

PAPER • OPEN ACCESS

Contact tracing in configuration models

To cite this article: Ivan Kryven and Clara Stegehuis 2021 *J. Phys. Complex.* **2** 025004

View the [article online](#) for updates and enhancements.

You may also like

- [Beyond COVID-19: network science and sustainable exit strategies](#)
J Bell, G Bianconi, D Butler et al.
- [Averaging dynamics, mortal random walkers and information aggregation on graphs](#)
Orowa Sikder
- [Based on mathematical epidemiology and evolutionary game theory, which is more effective: quarantine or isolation policy?](#)
Muntasir Alam, K M Ariful Kabir and Jun Tanimoto

OPEN ACCESS

PAPER



Contact tracing in configuration models

RECEIVED

12 October 2020

REVISED

16 November 2020

ACCEPTED FOR PUBLICATION

15 December 2020

PUBLISHED

2 February 2021

Ivan Kryven¹ and Clara Stegehuis^{2,*} ¹ Mathematical Institute and the Centre for Complex Systems Studies, Utrecht University, The Netherlands² Department of Electrical Engineering, Mathematics and Computer Science, Twente University, The Netherlands

* Author to whom any correspondence should be addressed.

E-mail: c.steghuis@utwente.nl**Keywords:** random graphs, configuration model, contact tracing, epidemic spreading

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

**Abstract**

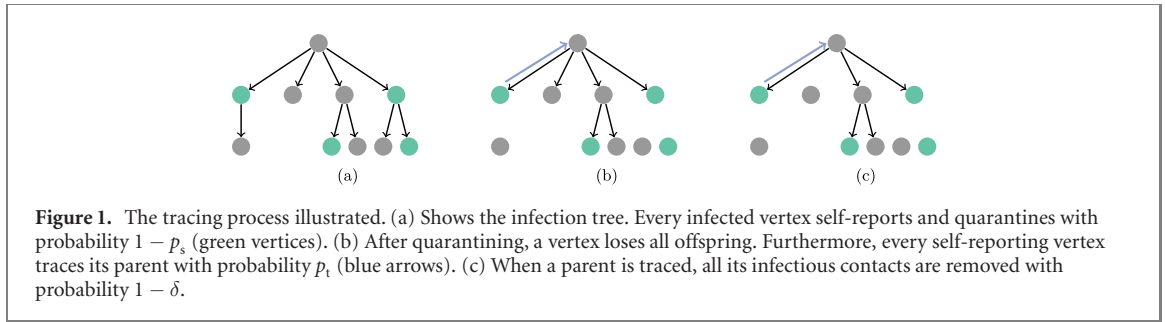
Quarantining and contact tracing are popular ad hoc practices for mitigating epidemic outbreaks. However, few mathematical theories are currently available to assess the role of a network in the effectiveness of these practices. In this paper, we study how the final size of an epidemic is influenced by the procedure that combines contact tracing and quarantining on a network null model: the configuration model. Namely, we suppose that infected vertices may self-quarantine and trace their infector with a given success probability. A traced infector is, in turn, less likely to infect others. We show that the effectiveness of such tracing process strongly depends on the network structure. In contrast to previous findings, the tracing procedure is not necessarily more effective on networks with heterogeneous degrees. We also show that network clustering influences the effectiveness of the tracing process in a non-trivial way: depending on the infectiousness parameter, contact tracing on clustered networks may either be more, or less efficient than on networks without clustering.

1. Introduction

Contact tracing is a frequently used method to control epidemic outbreaks. In this method, individuals who show symptoms of a disease, report themselves and identify their recent contacts which are then tested for the disease. If a contact tests positive, it is being isolated to prevent further spreading of the disease. In this way, an epidemic may be contained in its early stages.

The effect of contact tracing has mathematically been investigated by extending compartmental models, such as the SIR model, with an additional rule that infected individuals may be detected and removed with some rate that represents a tracing process [1, 3, 8, 15], or by other differential equation approaches [11, 12]. However, such compartmental models simplify the structure of contact networks by representing it with one numerical parameter. Complex networks on the other hand may have nontrivial structure, featuring heavy tailed degree distributions, clustering, and other phenomena. For example, the contact network of the HIV/AIDS epidemic in Cuba was found to be well-approximated by a power-law degree distribution [7], so that the proportion of vertices with k neighbors scale as $k^{-\tau}$. Such degree distributions feature a large variability of node degrees, with vertices of large degrees (also called hubs) being present along with large number of small degree nodes. We will refer to this phenomenon as *degree-heterogeneity*. Furthermore, power-law degree distributions were shown to cause important epidemiologic properties, such as vanishing epidemic thresholds [4, 18], strong finite-size effects [19], and novel universality classes for critical exponents [9].

A recent simulation study suggested that contact tracing is more effective on networks with high degree-heterogeneity [14]. Intuitively, high-degree vertices infect more others than low-degree vertices, so that they are also more likely to be traced. Furthermore, quarantining high-degree vertices has a larger effect on the spreading of epidemics than quarantining low-degree vertices. Thus, on these types of networks, contact tracing is expected to be more effective than is predicted in the standard SIR-models due to degree-heterogeneity. In [2], this expectation was made more formal by showing that the tracing process becomes more effective when high-degree vertices are likely to install contact tracing apps.



While approaches in [2, 14] rely on networks being locally tree-like, many real-world networks violate this condition and feature clustering: they contain a high density of triangles. Simulations suggest that network clustering has a strong positive impact on the effectiveness of the contact tracing process in homogeneous networks [13]. In general epidemics, clustering can either speed up, or slow down the spread of an epidemic process [20, 21].

In this paper, we quantify the network effect on the effectiveness of contact tracing, by mapping it to a combination of bond- and site-percolation models. We show that the extent to which contact tracing reduces the number of infections highly depends on the exact choice of tracing model. We show that when the tracing process is not immediate, but takes a nonzero amount of time, this drastically affects the outbreak size. We then investigate the effect of degree-heterogeneity and clustering on the effect of contact tracing on the final outbreak size using percolation models and find that clustering can either increase or decrease the effectiveness of tracing processes, depending on the infectiousness of the epidemic. This shows that the interplay between the underlying network structure and the exact choice of tracing process is delicate, and important to take into account.

We first describe the network model and define the tracing process in section 2. Then we show the relation between the success probability of tracing and the characteristic time of the tracing process. Section 3 analyzes the final outbreak size of our epidemic model with a generating function approach. We then study the effect of inducing clustering in the network in section 4.

2. Network and tracing model

In this paper, we assume that the underlying network is given by the configuration model, a network model that can generate networks with any prescribed degree distribution $(q_k)_{k \geq 1}$ [5]. In the configuration model, every vertex of degree k is equipped with k half-edges, which are paired uniformly at random. We assume that the disease spreads on this network as a bond percolation process: it removes each edge independently with probability $1 - \pi$. While this is a very simple variant of an epidemic process, the final size of an SIR epidemic with constant recovery duration can be identified as the size of the largest connected component after bond percolation [10]. In this setting, the effective basic reproductive number R_0 , or the average number of vertices infected by one infected vertex, is given by $R_0 = \pi \mathbb{E}[D(D-1)] / \mathbb{E}[D]$, where D denotes the degree of a uniformly chosen vertex [16].

We investigate the effect of the tracing process illustrated in figure 1 on the final size of the epidemic. The tracing process works in three steps. In the first step, every infected vertex ‘reports’ its infection independently with probability $1 - p_s$. If a vertex reports itself, it quarantines. Thus, a reported vertex is unable to infect other vertices, as shown in figure 1(b). The second step of the tracing process takes the contact tracing into account. In this second step, all vertices that have ‘reported’ themselves, list their recent contacts, who are being traced. For every reported vertex, their infector is identified independently with success probability p_t . In this case, we say that the infector vertex was ‘traced’.

The third step is the quarantining of all ‘traced’ vertices. Specifically, after a vertex is ‘traced’, it quarantines, so that it is unable to infect other vertices. However, the traced vertex may already have infected other vertices before it was traced. We therefore model such secondary quarantining of the traced vertices by removing each edge incident to a traced vertex with probability $1 - \delta$. These removed edges are the infections that have been removed due to the traced vertex being quarantined. Thus, the tracing process is modeled as an extra layer of bond percolation, see figure 1(c).

2.1. Immediate or delayed tracing: the impact of δ

The probability that the connection to a vertex is removed when its parent is traced, $1 - \delta$, depends on the parameters p_s and p_t . Here we show how δ relates these parameters under two assumptions on the tracing

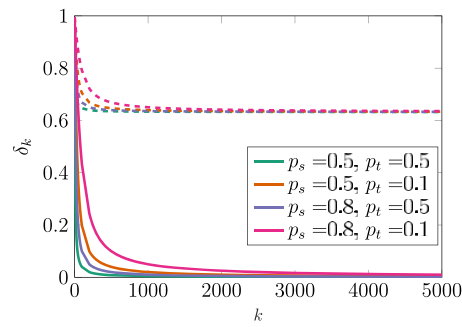


Figure 2. δ_k as a function of k for various values of p_s and p_t . The solid lines plot δ_k for immediate tracing (equation (1)), whereas the dotted lines plot δ_k for tracing with delay (equation (5)), using $\lambda = T = 1$.

process: immediate and delayed tracing, and discuss the impact of these assumptions on the effectiveness of the tracing process.

2.2. Immediate tracing

We first assume that the tracing process is immediate: once a vertex self-reports, it immediately traces its parent with probability p_t . If successful, the traced vertex immediately quarantines and cannot infect other vertices anymore. We now show that this assumption leads to a degree-dependent version of δ : δ_k .

Consider an outcome of the infection process as a tree composed of infected vertices. Tracing and self-reporting happens with the same probability, $(1 - p_s)p_t$, for all infected vertices. Therefore, for a given infected vertex in the tree that infects k neighbors of which d neighbors trace it, the first of these d ‘tracing’ contacts can be viewed as the first red ball drawn without replacement from an urn with d red balls and $k - d$ black balls. The number of black balls drawn before the first red ball is on average $(k - d)/(d + 1)$, which corresponds to the average number of infectious contacts of a vertex before it is first traced. Therefore, the average fraction of non-tracing contacts that occur before the vertex is traced equals $1/(d + 1)$.

The number of tracing vertices, d , is binomially distributed with parameters $(k, (1 - p_s)p_t)$, where k denotes the number of infectious contacts of the vertex. Using that $\mathbb{E}[(X + 1)^{-1}] = p^{-1}(1 + k)^{-1}(1 - (1 - p)^{k+1})$ when X is distributed as $\text{Bin}(k, p)$, we obtain that the average fraction of contacts that appear before the first tracing occurs, δ_k , equals

$$\delta_k = \frac{1 - (1 - (1 - p_s)p_t)^{k+1}}{(1 - p_s)p_t(1 + k)}, \tag{1}$$

so that δ_k is decreasing in k (see figure 2), and asymptotically, as k becomes large, we have:

$$\delta_k = \frac{1}{(1 - p_s)p_t k} (1 + o(1)). \tag{2}$$

Thus, we see that δ_k tends to zero when k becomes large, implying that for large values of k , only a vanishing fraction of contacts will not be traced.

Phase transition under immediate tracing. From (2) we obtain that the expected number of edges that remains for every vertex of degree k is asymptotically $\delta_k k \approx 1/((1 - p_s)p_t)$. As this quantity is independent of the vertex degree k , one might expect that the immediate tracing process removes the degree-heterogeneity. We will now show that the immediate tracing process is indeed very effective by calculating the critical value for the infectiousness parameter π , π_c after which the epidemic outbreak becomes extensive. That is, when $\pi < \pi_c$, the size of epidemic outbreaks are sub-linear in the total number of nodes, and when $\pi > \pi_c$, this size is linear. When the outbreak size scales linearly with the total number of vertices, we call such outbreak extensive or *giant*.

In appendix A, we show that there is a giant outbreak when

$$\frac{g'_D(1 - (1 - p_s)p_t\pi)}{\mathbb{E}[D]} < 1 - \frac{(1 - p_s)p_t}{p_s}, \tag{3}$$

where the random variable D denotes the degree of a randomly chosen vertex in the network, and $g_D(x)$ its probability generating function, $g_D(x) = \sum_k q_k x^k$. Thus, the critical value of the percolation parameter π_c at which a giant outbreak occurs is such that

$$\frac{g'_D(1 - (1 - p_s)p_t\pi_c)}{\mathbb{E}[D]} = 1 - \frac{(1 - p_s)p_t}{p_s}. \tag{4}$$

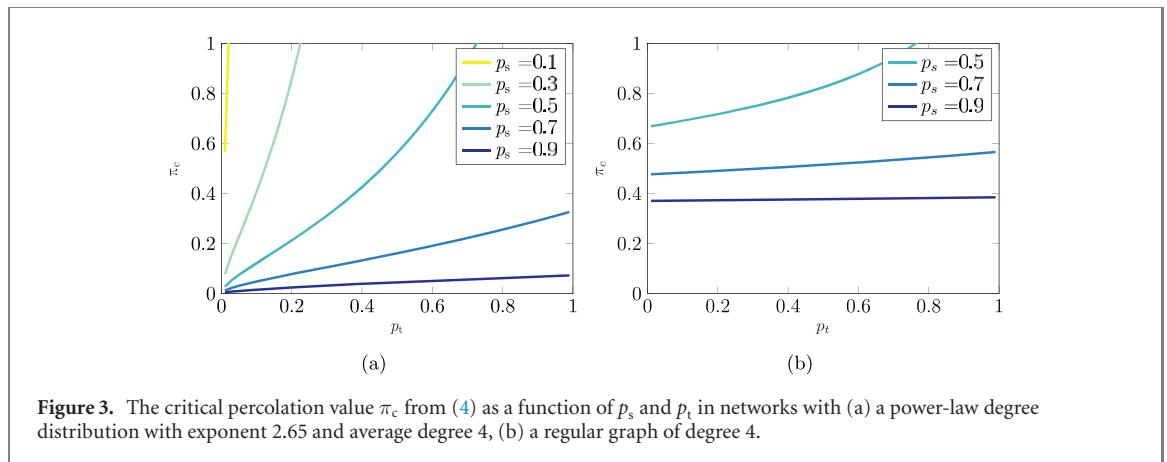


Figure 3 shows the value of π_c for two choices of the degree distribution: a regular graph where every vertex has degree 4 ($q_4 = 1$), and a power-law degree distribution with exponent 2.65 and average degree 4 ($q_k = Ck^{-2.65}$). Interestingly, we see a qualitative difference between the tracing and no-tracing scenarios. Figure 3(a) shows that $\pi_c > 0$ when $p_s, p_t > 0$ even for power-law distributions with degree exponent $\tau \in (2, 3)$. This means that under tracing, there is a regime for the infectiousness parameter π such that there are only small outbreaks. On the other hand, without tracing, $\pi_c = 0$ for power-law distributions with degree exponent $\tau \in (2, 3)$ [18], showing that a giant outbreak always occurs regardless of the value of infectiousness π . Thus, this tracing process is very effective: it can reduce an extensive outbreak to have a sub-extensive size.

In the standard SIR model, a comparable qualitative change in the size of the outbreak corresponds to a bifurcation taking place when the basic reproduction number $R_0 = 1$. In the regular graph, figure 3(b), decreasing p_s or increasing p_t increases the critical value π_c . Thus, when decreasing p_s or increasing p_t , there is a wider range of values of the infectiousness parameter π such that only small outbreaks occur, or alternatively, where the effective value of R_0 remains below one.

2.3. Tracing with delay

Even though the immediate tracing can result in a significant reduction of the giant outbreak, in practice, the tracing process may not be immediate. In what follows, we assume that there is a time-delay between the moment when a vertex self-reports and successfully traces its infector and the moment when the infector quarantines. We then again obtain an expression for the probability that the connection to a vertex is removed when its parent is traced, and obtain a degree-dependent version of the parameter δ : δ_k .

Suppose that it takes time T for a vertex to trace its infector (the step of figure 1(c)). Furthermore, assume that all infections from a degree- k vertex occur as independent exponential time clocks of rate λ . That is, an infected vertex still infects its neighbors independently with probability π . Furthermore, these infections of neighbors appear at independent, exponentially distributed time intervals with parameter λ . In the time-window of length T (the incubation period) in which an infector is not traced and therefore not quarantined yet, it can still infect others. Specifically, every remaining neighbor of the infector is infected independently in this time interval with probability $1 - e^{-\lambda T}$.

If we denote the number of neighbors of a degree- k vertex that are infected during the incubation period by N_q , and the number of vertices that were already infected before the incubation period started by N_b , then N_q is distributed as a Binomial($k - N_b, 1 - e^{-\lambda T}$) variable. Thus, we obtain

$$\mathbb{E}[N_q] = \mathbb{E}[\mathbb{E}[N_q|N_b]] = (1 - e^{-\lambda T})\mathbb{E}[k - N_b].$$

We then use that $\mathbb{E}[N_b] = k\delta_k$ with δ_k as in (1) to obtain

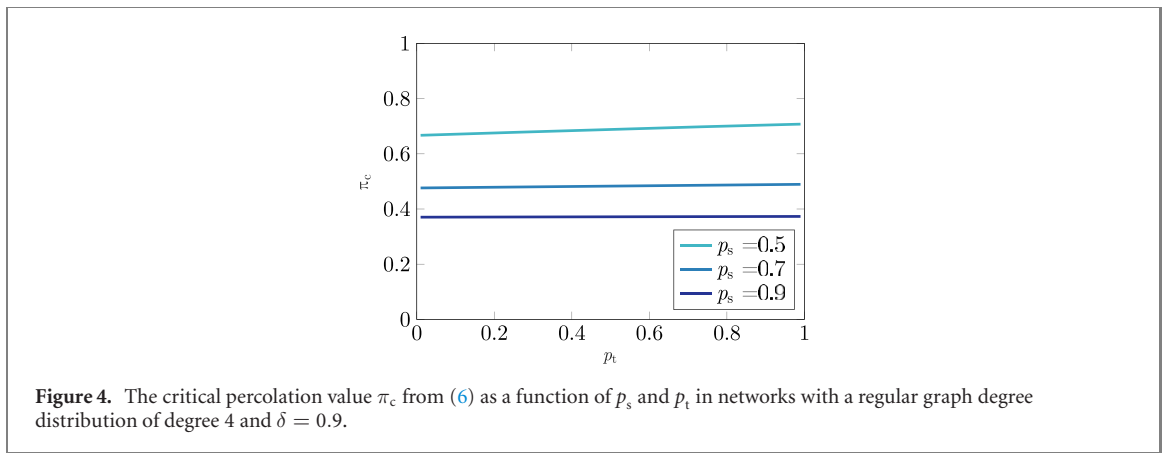
$$\mathbb{E}[N_q] = (1 - e^{-\lambda T})k \left(1 - \frac{1 - (1 - (1 - p_s)p_t)^{k+1}}{(1 - p_s)p_t(1 + k)} \right)$$

Then, the average number of vertices that are infected before tracing occurs is

$$\mathbb{E}[N_q] + \mathbb{E}[N_b] = (1 - e^{-\lambda T})k + k e^{-\lambda T} \frac{1 - (1 - (1 - p_s)p_t)^{k+1}}{(1 - p_s)p_t(1 + k)},$$

and the average fraction of neighbors that are infected before tracing occurs is

$$\delta_k(T) = 1 - e^{-\lambda T} \left(1 - \frac{1 - (1 - (1 - p_s)p_t)^{k+1}}{(1 - p_s)p_t(1 + k)} \right). \tag{5}$$



For large k ,

$$\delta_k(T) = 1 - e^{-\lambda T}(1 + o(1)),$$

which is independent of k . This implies that we can use $\delta = 1 - e^{-\lambda T}$ as a proxy, instead of having a k -dependent δ .

We therefore use a k -independent value of δ throughout the rest of the paper, which assumes a tracing process that is not immediate.

Phase transition under delayed tracing. In appendix B we show that the critical value of π beyond which a giant outbreak occurs, satisfies

$$(1 - \delta)g_D''(1 - \pi_c(1 - p_s)p_t) + \delta\mathbb{E}[D(D - 1)] = \frac{\mathbb{E}[D]}{\pi_c p_s}, \tag{6}$$

Equation (6) implies that $\pi_c = 0$ for power-law degree distributions with $\tau \in (2, 3)$, as then $\mathbb{E}[D^2]$, which appears on the left-hand side, diverges, so that the corresponding plot of figure 3(a) would be all lines at $\pi_c = 0$. We therefore only plot the critical percolation value π_c for regular graphs. This indicates that the epidemic threshold in power-law networks and regular graphs behave rather differently under delayed contact tracing. Figure 4 shows the value of π_c in regular graphs. We see that the value of π_c is more sensitive to p_s , the self-quarantining probability, than to p_t , the tracing probability. Thus, increasing the effectiveness of the tracing procedure barely influences the value of the epidemic threshold, though it may still influence the final size of the epidemic.

The influence of the tracing process on the critical value beyond which an epidemic becomes extensive is substantially more pronounced when the tracing process is immediate: under immediate tracing, even outbreaks in scale-free networks can become sublinear, while this is not possible under delayed tracing.

3. Final outbreak size under contact tracing

We now investigate the size of the remaining outbreak after tracing using a generating function approach under fixed δ , as described in section 2.3. In appendix C, we show that in the large-network limit, the fraction of vertices in the giant outbreak S is given by

$$S = p_s - p_s g_D(1 - \pi + \pi u), \tag{7}$$

where u is obtained by solving the implicit equation

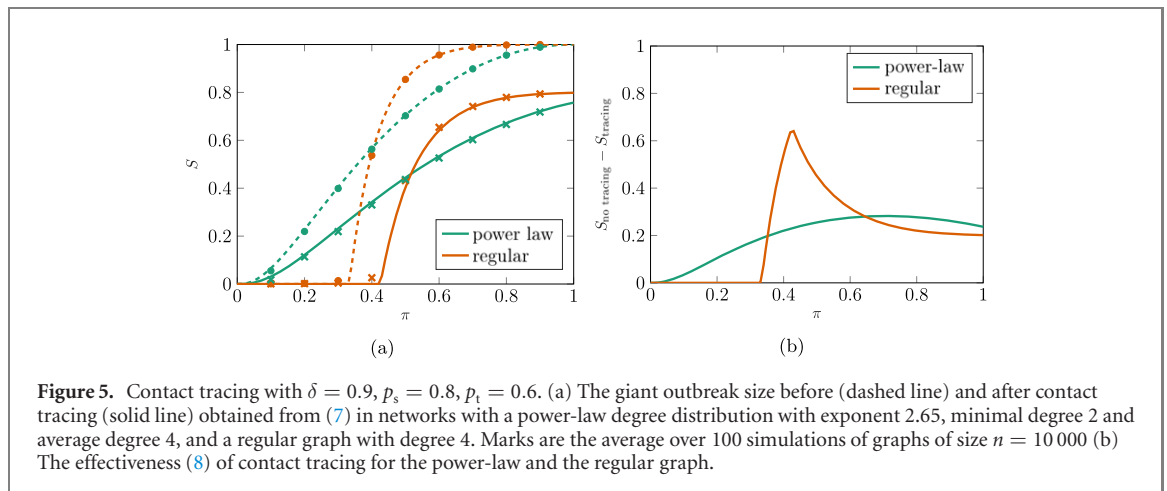
$$u = 1 - p_s + p_s [g_{D^*-1}(\delta\pi(u - 1) + 1) - g_{D^*-1}(((p_s - 1)p_t + 1)(\delta\pi(u - 1) + 1)) + g_{D^*-1}((\pi(u - 1) + 1)((p_s - 1)p_t + 1))],$$

where $g_{D^*-1}(x)$ is the generating function for the excess degree distribution: $g_{D^*-1}(x) = g_D'(x)/\mathbb{E}[D]$. Figure 5(a) plots the size of the giant outbreak for networks with two different degree distributions, and shows that the analytical results of (7) match well with numerical simulations.

By comparing the outbreak size with and without tracing, we can determine the effectiveness of contact tracing. That is,

$$\text{eff} = S_{\text{no tracing}} - S_{\text{tracing}}, \tag{8}$$

the outbreak size in an epidemic without contact tracing, minus the outbreak size in an epidemic with tracing. Here the outbreak size without tracing can be obtained by setting $p_s = 1$. Figure 5(b) plots the effectiveness



of contact tracing for two networks with the same average degree, but different degree distributions: a power-law degree distribution and a regular degree distribution. In both networks, the effectiveness of the tracing process depends on the infectiousness parameter π . In the regular network, the tracing process may shift the critical value of π_c where the giant outbreak occurs, so that tracing completely removes a giant outbreak. In that regime, tracing is very effective. When a giant outbreak occurs in both the epidemic with tracing and in the epidemic without tracing, the effectiveness of contact tracing decreases in π . That is, the more infectious the disease, the less effective the tracing procedure. In the power-law network, a giant outbreak is always present in both the traced and the non-traced version of the epidemic. In this situation, there seems to be an ‘optimal’ value of the infectiousness parameter π where the tracing process is most effective.

We see that tracing is not necessarily more effective in heterogeneous power-law networks compared to the homogeneous regular graph, in contrast with previous studies [2, 14]. For low and high values of π , tracing is more effective on power-law networks, while for intermediate values of π , tracing is more effective on the regular graph. This contrast with previous findings is caused by the immediate tracing assumption discussed in section 2.1. Under immediate tracing, most of the degree-heterogeneity is removed. Therefore, immediate tracing is extremely effective on heterogeneous networks, which were studied in [2, 14]. However, figure 5(b) shows the impact of contact tracing *with* delay. As (5) shows, tracing under delay is not as strongly influenced by the degree-heterogeneity. This makes the process of tracing with delay more intricate to understand, causing the non-trivial effects of figure 5(b).

The fact that for low values of π , tracing is more effective on power-law networks can be explained by the fact that for low values of π there is a giant outbreak on power-law networks, but not in regular graphs. Therefore, the tracing process is able to remove more infected vertices on the power-law network for low values of π , as more infected vertices are present. The sharp increase in the effectiveness on regular networks around $\pi \approx 0.35$ – 0.4 can be explained by the fact that this is the regime where without contact tracing, a giant outbreak would occur, but tracing reduces the giant outbreak to a small, sublinear one. Thus, in this regime, tracing is extremely effective. The peak of the curve in figure 5(b) for regular graphs is precisely the critical value where also under contact tracing a giant outbreak starts to occur of equation (6).

The fact that the tracing procedure on power-law networks is most effective for $\pi \approx 0.7$ is harder to explain, as the epidemic is extensive with and without contact tracing, but this value of the infectiousness π where the tracing process is most effective can be calculated from equation (7).

4. The effect of clustering on tracing

The configuration model is known to be locally tree-like: the fraction of triangles in the network vanishes asymptotically [6]. However, many real-world networks contain a non-trivial amount of triangles, which motivates studying the tracing process on a configuration model with enhanced clustering [17]. In this model, each vertex v has an edge-degree $d_v^{(1)}$ and a triangle degree $d_v^{(2)}$, denoting the number of triangles that the vertex is part of. Then a random graph is formed by pairing edges uniformly at random and pairing triangles uniformly at random.

Let the degree-triangle distribution be denoted by $q_{k,l}$, where k denotes the edge-degree, and l the triangle-degree. Let $g(x, y) = \sum_{k,l>0} q_{k,l} x^k y^l$ be the generating function of the edge and triangle degrees. Furthermore,

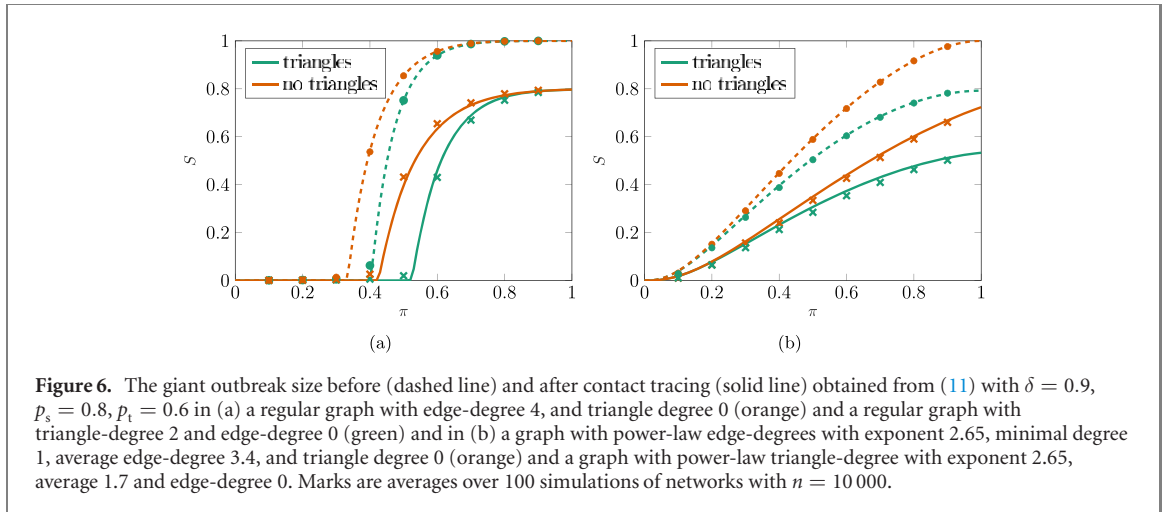


Figure 6. The giant outbreak size before (dashed line) and after contact tracing (solid line) obtained from (11) with $\delta = 0.9$, $p_s = 0.8$, $p_t = 0.6$ in (a) a regular graph with edge-degree 4, and triangle degree 0 (orange) and a regular graph with triangle-degree 2 and edge-degree 0 (green) and in (b) a graph with power-law edge-degrees with exponent 2.65, minimal degree 1, average edge-degree 3.4, and triangle degree 0 (orange) and a graph with power-law triangle-degree with exponent 2.65, average 1.7 and edge-degree 0. Marks are averages over 100 simulations of networks with $n = 10\,000$.

let

$$g_p(x, y) = \frac{1}{\langle k \rangle} \sum_{k,l>0} k q_{k,l} x^{k-1} y^l, \quad (9)$$

$$g_q(x, y) = \frac{1}{\langle l \rangle} \sum_{k,l>0} l q_{k,l} x^k y^{l-1}, \quad (10)$$

with

$$\langle s \rangle := \sum_{k,l>0} k q_{k,l}, \quad \langle l \rangle := \sum_{k,l>0} l q_{k,l}$$

be the generating functions of the number of edges and triangles that are reached by following a randomly chosen edge and a randomly chosen triangle respectively.

In appendix D, we show that the outbreak size after tracing equals

$$S = p_s - p_s g(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2 \pi v + \pi^2(3 - 2\pi)v^2), \quad (11)$$

where u and v are obtained by solving the system of implicit equations

$$\begin{aligned} u &= 1 - p_s + p_s (g_p(1 - \pi + \pi w u, (1 - \pi)^2 + 2(1 - \pi)^2 \pi w v + 2\pi^2(1 - \pi)v^2 w + \pi^2 w^2 v^2) \\ &\quad + g_p(1 - \pi \delta + \pi \delta u, (1 - \pi \delta)^2 + 2(1 - \pi \delta)^2 \pi \delta v + \pi^2 \delta^2(3 - 2\pi \delta)v^2) \\ &\quad - g_p(1 - \pi + \pi w(1 - \delta + \delta u), (\pi + \pi(\delta - 1)w - 1)^2 + 2\pi \delta w(\pi \delta - 1)(\pi + \pi(\delta - 1)w - 1)v \\ &\quad - w\pi^2 \delta^2(2(\pi - 1) + w(2\pi(\delta - 1) - 1))v^2)) \end{aligned}$$

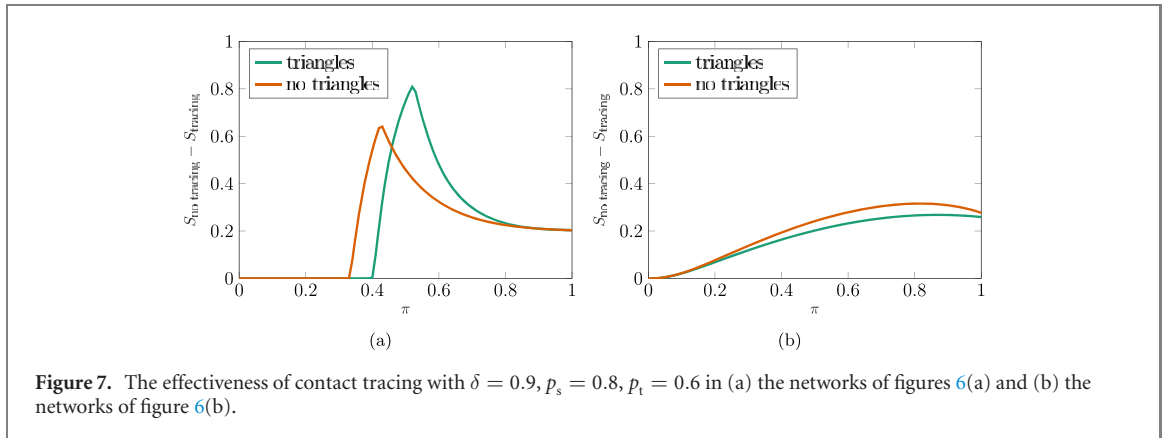
and

$$\begin{aligned} v &= 1 - p_s + p_s (g_q(1 - \pi + \pi w u, (1 - \pi)^2 + 2(1 - \pi)^2 \pi w v + \pi^2(1 - \pi)w^2 v^2) \\ &\quad + g_q(1 - \pi \delta + \pi \delta u, (1 - \pi \delta)^2 + 2(1 - \pi \delta)^2 \pi \delta v + \pi^2 \delta^2(3 - 2\pi \delta)v^2) \\ &\quad - g_q(1 - \pi + \pi w(1 - \delta + \delta u), (\pi + \pi(\delta - 1)w - 1)^2 + 2\pi \delta w(\pi \delta - 1)(\pi + \pi(\delta - 1)w - 1)v \\ &\quad - w\pi^2 \delta^2(2(\pi - 1) + w(2\pi(\delta - 1) - 1))v^2)), \end{aligned}$$

where $w = p_s + (1 - p_s)(1 - p_t)$.

Figures 6(a) and (b) show the epidemic size in networks with the same degree distribution but with a different amount of triangles. The analytic results for the final epidemic outbreak on networks with triangles obtained from (11) closely matches the results obtained by numerical simulations.

Furthermore, one may conclude from figures 7(a) and (b) that the effectiveness of the contact tracing non-trivially depends on the amount of clustering. In the regular graph, figure 7(a) shows that there is a range of the infectiousness parameter π where the tracing procedure is more effective on clustered networks than on tree-like networks, but there is also a range of parameters where the tracing procedure is more effective on the tree-like networks instead. On the heterogeneous power-law networks on the other hand, figure 7(b) shows that the effectiveness of tracing is always higher in the tree-like network than in the clustered networks. Furthermore, the difference between the clustered and non-clustered networks is less pronounced in the power-law network.



Intuitively, introducing triangles has two effects: on the one hand they make it easier for an epidemic to spread, as they induce multiple paths for a person i to infect another person j , but on the other hand, they reduce the number of vertices that the epidemic can reach from a given vertex in k steps compared to a tree. The latter effect makes it easier for the tracing process to stop the epidemic in the presence of triangles. For power-law vertices, this is less pronounced, as in the presence of high-degree vertices, it is likely that the vertex has already infected many other neighbors before being traced. This may intuitively explain the difference between introducing triangles in power-law networks compared to homogeneous networks.

In general, figure 7 shows that the effectiveness of contact tracing delicately depends on the interplay between the network degree distribution and its structure in terms of clustering.

5. Conclusion

In this paper, we have analytically studied a contact tracing process on networks with arbitrary degree distributions. In this process, infected vertices self-report and quarantine with some probability $1 - p_s$, and they trace their parent with probability p_t . Using generating functions, we derive analytical expressions of the giant outbreak size after the tracing process.

We investigated the effect of the network structure on the tracing process and found that degree heterogeneity may either enhance or diminish the effectiveness of tracing depending on the exact parameter values. In our tracing model, we assume that there is a time-delay between the time that a person is infected and the time that its infector is traced. This assumption makes the network heterogeneity non-trivially affect the tracing effectiveness.

Likewise, enhancing clustering in the network has a non-trivial effect on the effectiveness of contact tracing. Depending on the infectiousness of the epidemic, clustering may either increase or decrease the effectiveness of contact tracing, in contrast with conclusions from simulations on homogeneous networks [13]. This underlines the importance of taking the network structure into account when investigating such tracing processes.

In this paper, we investigated bond percolation, which can be mapped to the final size of an SIR epidemic with constant recovery duration. In further research, it would be interesting to investigate the entire time evolution of the number of infected vertices in the SIR process as well, and investigate the effect of the network structure on this time evolution. This would enable to answer the question whether the network structure affects the speed at which tracing processes slow down the spread of an epidemic.

Furthermore, our results on power-law networks suggest that there is an optimal value of the infectiousness parameter π such that tracing is the most effective. It would be interesting to investigate the relation between this optimal value of π and the parameters of the tracing process, to enable the design of optimally efficient tracing processes.

Data availability statement

No new data were created or analysed in this study.

Appendix A. Derivation of the critical value for immediate tracing

We model the infection tree as a branching process with a certain offspring distribution. A giant component emerges when the average offspring surpasses one. Thus, we calculate the average number of non-self-reporting offspring of a vertex N_f (as the self-reporting vertices will have zero offspring, and therefore do not contribute to creating a giant outbreak).

Let N_t the total number of offspring of a vertex in the second tier of the branching process in the original network. Then, N_t is distributed as $D^* - 1$, where D^* denotes the size-biased degree distribution: the distribution of the degree of a vertex that is found by following a uniformly chosen edge. Let $N_t^{(\pi)}$ be the number of neighbors after percolation with parameter π , so that $N_t^{(\pi)} = \text{Bin}(N_t, \pi)$. Let R denote the number of ‘reporting’ neighbors of the vertex, so that $R = \text{Bin}(N_t^{(\pi)}, 1 - p_s)$. Finally, let R_t the number of reporting vertices that also ‘trace’ its parent vertex, so that $R_t = \text{Bin}(R, p_t)$. Given $R_t, R, N_t^{(\pi)}$, the vertex infects $N_t^{(\pi)} - R$ non-reporting vertices before taking tracing into account. On average, $(N_t^{(\pi)} - R)/(R_t + 1)$ of these will remain after tracing (so when permuting randomly, and removing every neighbor after the first ‘tracing’ sibling). By the laws of conditional expectation

$$\begin{aligned} \mathbb{E}[N_f] &= \mathbb{E}\left[\mathbb{E}\left[N_f|N_t^{(\pi)}, R\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{E}\left[N_f|R_t\right]|N_t^{(\pi)}\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{N_t^{(\pi)} - R}{R_t + 1}|N_t^{(\pi)}, R\right]\right] \end{aligned}$$

Using that $\mathbb{E}\left[(X + 1)^{-1}\right] = p^{-1}(1 + k)^{-1}(1 - (1 - p)^{k+1})$ when X is distributed as $\text{Bin}(k, p)$ and that R is distributed as $\text{Bin}(N_t^{(\pi)}, 1 - p_s)$, we obtain

$$\begin{aligned} \mathbb{E}[N_f] &= \mathbb{E}\left[\mathbb{E}\left[\frac{(N_t^{(\pi)} - R)(1 - (1 - p_t)^{R+1})}{(R + 1)p_t}|N_t^{(\pi)}\right]\right] \\ &= \mathbb{E}\left[\frac{p_s(1 - (1 - (1 - p_s)p_t)^{N_t^{(\pi)}})}{(1 - p_s)p_t}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{p_s(1 - (1 - (1 - p_s)p_t)^{N_t^{(\pi)}})}{(1 - p_s)p_t}|N_t\right]\right], \end{aligned}$$

where the last equality follows again from the laws of conditional expectation. Further, as the probability generating function of a $\text{Bin}(k, p)$ random variable Y is $\mathbb{E}[x^Y] = (1 - p + px)^k$, and $N_t^{(\pi)}$ is distributed as $\text{Bin}(N_t, \pi)$, $\mathbb{E}[(1 - (1 - p_s)p_t)^{N_t^{(\pi)}}|N_t] = (1 - \pi + \pi(1 - (1 - p_s)p_t))^{N_t}$. Thus, we obtain

$$\begin{aligned} \mathbb{E}[N_f] &= \mathbb{E}\left[\frac{p_s(1 - (1 - (1 - p_s)p_t\pi)^{N_t})}{(1 - p_s)p_t}\right] \\ &= \frac{p_s}{(1 - p_s)p_t} - \frac{p_s}{(1 - p_s)p_t}\mathbb{E}[(1 - (1 - p_s)p_t\pi)^{N_t}] \\ &= \frac{p_s(1 - g_{D^*-1}(1 - (1 - p_s)p_t\pi))}{(1 - p_s)p_t}, \end{aligned}$$

where $g_{D^*-1}(x)$ is the probability generating function of the size-biased degree distribution minus 1 (as N_t is distributed as $D^* - 1$), so that $g_{D^*-1}(x) = g'(x)/\mathbb{E}[D]$.

A giant outbreak occurs when the expected number of offspring surpasses one, so when

$$\frac{g'_D(1 - (1 - p_s)p_t\pi)}{\mathbb{E}[D]} < 1 - \frac{(1 - p_s)p_t}{p_s}. \tag{12}$$

Thus, the critical value of the percolation parameter π_c is such that

$$\frac{g'_D(1 - (1 - p_s)p_t\pi_c)}{\mathbb{E}[D]} = 1 - \frac{(1 - p_s)p_t}{p_s}. \tag{13}$$

Appendix B. Critical value under delayed tracing

We now derive the critical percolation value under delayed tracing with fixed δ . We use the same notation as in appendix A. We use the same notation for the number of different types of offspring, see table 1. When R

Table 1. Definition of the variables used in appendices A and B

Variable	Definition
N_f	Number of non-reporting offspring of an infected vertex after tracing
N_t	Total number of offspring of an infected vertex before tracing
$N_t^{(\pi)}$	Number of infected offspring of an infected vertex before tracing
R	Number of ‘reporting’ offspring of an infected vertex
R_t	Number of offspring of an infected vertex that ‘trace’ their infector

vertices report themselves, the probability that their infector is not traced is $(1 - p_t)^R$. There are $N_t^{(\pi)} - R$ non-reporting vertices. When their infector is traced, on average a fraction of δ of them remain infected. Therefore,

$$\begin{aligned} \mathbb{E}[N_f] &= \mathbb{E}\left[\mathbb{E}\left[N_f | N_t^{(\pi)}, R\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[(1 - p_t)^R (N_t^{(\pi)} - R) + (1 - (1 - p_t)^R)(N_t^{(\pi)} - R)\delta | N_t^{(\pi)}\right]\right] \\ &= \mathbb{E}\left[N_t^{(\pi)} p_s (1 - (1 - p_s)p_t)^{N_t^{(\pi)} - 1}\right] + \delta \pi p_s \mathbb{E}[D^* - 1] \\ &\quad - \delta \mathbb{E}\left[N_t^{(\pi)} p_s (1 - (1 - p_s)p_t)^{N_t^{(\pi)} - 1}\right], \end{aligned}$$

where the last step used that given $N_t^{(\pi)}$, $N_t^{(\pi)} - R$ is binomially distributed with parameters $N_t^{(\pi)}$ and $(1 - p_s)$, and the probability generating function of a $\text{Bin}(k, p)$ random variable is given by $(1 - p + px)^k$. Also, given N_t , $N_t^{(\pi)}$ is binomially distributed with parameters N_t and π . Thus,

$$\begin{aligned} \mathbb{E}[N_f] &= p_s (1 - \delta) \mathbb{E}\left[\pi N_t (1 - \pi(1 - p_s)p_t)^{N_t - 1}\right] + \delta \pi p_s \mathbb{E}[D^* - 1] \\ &= p_s \pi (1 - \delta) g'_{D^* - 1}(1 - \pi(1 - p_s)p_t) + \delta \pi p_s \mathbb{E}[D^* - 1]. \end{aligned}$$

Because $\mathbb{E}[Yx^{Y-1}] = g'_Y(x)$ for any random variable Y , and N_t is distributed as $D^* - 1$, where D^* is the size-biased degree-distribution, Finally, using that $g_{D^* - 1}(x) = g'_D(x)/\mathbb{E}[D]$ and that $\mathbb{E}[D^* - 1] = \mathbb{E}[D(D - 1)]/\mathbb{E}[D]$, we obtain

$$\mathbb{E}[N_f] = \frac{p_s \pi (1 - \delta) g''_D(1 - \pi(1 - p_s)p_t)}{\mathbb{E}[D]} + \delta \pi p_s \mathbb{E}[D(D - 1)].$$

The critical value of π where a giant outbreak occurs, is when $\mathbb{E}[N_f] = 1$, which yields equation (6).

Appendix C. The giant outbreak size

In this section, we compute the giant outbreak size after tracing. A vertex does not trace its infector if it does not self-report, which happens with probability p_s , or if it does self-report, but does not successfully trace its infector, which happens with probability $(1 - p_s)(1 - p_t)$. Thus, the probability that a vertex of degree k is traced by none of its offspring equals

$$\mathbb{P}(\text{degree } k \text{ vertex not traced}) = (p_s + (1 - p_s)(1 - p_t))^k. \tag{14}$$

Let

$$p_k = \frac{(k + 1)q_{k+1}}{\sum_{k \geq 1} kq_k}$$

be the excess degree distribution, and p_k^* be the excess degree distribution after tracing. As a vertex loses all its offspring after self-reporting which happens with probability $1 - p_s$, $p(0)$ is given by

$$p(0) = 1 - p_s.$$

When a vertex is not traced, its degree remains the same. When a vertex is traced, an extra layer of percolation occurs with parameter δ . Thus,

$$p_k^* = \underbrace{p_s}_{\text{Not quarantined}} \left(\sum_{j=k}^{\infty} \underbrace{p_{k,j}^{\delta} (1 - (p_s + (1 - p_s)(1 - p_t))^j)}_{\text{At least one offspring traces this node}} + \underbrace{(p_s + (1 - p_s)(1 - p_t))^k}_{\text{None of the offsprings trace this node}} p_k \right), \quad k > 0$$

where $p_{k,j}^\delta$ is the probability that a vertex of degree j has remaining degree k after percolation with bond occupancy δ . The generating function for p_k^* is then given by:

$$\begin{aligned} \sum_{k=0}^{\infty} p_k^* x^k &= 1 - p_s + p_s \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} p_{k,j}^\delta x^k - p_s \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} x^k (p_s + (1 - p_s)(1 - p_t))^j p_{k,j}^\delta \\ &\quad + p_s \sum_{k=1}^{\infty} (x(p_s + (1 - p_s)(1 - p_t)))^k p_k. \end{aligned}$$

Now

$$\sum_{k=1}^{\infty} (x(p_s + (1 - p_s)(1 - p_t)))^k p_k = g_{D^*-1}(x(p_s + (1 - p_s)(1 - p_t))),$$

where $g_{D^*-1}(x)$ denotes the generating function of p_k . Furthermore,

$$\begin{aligned} \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} p_{k,j}^\delta x^k &= \sum_{k=1}^{\infty} \sum_{j=k}^{\infty} x^k p_j \mathbb{P}(\text{Bin}(j, \delta) = k) = \sum_{j=1}^{\infty} p_j \sum_{k=0}^j x^k \mathbb{P}(\text{Bin}(j, \delta) = k) \\ &= \sum_{j=1}^{\infty} p_j (1 - \delta + \delta x)^j = g_{D^*-1}(1 - \delta + \delta x). \end{aligned}$$

Similarly,

$$\sum_{k=1}^{\infty} \sum_{j=k}^{\infty} x^k (p_s + (1 - p_s)(1 - p_t))^j p_{k,j}^\delta = g_{D^*-1}((1 - \delta + \delta x)(p_s + (1 - p_s)(1 - p_t))).$$

Thus,

$$\begin{aligned} \sum_{k=0}^{\infty} p_k^* x^k &= 1 - p_s + p_s [g_{D^*-1}(1 - \delta + \delta x) - g_{D^*-1}((1 - \delta + \delta x)((1 - p_s)(1 - p_t) + p_s)) \\ &\quad + g_{D^*-1}(x((1 - p_s)(1 - p_t) + p_s))]. \end{aligned} \tag{15}$$

This is the generating function of the degree distribution of a tracing process on a network with excess degree distribution p_k . However, before the tracing process takes place, an epidemic modeled by a bond percolation process with occupancy π takes place. Thus, to obtain the generating function $G(x)$ of the degree distribution after the epidemic and the tracing process, we add the bond percolation process with bond occupancy probability π by substituting $x \rightarrow 1 - \pi + \pi x$ in (15):

$$\begin{aligned} G(x) &= 1 - p_s + p_s [g_{D^*-1}(\delta\pi(x - 1) + 1) \\ &\quad - g_{D^*-1}(((p_s - 1)p_t + 1)(\delta\pi(x - 1) + 1)) \\ &\quad + g_{D^*-1}((\pi(x - 1) + 1)((p_s - 1)p_t + 1))]. \end{aligned}$$

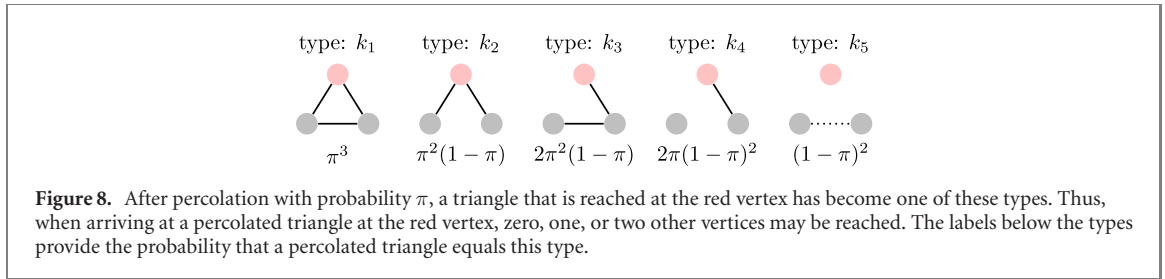
We then obtain the size of the giant outbreak $S = p_s - p_s g_D(1 - \pi + \pi u)$, where u is obtained by solving the implicit equation $u = G(u)$ and $g_D(x)$ is the generating function of the degree distribution.

Appendix D. Derivation of the giant outbreak size in clustered networks

Under bond percolation with probability π , a triangle from a given vertex can still be connected to its two triangle members, with probability $\pi^2(3 - 2\pi)$, it can connect to only one of its triangle members, with probability $2(1 - \pi)^2\pi$, or it can become disconnected from both other triangle members, with probability $(1 - \pi)^2$. Thus, for a vertex of triangle-degree k , the number of neighbors that are reachable through these triangles after bond percolation, has generating function $g_{D^*-1}(z) = ((1 - \pi)^2 + 2(1 - \pi)^2\pi z + \pi^2(3 - 2\pi)z^2)^k$. Let u denote the probability that a randomly chosen half-edge is not connected to the giant component. Similarly, let v denote the probability that following a randomly chosen triangle does not lead to the largest component. Then, after bond percolation with probability π ,

$$u = g_p(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2\pi v + \pi^2(3 - 2\pi)v^2), \tag{16}$$

$$v = g_q(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2\pi v + \pi^2(3 - 2\pi)v^2). \tag{17}$$



Adding site percolation with probability p_s results in

$$u = 1 - p_s + p_s g_p(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2 \pi v + \pi^2(3 - 2\pi)v^2), \tag{18}$$

$$v = 1 - p_s + p_s g_q(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2 \pi v + \pi^2(3 - 2\pi)v^2). \tag{19}$$

Let w denote that a vertex of degree 1 is traced by none of its offspring, so that

$$\mathbb{P}(\text{degree } k \text{ vertex not traced}) = (p_s + (1 - p_s)(1 - p_t))^k = w^k. \tag{20}$$

When a vertex is traced, an extra layer of percolation with parameter δ takes place, so that combined, this is percolation with parameter $\pi\delta$. However, the probability of this taking place, depends on the degree of the vertex after the first layer of percolation with parameter π . After the first layer of percolation with parameter π , triangles are percolated into 5 possible types, as illustrated in figure 8. In the leftmost two types, the percolated triangle contributes with two to the degree of the red vertex, the rightmost percolated triangle adds zero to the degree of the red vertex, and in the other two types, the percolated triangle adds one to the degree of the vertex. Thus, when we denote the number of percolated triangles of these types by k_1, k_2, \dots, k_5 , see figure 8, the degree of the vertex from the percolated triangles equals $2(k_1 + k_2) + k_3 + k_4$. The number of vertices that are reached through these percolated triangles equals $2(k_1 + k_2 + k_3) + k_4$.

Let $\hat{D}^{(1)}$ and $\hat{D}^{(2)}$ denote the degree of the number of edges and triangle-edges respectively after the tracing process. Furthermore, let k_6 denote the number of remaining half-edges attached to a vertex after percolating the half-edges with probability π . As the probability of a vertex not being traced equals w to the power of the degree after percolation with parameter π ,

$$\begin{aligned} \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} \mathbb{1}_{\text{not traced}} \right] &= \mathbb{E} \left[\mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} \mathbb{1}_{\text{not traced}} | k_1, \dots, k_5, k_6 \right] \right] \\ &= \mathbb{E} \left[x^{k_6} y^{2(k_1+k_2+k_3)+k_4} w^{2(k_1+k_2)+k_3+k_4+k_6} \right] \\ &= g_p \left(1 - \pi + \pi w x, (1 - \pi)^2 + 2(1 - \pi)^2 \pi w y \right. \\ &\quad \left. + 2\pi^2(1 - \pi)y^2 w + \pi^2 w^2 y^2 \right), \end{aligned}$$

where the last step used the probabilities in figure 8, and the generating function of the multinomial distribution. Also, when a vertex is traced, its neighbors are percolated with parameter δ . Therefore,

$$\begin{aligned} \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} \mathbb{1}_{\text{traced}} \right] &= \mathbb{E} \left[\mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} \mathbb{1}_{\text{traced}} | k_1, \dots, k_5, k_6 \right] \right] \\ &= \mathbb{E} \left[(1 - w^{2(k_1+k_2)+k_3+k_4+k_6}) \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} | k_1, \dots, k_5, k_6, \text{traced} \right] \right] \\ &= \mathbb{E} \left[x^{D^{(1,\pi\delta)}} y^{D^{(2,\pi\delta)}} \right] - \mathbb{E} \left[w^{2(k_1+k_2)+k_3+k_4+k_6} \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} | k_1, \dots, k_5, k_6, \text{traced} \right] \right], \end{aligned}$$

where $D^{(1,\pi\delta)}$ and $D^{(2,\pi\delta)}$ denote the degree and triangle-degree respectively of a vertex after percolation with parameter $\pi\delta$. Thus,

$$\mathbb{E} \left[x^{D^{(1,\pi\delta)}} y^{D^{(2,\pi\delta)}} \right] = g_p(1 - \pi\delta + \pi\delta x, (1 - \pi\delta)^2 + 2(1 - \pi\delta)^2 \pi\delta y + \pi^2 \delta^2 (3 - 2\pi\delta)y^2).$$

For the second term, we have to take into account that the percolated triangles of figure 8 are again percolated with parameter δ . Let $\phi_{t,i}$ denote the probability that a percolated triangle of type k_t reaches i neighbors after an extra layer of percolation with probability δ . For example, $\phi_{1,0} = (1 - \delta)^2$, the probability that a full triangle does not reach both its neighbors after percolation with parameter δ . Furthermore, let ζ_t denote the probability that after percolation of a full triangle with probability π , the percolated triangle is of type k_t . The ζ_t are given

in figure 8, and for example $\zeta_2 = \pi^2(1 - \pi)$. Then we obtain

$$\begin{aligned} & \mathbb{E} \left[w^{2(k_1+k_2)+k_3+k_4+k_6} \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} | k_1, \dots, k_5, k_6, \text{traced} \right] \right] \\ &= \mathbb{E} \left[w^{2(k_1+k_2)+k_3+k_4+k_6} (1 - \delta + \delta x)^{k_6} \prod_{t=1}^5 \left(\sum_{i=0}^2 \phi_{t,i} y^i \right)^{k_t} \right] \\ &= g_p \left(1 - \pi + \pi w(1 - \delta + \delta x), \sum_{t=1}^5 \zeta_t w^{a_t} \sum_{i=0}^2 \phi_{t,i} y^i \right), \end{aligned}$$

where $a_t = 2$ for $t = 1, 2$, $a_t = 1$ for $t = 3, 4$ and $a_t = 0$ for $w = 5$. Plugging in the expressions for ζ_t and $\phi_{t,i}$ and simplifying, yields

$$\begin{aligned} & \mathbb{E} \left[w^{2(k_1+k_2)+k_3+k_4+k_6} \mathbb{E} \left[x^{\hat{D}^{(1)}} y^{\hat{D}^{(2)}} | k_1, \dots, k_5, k_6, \text{traced} \right] \right] \\ &= g_p \left(1 - \pi + \pi w(1 - \delta + \delta u), (\pi + \pi(\delta - 1)w - 1)^2 + 2\pi\delta w(\pi\delta - 1)(\pi + \pi(\delta - 1)w - 1)v \right. \\ & \quad \left. - w\pi^2\delta^2(2(\pi - 1) + w(2\pi(\delta - 1) - 1))v^2 \right) \end{aligned}$$

Thus, when we let u denote the probability that a vertex that is reached by following a randomly chosen half-edges is not connected to the giant component, we obtain

$$\begin{aligned} u &= 1 - p_s + p_s \left(g_p(1 - \pi + \pi w u, (1 - \pi)^2 + 2(1 - \pi)^2\pi w v + 2\pi^2(1 - \pi)v^2 w + \pi^2 w^2 v^2) \right. \\ & \quad + g_p(1 - \pi\delta + \pi\delta u, (1 - \pi\delta)^2 + 2(1 - \pi\delta)^2\pi\delta v + \pi^2\delta^2(3 - 2\pi\delta)v^2) \\ & \quad - g_p \left(1 - \pi + \pi w(1 - \delta + \delta u), (\pi + \pi(\delta - 1)w - 1)^2 + 2\pi\delta w(\pi\delta - 1)(\pi + \pi(\delta - 1)w - 1)v \right. \\ & \quad \left. - w\pi^2\delta^2(2(\pi - 1) + w(2\pi(\delta - 1) - 1))v^2 \right) \end{aligned}$$

Similarly

$$\begin{aligned} v &= 1 - p_s + p_s \left(g_q(1 - \pi + \pi w u, (1 - \pi)^2 + 2(1 - \pi)^2\pi w v + \pi^2(1 - \pi)w^2 v^2) \right. \\ & \quad + g_q(1 - \pi\delta + \pi\delta u, (1 - \pi\delta)^2 + 2(1 - \pi\delta)^2\pi\delta v + \pi^2\delta^2(3 - 2\pi\delta)v^2) \\ & \quad - g_q \left(1 - \pi + \pi w(1 - \delta + \delta u), (\pi + \pi(\delta - 1)w - 1)^2 + 2\pi\delta w(\pi\delta - 1)(\pi + \pi(\delta - 1)w - 1)v \right. \\ & \quad \left. - w\pi^2\delta^2(2(\pi - 1) + w(2\pi(\delta - 1) - 1))v^2 \right). \end{aligned}$$

We can then find the remaining component size from $S = p_s - p_s g(1 - \pi + \pi u, (1 - \pi)^2 + 2(1 - \pi)^2\pi v + \pi^2(3 - 2\pi)v^2)$.

ORCID iDs

Ivan Kryven  <https://orcid.org/0000-0002-3964-2196>

Clara Stegehuis  <https://orcid.org/0000-0003-3951-5653>

References

- [1] Ball F G, Knock E S and O'Neill P D 2011 Threshold behaviour of emerging epidemics featuring contact tracing *Adv. Appl. Probab.* **43** 1048–65
- [2] Bianconi G, Sun H, Rapisardi G and Arenas A 2014 A message-passing approach to epidemic tracing and mitigation with apps (arXiv:2007.05277)
- [3] Blum M G B and Tran V C 2010 HIV with contact tracing: a case study in approximate bayesian computation *Biostatistics* **11** 644–60
- [4] Boguñá M, Pastor-Satorras R and Vespignani A 2003 Absence of epidemic threshold in scale-free networks with degree correlations *Phys. Rev. Lett.* **90** 028701
- [5] Bollobás B 1980 A probabilistic proof of an asymptotic formula for the number of labelled regular graphs *Eur. J. Comb.* **1** 311–6
- [6] Bollobás B and Riordan O 2015 An old approach to the giant component problem *J. Comb. Theory B* **113** 236–60
- [7] Cléménçon S, Arazoza H D, Rossi F and Tran V C 2015 A statistical network analysis of the HIV/AIDS epidemics in Cuba *Social Network Analysis and Mining* **5** 58
- [8] Cléménçon S, Tran V C and de Arazoza H 2008 A stochastic SIR model with contact-tracing: large population limits and statistical inference *J. Biol. Dynam.* **2** 392–414
- [9] Dhara S, van der Hofstad R and van Leeuwen J S H 2019 Critical percolation on scale-free random graphs: new universality class for the configuration model (arXiv:1909.05590)

- [10] Durrett R 2006 *Random Graph Dynamics* (Cambridge: Cambridge University Press)
- [11] Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, Parker M, Bonsall D and Fraser C 2020 Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing *Science* **368** eabb6936
- [12] Fraser C, Riley S, Anderson R M and Ferguson N M 2004 Factors that make an infectious disease outbreak controllable *Proc. Natl Acad. Sci.* **101** 6146–51
- [13] Kiss I Z, Green D M and Kao R R 2005 Disease contact tracing in random and clustered networks *Proc. R. Soc. B.* **272** 1407–14
- [14] Kojaku S, Hébert-Dufresne L and Ahn Y-Y 2020 The effectiveness of contact tracing in heterogeneous networks (arXiv:2005.02362)
- [15] Kretzschmar M E, Rozhnova G, Bootsma M C J, van Boven M, van de Wiggert J H H M and Bonten M J M 2020 Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study *The Lancet Public Health* **5** e452–9
- [16] May R M and Lloyd A L 2001 Infection dynamics on scale-free networks *Phys. Rev. E* **64** 066112
- [17] Newman M E J 2009 Random graphs with clustering *Phys. Rev. Lett.* **103** 058701
- [18] Pastor-Satorras R and Vespignani A 2001 Epidemic spreading in scale-free networks *Phys. Rev. Lett.* **86** 3200
- [19] Pastor-Satorras R and Vespignani A 2002 Epidemic dynamics in finite size scale-free networks *Phys. Rev. E* **65** 035108
- [20] Stegehuis C, van der Hofstad R and van Leeuwen J S H 2016 Epidemic spreading on complex networks with community structures *Sci. Rep.* **6** 29748
- [21] Wang B, Cao L, Suzuki H and Aihara K 2012 Impacts of clustering on interacting epidemics *J. Theor. Biol.* **304** 121–30