Heterogeneous Bond Percolation on Complex Networks: Application to Epidemiology

Antoine Allard[1], Pierre-André Noël[1], Louis J. Dubé[1] and Babak Pourbohloul[2,3]

[1]*Département de Physique, de Génie Physique et d'Optique, Université Laval, Québec, Canada*
[2]*University of British Columbia Centre for Disease Control, British Columbia, Canada*
[3]*Department of Health Care and Epidemiology, University of Bristish Columbia, British Columbia, Canada*

Abstract

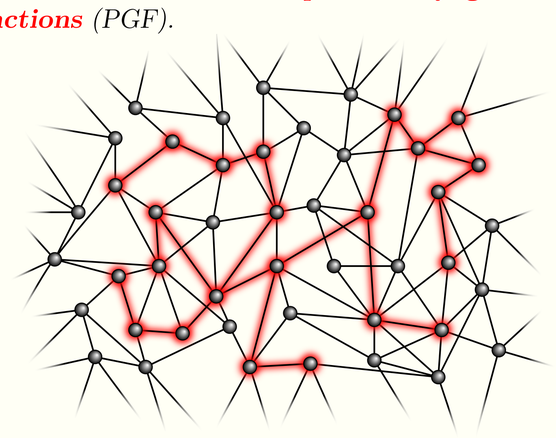
In the past decade, contact network epidemiology has revealed itself as a powerful approach to model the spread of infectious diseases in population. Using complex networks to capture the patterns of interactions between individuals that can lead to the transmission of an infectious disease, the bond percolation of such networks allows one to predict the outcome of outbreaks and therefore can help testing and improving public health strategies. We present a new bond percolation model of complex networks with an arbitrary degree distribution that explicitly incorporates heterogeneity in the bond occupation probability. Since several infectious diseases show heterogeneity in their transmissibility, mainly caused by intrinsic physiological and behavioural differences among individuals, this model will permit more realistic investigations and potentially lead to more precise intervention/prevention scenarios.

Contact Network Epidemiology

Mathematical modeling has gained credibility during the past two decades as ***a systematic quantitative decision-making tool for testing and optimizing intervention scenarios*** *before any disease strikes. Among the different methodologies, contact network epidemiology is a powerful analytical framework that applies bond percolation on complex networks to make epidemiological predictions and intervention recommendations:*

1) *Potential disease-causing contacts (****edges****) between individuals (****nodes****) are represented by a complex network, called the* ***contact network****, which is entirely random in all respects other than its specified distribution of* ***degrees*** *(number of edges connected to a node).*
2) *The disease is transmited from infectious nodes to their suceptible neighbours with a probability of transmission* ***T*** *called the* ***transmissibility****.*
3) *The final state of an outbreak (or epidemic) caused by an initial infected node (patient zero) is therefore given by the size of the* ***cluster*** *(group of nodes linked by occupied edges) it belongs to (see figure below).*

The probability that patient zero belongs to a cluster of a given size is analytically obtained with a bond percolation formalism based on ***probability generating functions*** *(PGF).*

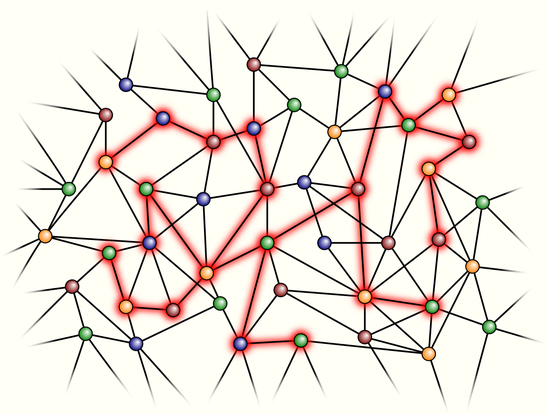


The Model

Existing models however still stand on simplifying assumptions restricting the categories of diseases that can be accurately described. For example, most existing models fail to adequately simulate the spread of diseases whose probability of transmission is correlated with intrinsic physiological and behavioural characteristics of individuals. In order to apply the right probability of transmission, one needs to know ***who infects whom****. Thus, we have considered:*

1) ***Networks with nodes divided into*** M ***types*** *(multitype networks) where w_i is the fraction of the network occupied by type-i nodes and $P_i(\boldsymbol{k})$ is their degree distribution.*
2) *A* ***transmissibility matrix*** $\mathbf{T}$ *whose elements, T_{ij}, are the probability that an infected type-i nodes infects its susceptible type-j neighbour.*
3) *The probability for an outbreak of size s caused by an initial infected node (patient zero) is given by the probability that it belongs to a cluster of the same size s (see figure below with $M = 4$).*

We have developped a generalized bond percolation formalism capable of predicting the outcome of outbreaks on multitype networks.



Numerical Results

In order to validate our formalism, we have performed extensive numerical simulations. To do so, we have generated ***networks of 100 000 nodes*** *on which we have simulated several outbreaks. We have used the following parameters:*

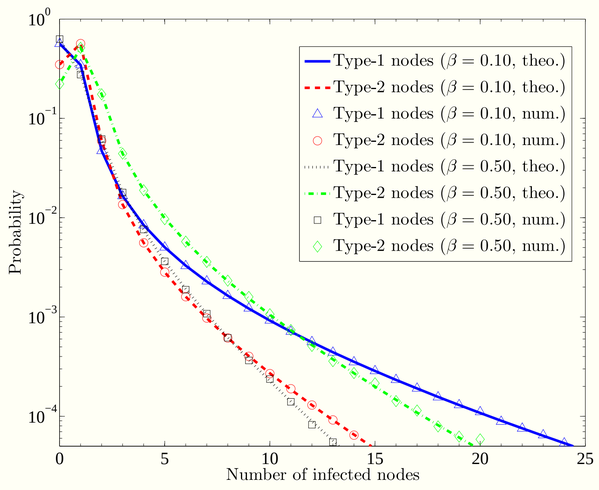
$$P_i(k_1,k_2) \propto (k_1+k_2)^{-\alpha_i}\, \exp\left(-\frac{(k_1+k_2)}{\kappa_i}\right) \cdot \binom{k_1+k_2}{k_1} p_{i1}^{k_1} p_{i2}^{k_2}$$

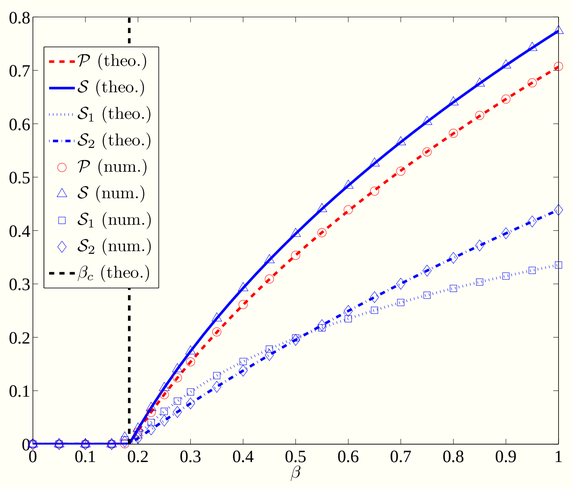
$$M = 2; \quad \boldsymbol{\alpha} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \boldsymbol{\kappa} = \begin{bmatrix} 8 \\ 10 \end{bmatrix}; \quad \boldsymbol{p} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}; \quad \mathbf{T} = \beta \begin{bmatrix} 0.95 & 0.99 \\ 0.61 & 1.00 \end{bmatrix}.$$

where β is a scaling parameter.

The figure on the left shows the outbreak size distribution for each node types obtained by numerical simulations and compared to the theoretical predictions of (1).

The figure on the right shows the probability of an epidemic ($\mathcal{P}$) and its relative size ($\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{S} \equiv \mathcal{S}_1 + \mathcal{S}_2$) obtained by numerical simulations for several values of β between 0 and 1. These results are compared to the theoretical values predicted by (2) and (3). The critical transmisibility parameter β_c has been obtained by solving the equation $\zeta_2 = 0$ (ζ_2 is not shown explicitly here).





PGF Formalism

Let us define $G_i(\boldsymbol{x}, \mathbf{T})$ as the PGF generating the occupied degree distribution of type-i nodes

$$G_i(\boldsymbol{x}, \mathbf{T}) = \sum_{\boldsymbol{k}=\boldsymbol{0}}^{\infty} P_i(\boldsymbol{k}) \prod_{l=1}^{M} \left[1 + (x_l - 1) T_{il}\right]^{k_l}$$

and $F_{ij}(\boldsymbol{x}, \mathbf{T})$ as the PGF generating the occupied excess degree distribution of type-j nodes reached from a type-i node

$$F_{ij}(\boldsymbol{x}, \mathbf{T}) = \frac{\partial G_j(\boldsymbol{x}, \mathbf{T})/\partial x_i}{\partial G_j(\mathbf{1}, \mathbf{T})/\partial x_i}.$$

Outbreak Size Distribution

The outbreak size distribution (i.e. the probability that a randomly chosen node belongs to a given finite size component) is generated by

$$K(\boldsymbol{x}; \mathbf{T}) = \sum_{i=1}^{M} w_i x_i G_i\big(\boldsymbol{H_i}(\boldsymbol{x}; \mathbf{T}); \mathbf{T}\big) \qquad (1)$$

where $\boldsymbol{H_i}(\boldsymbol{x}; \mathbf{T})$ stands for the vector

$$\big(H_{i1}(\boldsymbol{x}; \mathbf{T}), \ldots, H_{iM}(\boldsymbol{x}; \mathbf{T})\big)$$

whose elements are the solutions of the self-consistency equation

$$H_{ij}(\boldsymbol{x}; \mathbf{T}) = x_j F_{ij}\big(\boldsymbol{H_j}(\boldsymbol{x}; \mathbf{T}); \mathbf{T}\big).$$

Epidemic Threshold

Near the epidemic (or percolation) threshold, it can be shown that the average outbreak size $\langle s \rangle$ diverges like

$$\langle s \rangle \sim \zeta_M^{-1}$$

where ζ_M is a polynomial in the elements of $\mathbf{T}$ whose coefficients depend only of the network topology. Therefore the phase transition happens at $\zeta_M = 0$ which marks the point where the giant component first appears and defines the ***percolation threshold hypersurface****. This in turn defines the* ***critical transmissibility set*** *over which, for a given network structure $\big(w_i$ and $P_i(\boldsymbol{k})\big)$, there will be* ***a risk of a large-scale epidemic****.*

Epidemic Probability

The probability for a large-scale epidemic to occur (i.e. the probability that a randomly chosen node leads to the giant component) is given by

$$\mathcal{P} = 1 - \sum_{i=1}^{M} w_i G_i\big(\overrightarrow{\boldsymbol{h_i}}; \mathbf{T}\big) \qquad (2)$$

where $\overrightarrow{h_{ij}}$, the probability that an outgoing $i \to j$ edge does not lead to the giant component, is obtained by solving

$$\overrightarrow{h_{ij}} = F_{ij}\big(\overrightarrow{\boldsymbol{h_j}}; \mathbf{T}\big).$$

Epidemic Size

If such an epidemic should occur, the fraction of the network occupied by infected type-i nodes is

$$\mathcal{S}_i = w_i\big[1 - G_i\big(\overleftarrow{\boldsymbol{h_i}}; \mathbf{T}^{\dagger}\big)\big] \qquad (3)$$

where $\overleftarrow{h_{ij}}$, the probability that an incoming $j \to i$ edge does not link to the giant component, is obtained by solving

$$\overleftarrow{h_{ij}} = F_{ij}\big(\overleftarrow{\boldsymbol{h_j}}; \mathbf{T}^{\dagger}\big).$$

Conclusion

We have introduced a generalized formalism that successfully simulate the spread of infectious diseases with heterogeneous transmissibility on complex networks. This new model will allow more realistic simulations and will hopefully help improving public health strategies.

IRSC CIHR Instituts de recherche en santé du Canada / Canadian Institutes of Health Research
NSERC CRSNG
Michael Smith Foundation for Health Research
B.C. Ministry of Health