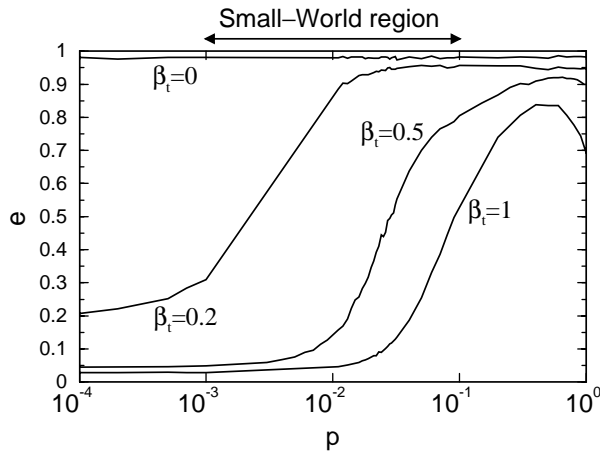


Fig. 3. The epidemic size as a function of the rewiring probability  $p$  for Sitr model in an SW network. Parameters of the Sitr model are the same as in Fig. 1.

both. We performed direct numerical simulations of the contact tracing in both SW and SF networks. The epidemic dynamics in SW networks remains qualitatively similar to random graphs, since they possess a well-defined epidemic threshold in both  $\beta_r$  and  $\beta_t$ .

Fig. 3 shows the dependence of the epidemic size  $e$  on the rewiring probability  $p$  which characterizes the degree of randomness in an SW network. As we can see,  $e$  changes mostly within the SW range ( $0.001 < p < 0.1$ ) where the clustering coefficient and the average path undergo large variations.

The situation in SF networks is more subtle. The presence of highly connected nodes leads to the absence of the epidemic threshold in simple SIR and SIS models in SF networks, so for an arbitrary small infection rate the epidemics affect a finite portion of the population [5,13]. Thus, it is of interest to see whether contact tracing can lead to an effective limitation of the epidemic size and in particular to the re-emergence of the epidemic threshold in SF networks, because highly connected nodes can be quickly traced and removed. Our preliminary simulations suggest that contact tracing do not lead to the return of the epidemic threshold; however, the size of the epidemics can be greatly reduced by the contact tracing.

## 5. Conclusions

In this paper, we studied the role of contact tracing as a part of the epidemic control strategy. By nature of this strategy, it can only be effective in structured networks when every node is connected to a certain group of other nodes. We demonstrated that by applying this strategy a major outbreak can be significantly reduced or even eliminated at a small additional cost. Using multi-site mean-field theory, we derived a model of

contact tracing for random Bernoulli graphs which are characterized by small clustering and narrow degree distribution. For other types of networks which either have wide degree distribution or large clustering, the theory has to be modified. The work in this direction is now in progress.

## Acknowledgements

This work was supported by the US DOE grant DE-FG03-96ER14592, ARO MURI-ARO grant DAAG55-98-0269, and NATO Collaborative Linkage Grant PST.CLG. 978512.

## References

- [1] M.E.J. Newman, *J. Stat. Phys.* 101 (2000) 819;  
S.H. Strogatz, *Nature* 410 (2001) 268.
- [2] N.T.J. Bailey, *The Mathematical Theory of Infectious Diseases*, 2nd Edition, Griffin, London, 1975.
- [3] A. Johansen, *Physica D* 78 (1994) 186.
- [4] C. Moore, M.E.J. Newman, *Phys. Rev. E* 61 (2000) 5678.
- [5] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* 86 (2001) 3200.
- [6] D. Rand, in: J. McGlade (Ed.), *Advanced Ecological Theory. Principles and Applications*, Blackwell Science, Oxford, 1999, pp. 100–142.
- [7] H. Andersson, T. Britton, *Epidemics: Stochastic models and their statistical analysis*, in: *Springer Lecture Notes in Statistics*, Vol. 151, Springer, New York, 2000.
- [8] J. Müller, M. Kretzschmar, K. Dietz, *Math. Biosci.* 164 (2000) 39.
- [9] R. Huerta, L. Tsimring, *Phys. Rev. E* 66 (2002) 056115.
- [10] D. ben-Avraham, J. Köhler, *Phys. Rev. A* 45 (1992) 8358.
- [11] R. Albert, A. Barabasi, *Rev. Mod. Phys.* 74 (2002) 47.
- [12] D.J. Watts, S.H. Strogatz, *Nature (London)* 393 (1998) 130.
- [13] R.M. May, A.L. Lloyd, *Phys. Rev. E* 64 (2001) 066112.