Control of epidemics on complex networks: Effectiveness of delayed isolation

Tiago Pereira[1,2,*] and Lai-Sang Young[3,†]

[1]*Department of Mathematics, Imperial College London, London SW7 2AZ, United Kingdom*

[2]*Institute of Mathematical and Computer Sciences, Universidade de São Paulo, São Carlos 13566-590, São Paulo, Brazil*

[3]*Courant Institute of Mathematical Sciences, New York University, New York 10003, New York, USA*

(Received 28 May 2015; published 31 August 2015)

We study isolation as a means to control epidemic outbreaks in complex networks, focusing on the consequences of delays in isolating infected nodes. Our analysis uncovers a tipping point: if infected nodes are isolated before a critical day d_c, the disease is effectively controlled, whereas for longer delays the number of infected nodes climbs steeply. We show that d_c can be estimated explicitly in terms of network properties and disease parameters, connecting lowered values of d_c explicitly to heterogeneity in degree distribution. Our results reveal also that initial delays in the implementation of isolation protocols can have catastrophic consequences in heterogeneous networks. As our study is carried out in a general framework, it has the potential to offer insight and suggest proactive strategies for containing outbreaks of a range of serious infectious diseases.

DOI: 10.1103/PhysRevE.92.022822 PACS number(s): 89.75.Hc, 64.60.Ht, 87.19.X−, 87.23.Ge

I. INTRODUCTION

Epidemics have a major impact on our society. Nearly 15 000 000 annual deaths worldwide can be attributed to infectious diseases [1], and more than 300 new diseases emerged in a span of 60 years [2]. Previous studies on epidemics have focused mainly on epidemic thresholds [3–5], total number of infected individuals during the outbreak [6,7], and rates of spreading [8]. Immunization as a means to prevent or lessen the severity of outbreaks has also been studied [9–12], and for diseases ranging from smallpox to influenza, immunization is effective. But when an outbreak occurs for which the population is not immunized the task of containing it can be daunting in today's highly interconnected world [13].

Drastic measures must be taken to contain outbreaks of serious diseases. A well documented strategy, that was widely used during the recent outbreaks of both SARS and Ebola, is the *isolation* of infected individuals [14–19]. While immediate isolation of all those potentially infected can extinguish an outbreak, in practice there are a number of factors which induce delays in the system; first, it takes time for public health officials to recognize the urgency of the situation. Second, local medical facilities may be underequipped. Finally, individuals often fail to recognize that they are infectious, or choose not to seek medical attention for as long as possible, infecting people with whom they come into contact in the meantime. These phenomena were, unfortunately, all too evident in the recent Ebola outbreak [16,17].

While some delay in the removal of infected individuals is unavoidable, after too long of a delay no isolation policy can be effective [15]. Education and preparedness can shorten these delays, but the latter can also incur major economic costs, and aggressive tactics to enforce isolation can be controversial. As governments weigh the costs and benefits of preparedness and emergency responses, an important question is: *How much delay can be tolerated without substantial impact on the effectiveness of isolation?*

In this paper, we address this question in the framework of complex networks. We consider two sources of delays: d_0 represents the initial delay, i.e., the number of days between the infection of "patient zero" and the start of an isolation policy, while d represents the average number of days between a node's infection and its removal, simulating the inadvertent delays after an isolation policy has been declared.

Our main message is that *there is a tipping point that can be estimated in terms of network and disease parameters.* That is, there is a critical day d_c such that for $d < d_c$, the total number of infected nodes is small and does not scale with system size, while it climbs steeply for $d > d_c$, and a large fraction of the population is infected. Moreover, our analysis reveals that *the more heterogeneous the network structure, the more detrimental time delays can be, and the differences are especially stark for* d_0. Thus for scale-free networks, smaller values of both d_0 and d are required to achieve the same desired effects.

This works focusses on time sensitivity: it reveals that the effectiveness of isolation must be evaluated alongside realistic time delays in its implementation. Our findings are based on the use of branching processes to model disease propagation in early stages of an outbreak. Using these probabilistic tools, we can analyze the effects of isolation delay in a general framework. So our results are not tied to a specific disease and offer insight into how to maximize the delays while keeping the number of infected nodes to a minimum.

II. MODEL DESCRIPTION

We consider a variant of the SIS model [20]. Each node of the network represents an individual and each link is a connection establishing the node's neighbors. Individuals exist in one of two discrete states, *healthy*, or *infected*. At each time step, which we think of as a day, a healthy, hence susceptible, node j is infected with probability $p_j = 1 - e^{-\beta n_j}$, where β is the infection rate and n_j is its number of infected neighbors. At the same time, each infected node recovers with rate γ. We assume $\gamma < \beta \ll 1$ are chosen above the epidemic threshold, i.e., the disease invades the entire population. This

*tiago.pereira@imperial.ac.uk

†lsy@cims.nyu.edu

model is extensively studied in the literature of complex networks [7,9,11].

We consider exclusively epidemics with point sources, that is to say, we initialize the model by setting exactly one node as infected; this is our "patient zero." In the absence of any isolation protocols, the system evolves according to the rules above.

Next we define what it means to impose an *isolation protocol with delay* $d = 1,2,\ldots$ (assuming for the moment $d_0 = 0$). Suppose the ith node is infected on day t. If it recovers on or before day $t + d$, then no action is required. If at the end of day $t + d$ it is still not recovered, then we put it in isolation, i.e., we remove it from the network, in the sense that from step $t + d + 1$ on, node i cannot infect any other node: it does not count toward n_j for any node j even though it is infected [21]. The period of isolation can be long but finite (after which the node is assumed to have recovered) or it can be indefinite; it does not matter much since we are primarily interested in short-time dynamics. For the same reason, our results are not appreciably affected whether our recovered nodes are susceptible, or forever immune to this disease, as in the SIR model. We also consider an *initial delay* of d_0 days before the protocol above is implemented.

Whether or not an isolation protocol is imposed, we keep track of the following statistics: Consider an infection starting from node i. We let $X_i(t)$ be the random variable giving the number of nodes that have been infected at or before time t (whether or not they have recovered), and let

$$\rho_i(t) = \frac{1}{N}\mathbb{E}[X_i(t)],$$

where N is the total number of nodes in the network. Finally, we average over all point sources, letting

$$\rho(t) = \frac{1}{N}\sum_i \rho_i(t) \quad \text{and} \quad \rho_* = \lim_{t\to\infty} \rho(t).$$

III. EFFECTS OF DELAYS WITH PARAMETER *d*

Our first goal is to understand the behavior of ρ_* in terms of network structure and the delay d. To this end, we consider two different kinds of networks: random Erdós-Renyi (ER) networks [22], and Barabási-Albert scale-free (SF) networks [23]. In our simulations, we use $N = 10^4$, $\beta = 0.04$, $\gamma = 0.01$. Our simulation results, which are shown in Fig. 1, suggest that there is a critical value d_c that depends on network structure: isolation protocols with $d < d_c$ are very effective, whereas ρ_* climbs steeply for protocols using $d > d_c$.

While a precise analysis of ρ_* is quite involved, we propose that early spread dynamics can be modeled effectively as *branching processes.*

Suppose the outbreak starts on day 0 at a node we will refer to as node 0. A simple computation shows that under the model rules above and an isolation protocol of $d \geqslant 1$, each neighbor of node 0 has probability $p \approx (1-\gamma)\beta d$ of getting infected: Neglecting first the recovery rate of node 0, a neighbor stays healthy for d days with probability $p \approx (1-\beta)^d$, so the probability of its getting infected is $p \approx 1-(1-\beta)^d \approx d\beta$. Including the recovery rate of node 0 gives an additional factor of $1-\gamma$. Thus the expected number of neighbors of node

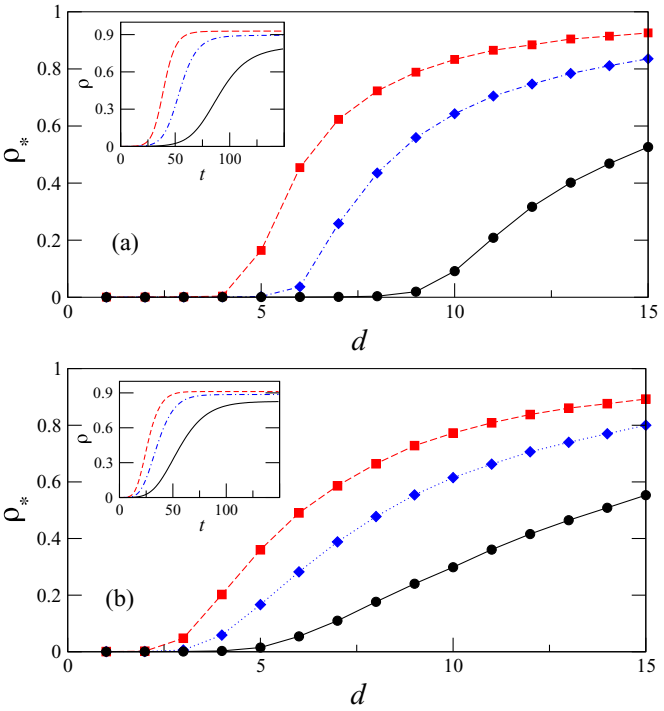

FIG. 1. (Color online) Effects of the isolation delay d on the maximum value ρ_*. Simulations reveal that there is a critical value d_c that depends on network structure: isolation protocols with $d < d_c$ are very effective, whereas ρ_* climbs steeply for protocols using $d > d_c$. Panel (a) shows plots for three ER networks with mean degrees $\langle k\rangle$ equal to 4 (•), 6 (◊), and 8 (□). Inset shows evolution of $\rho(t)$ as functions of t for mean degrees $\langle k\rangle$ equal to 4 (full line), 6 (dashed doted), and 8 (dashed). Panel (b) shows corresponding plots for SF networks with the same mean degrees. The SF network has degree distribution $p(k) \propto k^{-\alpha}$, with $\alpha = 3$.

0 who will be infected is $(1-\gamma)\beta d\langle k\rangle$ where $\langle k\rangle$ is the mean degree of the network. Each neighbor will infect, on average, another $(1-\gamma)\beta d\langle k-1\rangle$ of its neighbors, and the transmission is continued at this rate.

Here we have assumed that the relevant parts of the network have the structure of a tree, and the number of offspring is close to independent and identically distributed (i.i.d.) Under these assumptions, the dynamics of our model are well approximated by a branching process with the branching numbers above. Let $\mathbb{E}[Y_*]$ denote the expected total number of progeny for such a branching process. Then $\mathbb{E}[Y_*] = \infty$ if and only if $d \geqslant d_c$ where

$$d_c = \frac{1}{(1-\gamma)\beta\langle k-1\rangle}. \tag{1}$$

For $d < d_c$, $\mathbb{E}[Y_*] = 1 + \mathbb{E}[Y_1]\sum_{n=0}^{\infty}[(1-\gamma)\beta d\langle k-1\rangle]^n$ where Y_1 is the number of offspring of generation 1, giving

$$\mathbb{E}[Y_*] = 1 + \frac{(1-\gamma)\beta d\langle k\rangle}{1-(1-\gamma)\beta d\langle k-1\rangle}.$$

Plugging in $d = d_c - a$, $a > 0$, one obtains

$$\mathbb{E}[Y_*] \approx \frac{d_c}{a}.$$

Thus for $d \leqslant d_c - \frac{1}{2}$, for example, $\mathbb{E}[Y_*] = \mathcal{O}(d_c)$. Returning to our network, this means $\rho_* = \mathcal{O}(\frac{1}{N})$ for $d \leqslant d_c - \frac{1}{2}$, and $\rho_* = \mathcal{O}(1)$ for $d \geqslant d_c$. Notice also that d_c depends only on infection and recovery rates and mean degree, and is independent of system size.

For $\langle k \rangle = 4,6,8$, the computation above gives $d_c \approx 8.4$, 5, and 4 respectively. These values compare well with our simulation results for ρ_* for the ER networks in Fig. 1(a). While early dynamics are well described by a branching process, as ρ_* gets larger, it becomes increasingly likely that some of the neighbors of a newly infected node have already been infected, leading to a decline in branching number.

For the SF networks, on the other hand, the critical values for d_c revealed by simulations are significantly smaller than for ER networks; see Fig. 1(b). We attribute that to the presence of hubs, or highly connected nodes, in the network, and propose below a second order correction to the estimation of d_c for networks with heavy-tailed degree distributions.

The idea is to consider not only first neighbors but also *second neighbors*, i.e., nodes that are linked to a given node in exactly two steps. Continuing to study the tree along which the infection spreads, we let i be one of the infected nodes. For a second neighbor, the probability of getting infected by node i is $p \approx [(1-\gamma)\beta d]^2$. Thus the number of offspring (every two generations) is approximately $[(1-\gamma)\beta d]^2 N_2(i)$ where $N_2(i)$ is the mean number of second neighbors of node i. Taking the average over all nodes, the mean number of offsprings is $[(1-\gamma)\beta d]^2 \langle N_2 \rangle$, and $\langle N_2 \rangle \approx \langle k^2 \rangle$ [24]. Assuming that the number of second neighbors infected is significantly larger than the number of first neighbors infected, the argument above gives a revised value of

$$d_c^{\mathrm{cor}} \approx \frac{1}{(1-\gamma)\beta(\mu^2+\sigma^2)^{1/2}}, \tag{2}$$

where μ and σ^2 are the mean and variance of the network's degree distribution.

For the SF networks simulated with $\langle k \rangle = 4,6,8$, we have computed that $\sigma^2 \approx 36$, 72, and 123 respectively. Plugging into the formula above, we obtain $d_c^{\mathrm{cor}} \approx 3.5, 2.4, 1.8$, in excellent agreement with the simulation results in Fig. 1(b). Comparing (2) with (1), we see that the larger σ^2 is (relative to μ^2), the larger is the discrepancy between d_c^{cor} and d_c. For our ER networks the correction is about 5%, which is immaterial.

IV. EFFECTS OF INITIAL DELAYS WITH PARAMETER d_0

As explained earlier, this d_0-day delay is intended to capture the time it takes for public health authorities to recognize the need for imposing isolation protocols. We remark that our discussion of d_0 here is simplified by the absence of an incubation period, the length of which must be factored in when computing tolerable initial delays in real-life situations; our d_0 corresponds to longer (and more realistic) initial delays when incubation period is taken into account. Figure 2 shows, for an ER and an SF network, the effects of initial delays on ρ_*.

For purposes of investigating ρ_*, systems with $d_0 > 0$ can be modeled as branching processes with $\rho(d_0)N$ nodes in the zeroth generation. The critical delay d_c is unchanged, and for any d, if we let $\mathbb{E}_{d_0}[X_*]$ denote the total number of infected nodes with initial delay d_0, then

$$\mathbb{E}_{d_0}[X_*] = \rho(d_0) N \mathbb{E}_0[X_*].$$

When $\mathbb{E}_0[X_*] = \mathcal{O}(1)$, ρ_* remains small for d_0 not so large, and for fixed $d < d_c$, ρ_* is proportional to $\rho(d_0)$. This is seen in the picture for the ER network in Fig. 2.

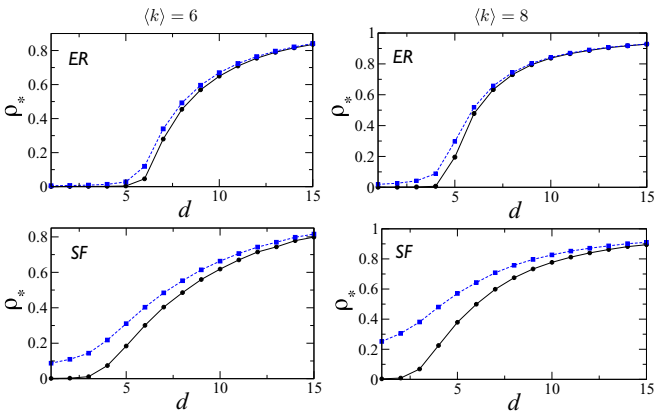

FIG. 2. (Color online) Effects of initial delay d_0 before the imposition of isolation protocols. Our simulations show effects are much more severe for SF than for ER networks. In the left column we show plots of ρ_* as functions of d for an ER and SF network with mean degree $\langle k \rangle = 6$, for $d_0 = 7$ (○) and 21 (□). In the right column we show the corresponding plots for ER and SF networks with $\langle k \rangle = 8$.

For SF networks, ρ_* can be very large for even moderate size d_0, as can be seen in Fig. 2. Once a hub is reached during the initial time period when the epidemic spreads unchecked, $\rho(d_0)$ can be so large that the situation is out of control. In the SF network used in Fig. 2, where $N = 10^4$ and $\langle k \rangle = 6$, the diameter of the network is 5, and 1% of the nodes have degrees larger than 40. Designating this top 1% of nodes as *hubs*, we find that nearly 50% of nodes are immediate neighbors of hubs, and nearly 97% lie within distance 2 of a hub. Thus it is quite easy to reach a hub during this initial period of unchecked growth. This is what makes initial delays so much more dangerous for SF networks than for ER ones.

V. DISTRIBUTIONS OF NUMBERS OF INFECTED NODES

Up until now we have focused on ρ_*, which involves averaging over all point sources i of outbreaks and all sample paths of the spread dynamics. To investigate the role played by hubs and to identify situations that are potentially explosive, it is instructive to view ϱ_t as a random variable in i (the first infected node) and the sample path (noise realization), so that $\mathbb{E}(\varrho_t) = \rho(t)$. We find that while in ER networks the probability density function (PDF) ν of ϱ_t has a profile consistent with exponential growth, the situation is quite different for SF networks, where some infection sources and sample paths lead quickly to containment and others lead to full-blown epidemics; see Fig. 3.

For the SF network at $t = 90$, there is the following clear dichotomy: ϱ_t falls in $A = [0,0.01]$ with probability $\sim 1/3$ and in $B = [0.2,0.28]$ with probability $\sim 2/3$. Moreover, there is a clear dependence on the structural role of the node

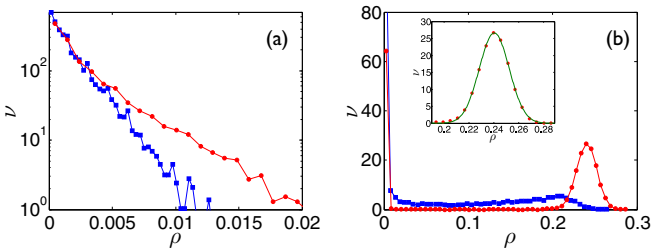

FIG. 3. (Color online) PDF ν of ϱ_t for $t = 30$ (□) and $t = 90$ (○). All networks have $\langle k\rangle = 6$, $d_0 = 0$, and $d = 5$. In (a) the distribution for ER network exhibit an exponential behavior. In (b) the distribution for SF networks show a bimodal behavior. In the inset of (b), we show a Gaussian fit of high end of the distribution ϱ_t for $t = 90$.

where the infection starts. For outbreaks that originate from hubs, $\varrho_t \in B$ with probability 1, while for first neighbors of hubs, the corresponding probability is roughly 0.8. In fact, an overwhelming proportion of the nodes for which $\varrho_t \in A$ are low degree nodes, their mean degree being about 4. These statistics suggest that imposing more stringent protocols on hubs and neighbors of hubs can be an effective strategy.

The bimodal distributions of ϱ_t for SF networks in Fig. 3 can be explained as follows: Nodes with small degrees are natural barriers for disease transmission. For example, in an SF network with $\langle k\rangle = 4$, many nodes have degree $\leqslant 2$. For d_0,d not so large, such nodes offer a good chance to stop the propagation. On the other hand, once a hub is reached, the spread will likely continue for some time even with quite small d, leading to the Gaussian-like distribution at the high end. Finally, the accessibility of hubs in SF networks discussed earlier explains why with high probability one of these two scenarios will prevail.

VI. CONCLUSIONS

We have shown that isolation can be an extremely effective measure if taken quickly, namely, before a critical day d_c, and have uncovered the value of d_c in terms of the network structure and disease parameters. We have also demonstrated the perils of initial delays in SF networks. Our analysis extends naturally to networks with communities: as a random walker is initially trapped within its community [25], we can extend our analysis by taking into account local community structures. We expect that in scale-free networks with assortative mixing, d_c is further reduced as a consequence of correlation hubs and low degree nodes. Furthermore, we can weaken our assumption on local tree structures by introducing effective numbers of free neighbors to tackle small-world networks. Our results are general and provide a theoretical basis for assessing the impact of delays on the effectiveness of isolation protocols. They serve as a starting point for case studies once specific details of the disease are incorporated.

T.P. was partially supported by the European Union's Seventh Framework Programme, ERC AdG Grant No. 339523 RGDD. L.S.Y. was supported in part by NSF Grant No. DMS-1363161.

[1] D. M. Morens, G. K. Folkers, and A. S. Fauci, Nature (London) **430**, 242 (2004).
[2] K. E. Jones, N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman, and Peter Daszak, Nature (London) **451**, 990 (2008).
[3] C. Castellano and R. Pastor-Satorras, Phys. Rev. Lett. **105**, 218701 (2010).
[4] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[5] S. Bowong, J. J. Tewa, and J. Kurths, Int. J. Bifurcation Chaos **23**, 1350128 (2013).
[6] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).
[7] M. J. Keeling and K. T. D. Eames, J. R. Soc., Interface **2**, 295 (2005).
[8] M. Barthelemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).
[9] X. Fu, M. A. Small, and G. Chen, *Propagation Dynamics on Complex Networks Models, Methods and Stability Analysis* (John Wiley & Sons, United Kingdom, 2014).
[10] R. Cohen, S. Havlin, and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003).
[11] J. C. Miller and E. M. Volz, PLoS ONE **8**, e69162 (2013).
[12] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
[13] M. A. Small, D. M. Walker, and C. K. Tse, Phys. Rev. Lett. **99**, 188702 (2007).
[14] T. Day, A. Park, N. Madras, A. Gumel, and J. Wu, Am. J. Epidemiol. **163**, 479 (2006).
[15] G. Chowell and H. Nishiura, BMC Medicine **12**, 196 (2014).
[16] S. I. Okware, F. G. Omaswa, S. Zaramba, A. Opio, J. J. Lutwama, J. Kamugisha, E. B. Rwaguma, P. Kagwa, and M. Lamunu, Trop. Med. Int. Health **7**, 1068 (2002).
[17] A. Pandey, K. E. Atkins, J. Medlock, N. Wenzel, J. P. Townsend, J. E. Childs, T. G. Nyenswah, M. L. Ndeffo-Mbah1, and A. P. Galvani, Science **346**, 991 (2014).
[18] S. Riley, C. Fraser, C. A. Donelly *et al.*, Science **300**, 1961 (2003).
[19] N. M. Ferguson, D. A. T. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke, Nature (London) **442**, 448 (2006).
[20] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, Phys. Rev. E **86**, 041125 (2012).
[21] For mathematical simplicity we have set incubation period to 0; otherwise d would be the average number days since the end of incubation.
[22] V. Batagelj and U. Brandes, Phys. Rev. E **71**, 036113 (2005).
[23] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[24] Notice that $N_2(i) = \sum_\ell \sum_{j\neq i} A_{i\ell}A_{\ell j}$, and the average $\langle N_2\rangle = [\sum_i N2(i)]/N$. Performing the sum first in the index j, then in the index i we obtain $\langle N_2\rangle \approx [\sum_\ell k_\ell^2]/N$.
[25] J.-C. Delvenne, S. N. Yaliraki, and M. Barahona, Proc. Natl. Acad. Sci. USA **107**, 12755 (2010).