



Pierre-André Noël¹, Bahman Davoudi², Robert C. Brunham², Louis J. Dubé¹ and Babak Pourbohloul^{2,3}

¹Département de Physique, de Génie Physique, et d'Optique, Université Laval, Québec, Canada

²University of British Columbia Centre for Disease Control, British Columbia, Canada

³Department of Health Care and Epidemiology, University of British Columbia, British Columbia, Canada



BC Centre for Disease Control
AN AGENCY OF THE PROVINCIAL HEALTH SERVICES AUTHORITY

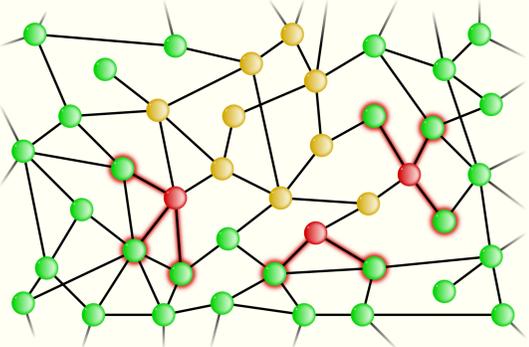
Abstract

In the last decade, many real-world systems have been shown to display complex network structures [1,2,3]. The dynamics on these networks has attracted considerable attention: for instance, the propagation in human populations of infectious diseases or of rumours indicates how crucial a good dynamical understanding is. While numerical simulations offer great generality, they are generally difficult to interpret and one often relies on analytical approaches to provide the necessary insights. Existing formalisms [4,5,6,7] partly include the full complexity of the systems at hand: structure of the networks (realistic, finite-size), time evolution and characterization of outcomes (e.g. outbreak vs. epidemics) to name a few. We have developed an analytical framework that improves over previous works in two complementary directions: **i.** finite-size effects have been identified and taken into account for discrete dynamics; and **ii.** continuous time evolution has been formulated for infinite networks. These are the first steps towards a formalism unifying continuous dynamics and finite-size networks. We will discuss the quantitative and qualitative differences with earlier studies and point out directions for further improvements.

The problem

We use a **contact network** in order to take into account the behaviour and interactions of individuals in the population.

- **Susceptible nodes** (○) represent individuals that have not acquired the disease yet.
- **Infectious nodes** (●) represent individuals that can currently transmit the disease.
- **Removed nodes** (○) represent individuals that have once acquired the disease but that can no longer transmit it (e.g. recovered, dead, quarantined).
- Nodes joined by a **link** (—) are said to be **neighbours**.
- The **degree** k of a node is its number of neighbours.
- The set $\{p_k\}_{k \in \mathbb{N}}$ such that a node selected at random has probability p_k of being of degree k is the **degree distribution** of the contact network.



We start with a network of N susceptible nodes respecting a given degree distribution. One of these nodes is randomly selected and becomes infectious; we refer to that node as **generation 0**. The term “generation g ” also refers to the time period when nodes of generation g are infectious.

Susceptible nodes that are neighbours to an infectious one of generation g (○) have probability T (**transmissibility**) to be infected and become part of generation $g+1$. Nodes of generation g then become removed.

Infinite networks

Counting neighbours

We define the **probability generating function** (PGF) for the degree of a random node [4,5] as

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k$$

respecting the normalization $G_0(1) = \sum_k p_k = 1$. The average degree is simply $z_1 = \sum_k k p_k = G_0'(1)$.

Except for generation 0, nodes have **higher probability to become infectious the more so the higher their degree**. The PGF for the number of neighbours of a generation g node (excluding the node from which it has been infected, when applicable) is thus

$$G_g(x) = \begin{cases} G_0(x) & (g=0) \\ \frac{\sum_k (k+1)p_{k+1}x^k}{\sum_k (k+1)p_{k+1}} = \frac{1}{z_1} G_0'(x) & (g \geq 1) \end{cases}$$

Evolution

We introduce the two variables PGF

$$\Psi_0^g(x, y) = \sum_{s, m=0}^{\infty} \psi_{sm}^g x^s y^m$$

generating the probability ψ_{sm}^g that a total of s nodes have been infected **before or at** generation g , m of which have been infected **at** generation g .

The probability for neighbours of a generation g node to be **already infected** at that generation **vanishes** like $1/N$ for arbitrarily large networks.

The number of infections directly caused by m' generation $g-1$ nodes is generated by $[G_{g-1}(1 + (xy-1)T)]^{m'}$. We thus have the recurrence relationship

$$\Psi_0^g(x, y) = \sum_{s', m'=0}^{\infty} \psi_{s'm'}^{g-1} x^{s'} y^{m'} [G_{g-1}(1 + (xy-1)T)]^{m'}$$

with the initial condition $\Psi_0^0(x, y) = xy$. The probability for s nodes to be infected at generation g is given by $p_s = \sum_m \psi_{sm}^g$. Although obtained differently, this result is identical to the one in [6].

Finite networks

Degree distribution in susceptibles

Since **high degree** nodes are more likely to get infected, their **susceptible** population **decreases faster** than the low degree one. To take this into account, we define susceptible only quantities $G_0^S(x; s)$, $p_k^S(s)$ and $z_1^S(s)$, functions of s and similar to their complete network counterpart, i.e. $G_0^S(x; s) = \sum_k p_k^S(s) x^k$ and $z_1^S(s) = G_0^S(1; s)$.

Normalization considerations and continuity approximations give the differential equation and solution

$$\frac{dp_k^S(s)}{ds} = \frac{p_k^S(s)}{N-s} \left(1 - \frac{k}{z_1^S(s)}\right) \xrightarrow{p_k^S(1)=p_k} G_0^S(x; s) = \frac{N-1}{N-s} G_0(x\theta(s))$$

$\theta(s)$ is fixed by evaluation at $x=1$, i.e. $G_0(\theta(s)) = \frac{N-s}{N-1}$.

It follows that the PGF for the potential number of infections directly caused by a generation g infectious node (the **finite-network counterpart** of $G_g(x)$) is given by

$$\tilde{G}_g(x; s, m) = \begin{cases} G_0(x) & (g=0) \\ \frac{(N-1)[G_0^S(x; s-m) - G_0^S(x; s)]}{x^m} & (g \geq 1) \end{cases}$$

Additional loss of transmissions

In contrast to the infinite-size case, **neighbours of an infectious node can be already infected**. Hence, transmissions are “lost” and this lowers the effective transmissibility (see details in [8])

$$\tilde{T}_{s'm'} = \frac{(N-s')}{m'} \frac{[1 - G_0^S(1 - \lambda_{s'm'}^S)]}{\tilde{G}_{g-1}^S(1; s', m')}$$

$$\lambda_{s'm'} = \frac{m'}{n_S + n_I + n_R} \quad \begin{matrix} n_S = (N-s')z_1^S(s') \\ n_I = m' \tilde{G}_{g-1}^S(1; s', m') \\ n_R = \eta(s' - m') \end{matrix}$$

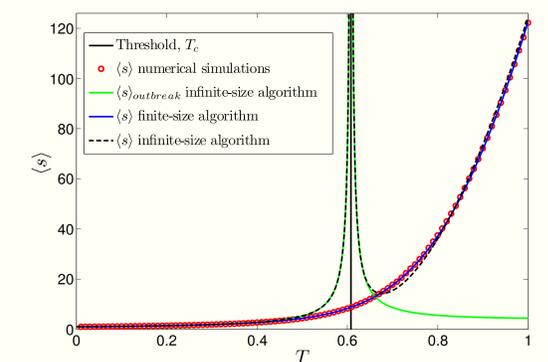
$$\frac{d\eta(s)}{ds} = \frac{z_1^S(s)}{z_1^S(s)} \left[(1-T) - \frac{(2-T)\eta(s)}{(N-s)z_1^S(s)} \right] \quad \begin{matrix} z_1^S(s) = G_0^S(1; s) \\ \eta(1) = (1-T)z_1^S(1) \end{matrix}$$

New recurrence relationship

Combining these two finite-size effects, we obtain

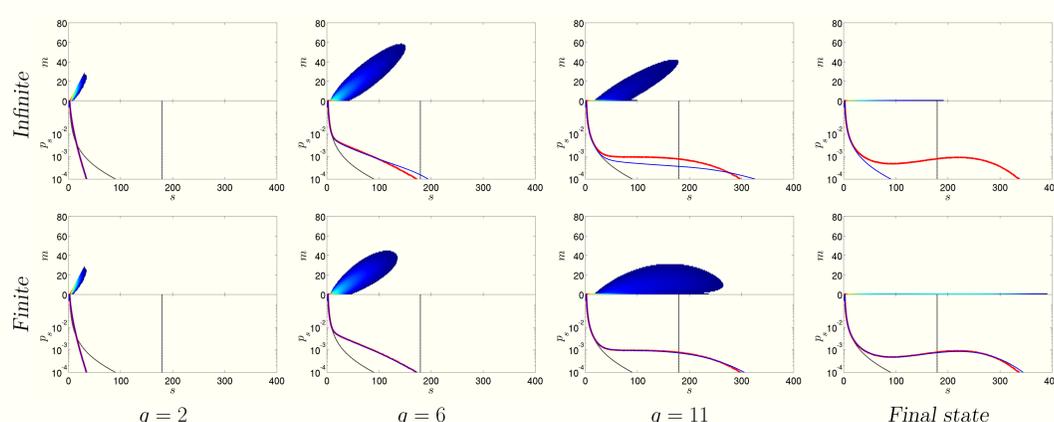
$$\Psi_0^g(x, y) = \sum_{s', m'=0}^{\infty} \psi_{s'm'}^{g-1} x^{s'} y^{m'} [\tilde{G}_{g-1}(1 + (xy-1)\tilde{T}_{s'm'}; s', m')]^{m'}$$

for the recurrence relationship in finite networks. This is a major improvement over the results of [7].



Results

Results obtained for a power-law network $p_k \propto k^{-\tau} e^{-k/\kappa}$ of $N=1000$ nodes with $\tau=2$, $\kappa=5$ and $T=0.8$. Red: numerical simulations; blue: present analytical results [8]; black: previous analytical results (outbreaks) [5].



Conclusion

Other developments (not presented here, see [8])

- Mapping of generations to continuous time.
- Effect of correlations when using discrete representation of continuous dynamics [9].

Future perspectives

- Development of a formalism simultaneously allowing continuous dynamics **and** finite-size effects.
- Additional network characteristics (e.g. clustering).
- Asymptotic limits and other analytical analysis.

