APPENDIX A: Simulation Methods

Network Generation

We use the configuration model (Molloy & Reed, 1995) to generate uncorrelated, random networks. This algorithm assigns d_i "half-edges" to each node i where d_i is the degree of node i, and then randomly connects pairs half-edges to create edges until there are no half-edges left. Self-loops (an edge from a node to itself) and duplicate edges are removed subsequently by randomly rewiring the edges.

Stochastic Epidemic Simulations

The stochastic epidemiological simulations in this work are based on a discrete-time, chain binomial, SIR model (Bailey, 1957.) At each time step, infected nodes recover with a probability γ; and susceptible nodes become infected with a probability $\left(1 - e^{-\tau n_i dt}\right)$, where τ is the transmission rate and n_i is the number of infectious nodes to which the susceptible individual is connected. All of our epidemic simulations were carried out on networks of size 10,000, and replicated 200 times. The recovery probability, γ, was fixed at 0.1 per unit time, while the transmission probability, τ, varied with the per contact probability of transmission, $T = \dfrac{\tau}{\tau + \gamma}$.

APPENDIX B: Statistical analysis of degree distributions from empirical contact networks

<u>Network Degree Distributions</u>

(a) *Poisson-distributed networks*: Poisson-distributed networks, also known as Erdos-Renyi random graphs, are characterized by the property that any two of the *N* vertices in the network are connected by an independent probability *p*. This property results in a degree distribution which is binomial, or Poisson in the limit. Thus, a Poisson network has degree distribution:

$$p(k) = \frac{\theta^k e^{-\theta}}{k!},$$

where $\lambda = pN$ is the average degree in the network.

(b) *Exponentially-distributed networks*: Exponentially-distributed random networks have more degree heterogeneity than the Poisson random network but a shorter tail than the scale-free network. An exponential random graph has degree distribution:

$$p_k = \left(1 - e^{\frac{-1}{\theta}}\right) e^{\frac{-1}{\theta}(k-1)},$$

where λ is a constant. The (k-1) term shifts the origin of the distribution to 1 since we are only considering connected graphs where all vertices have minimum degree equal to one.

(c) *Power-law-distributed network*: Many realistic networks have been identified recently as having power-law distributions, or being scale-free. Such networks are described by degree distributions of one of two forms. The first is a pure power-law distribution:

$$p(k) = \frac{k^{-\theta}}{\zeta(\theta)},$$

where $k > 1$; $\theta > 1$ is a constant, and $\zeta(\theta)$ is the Riemann Zeta function defined as $\sum_{k=1}^{\infty} k^{-\theta}$. The second form is a truncated power-law distribution:

$$p(k) = \frac{k^{-\theta} e^{-k/\kappa}}{C(\theta, \kappa)},$$

where θ is a constant, and $C(\theta, \kappa)$ is normalization factor defined as $Li_\theta\left(e^{-1/\kappa}\right)$, where $Li_n(x)$ is the nth polylogarithm of *x*. The exponential term results in an exponential cutoff around some finite degree, which is a property observed in many real-world networks due to system constraints. We choose $\kappa = 200$ since it is the largest degree observed in any of our data sets.

Maximum-Likelihood Parameter Estimation

In this section, we estimate parameters for the Poisson, exponential, pure power-law and truncated power-law using maximum-likelihood methods. Given a probability mass function, p_θ, and independently observed degrees $\{k_i\}$ from an empirical contact network, we can compute the probability mass associated with the observed data, parameterized by some parameter, θ: $p(k_1, k_2, ..., k_N | \theta)$. As a function of θ, $l(\theta) = p(k_1, k_2, ..., k_N | \theta)$ is the likelihood function. The maximum-likelihood estimate of parameter θ is then the value of that maximizes $L(\theta) = \ln(l(\theta))$.

(a) *Estimating parameter λ for the Poisson distribution*:

$$p_\theta(k) = \frac{\theta^k e^{-\theta}}{k!}$$

$$l(\theta) = p_\theta(k_1, k_2, ..., k_N | \theta) = \prod_{i=1}^{N} p_\theta(k_i) = \prod_{i=1}^{N} \frac{\theta^{k_i} e^{-\theta}}{k_i!}$$

$$L(\theta) = \sum_{i=1}^{N} \ln\left(\frac{\theta^{k_i} e^{-\theta}}{k_i!}\right)$$

The θ which maximizes $L(\theta)$ is found by solving $\frac{d}{d\theta} L(\theta) = 0$. This yields:

$$\hat{\theta}_{MLE} = \frac{1}{N} \sum_{i=1}^{N} k_i .$$

(b) *Estimating parameter λ for the Exponential distribution*:

$$p_k = \left(1 - e^{\frac{-1}{\theta}}\right) e^{\frac{-1}{\theta}(k-1)}$$

$$l(\theta) = p_\theta(k_1, k_2, ..., k_N | \theta) = \prod_{i=1}^{N} p_\theta(k_i) = \prod_{i=1}^{N} \left(1 - e^{-1/\theta}\right) e^{-1/\theta(k_i - 1)}$$

$$L(\theta) = \sum_{i=1}^{N} \ln\left(\left(1 - e^{-1/\theta}\right) e^{-1/\theta(k_i - 1)}\right)$$

The θ which maximizes $L(\theta)$ is found by solving $\frac{d}{d\theta} L(\theta) = 0$. This yields:

$$\left(\frac{1}{\hat{\theta}}\right)_{MLE} = \ln\left(\frac{\sum_{i=1}^{N} k_i}{\sum_{i=1}^{N} k_i - N}\right).$$

(c) *Estimating parameter γ for the pure power-law distribution*:

$$p_\theta(k) = \frac{k^{-\theta}}{\zeta(\theta)}$$

$$l(\theta) = p_\theta(k_1, k_2, ..., k_N \mid \theta) = \prod_{i=1}^{N} p_\theta(k_i) = \prod_{i=1}^{N} \frac{k_i^{-\theta}}{\zeta(\theta)}$$

$$L(\theta) = \sum_{i=1}^{N} \ln\left(\frac{k_i^{-\theta}}{\zeta(\theta)}\right)$$

The θ which maximizes $L(\theta)$ is found by solving $\frac{d}{d\theta} L(\theta) = 0$. This yields:

$$\frac{\zeta'(\theta)}{\zeta(\theta)} = \frac{-\sum_{i=1}^{N} \ln(k_i)}{N},$$

where, $\zeta(\theta)$ is the Riemann Zeta function. A table of values for the value of the ratio $\zeta'(\theta)/\zeta(\theta)$ can be generated using many computational software programs like Matlab and Mathematica.

(d) *Estimating parameter γ for the truncated power-law distribution*:

$$p_\theta(k) = \frac{k^{-\theta} e^{-k/\kappa}}{C(\theta, \kappa)}$$

$$l(\theta) = p_\theta(k_1, k_2, ..., k_N \mid \theta) = \prod_{i=1}^{N} p_\theta(k_i) = \prod_{i=1}^{N} \frac{k_i^{-\theta} e^{-k_i/\kappa}}{C(\theta, \kappa)}$$

$$L(\theta) = \sum_{i=1}^{N} \ln\left(\frac{k_i^{-\theta} e^{-k_i/\kappa}}{C(\theta, \kappa)}\right)$$

The θ which maximizes $L(\theta)$ is found by solving $\frac{d}{d\theta} L(\theta) = 0$. This yields:

$$\frac{\partial C/\partial \theta}{C(\theta, \kappa)} = \frac{-\sum_{i=1}^{N} \ln(k_i)}{N},$$

where, $C(\theta,\kappa)$ is defined as $Li_\theta\left(e^{-1/\kappa}\right)$, and κ is chosen as 200. A table of values for the value of the ratio $\dfrac{\partial C/\partial \theta}{C(\theta,\kappa)}$ can be generated using many computational software programs such as Matlab and Mathematica.

Standard Error for MLE Parameter Estimates

We can compute the expected error in the maximum-likelihood parameter estimates based on the curvature method (Schrago, 2006). The maximum-likelihood estimate of parameter θ is the value of θ that maximizes $L(\theta) = \ln\left(l(\theta)\right)$, where $l(\theta) = p(k_1,k_2,...,k_N \mid \theta)$ as described in the previous section.

We obtain the Taylor series approximation of $L(\theta)$ at $\hat{\theta}$:

$$L(\theta) = L(\hat{\theta}) + (\theta - \hat{\theta})\frac{dL}{d\theta} + \frac{1}{2}(\theta - \hat{\theta})^2 \frac{d^2L}{d\theta^2}.$$

Since, $\dfrac{dL(\hat{\theta})}{d\theta} = 0$, this reduces to:

$$L(\theta) = L(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^2 \frac{d^2L}{d\theta^2}.$$

Thus the standard error in the MLE parameter estimate is taken to be

$$SE = \frac{1}{\sqrt{-\dfrac{d^2L(\hat{\theta})}{d\theta^2}}}.$$

Goodness-of-Fit

To measure the goodness-of-fit of our proposed distributions and to select the best fitting distribution for the empirical data, we compute the Akaike Information Criterion (AIC) (Akaike, 1974) which is defined as follows:
$$AIC = 2k - 2\ln L.$$
Here, k is the number of parameters in the proposed distribution and L is the likelihood function. The AIC is based on the principle of entropy maximization and provides an absolute measure for model selection. For small sample sizes, a modified version of AIC can be used:

$$AIC = 2k - 2\ln L + \frac{2k(k+1)}{n-k-1}.$$

Once the AIC has been computed, we can compute Akaike weights (w_i) to provide a relative measure of the strength of evidence for each model. We first measure delta AIC, a measure of each model relative to the best model :

$$\Delta_i = AIC_i - \min_i(AIC_i).$$

Then, Akaike weights are computed as the ratio of delta AIC values for each model relative to the whole set of R candidate models:

$$w_i = \frac{\exp(-\Delta_i/2)}{\sum_{i=1}^{R}\exp(-\Delta_i/2)}$$

The Akaike weight for a model thus indicates the probability that the model is the best among the whole set of candidate models. In Table 1, we provide the Akaike weight results for our four candidate models. Similar results were attained with other tests such as the Kolmogorov-Smirnov test and Kullback-Liebler information test.

	Poisson	Exponential	Pure power law	Trunc. power law
Vancouver	0.00	100.00	0.00	0.00
Portland	0.00	100.00	0.00	0.00
Karate Club	0.01	99.99	0.00	0.00
Atlanta HS	0.00	96.49	2.29	1.22
Midwest Town	0.00	100.00	0.00	0.00
Colorado Springs	0.00	100.00	0.00	0.00

Table B1: The Akaike weights for the four candidate distributions for each of the datasets. The Akaike weight (shown here as a percentage) for a model distribution represents the chances of that model being the best one among the set of candidate models.

References

Akaike, H. 1974 A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19(6)**, 716-723.

Schrago, C. 2006 An empirical examination of the standard errors of maximum likelihood phylogenetic parameters under the molecular clock via bootstrapping. *Genet. Mol. Res.* **5(1)**, 233-241

Appendix C: Greedy Rewiring Algorithm

Our greedy rewiring algorithm takes as input a connected, undirected network G, and rewires selected edges such that the total number of edges in the network and thus the mean of the degree distribution remain unchanged. The procedure avoids duplicating edges, creating self loops, and disconnecting nodes from the network.

For our analysis, we begin with a regular random network and rewire edges until the degree distribution becomes approximately exponential. In particular, the algorithm runs until the coefficient of variation (standard deviation/mean) of the degree distribution is equal to one (this characterizes an exponential distribution). The specific algorithm is described below and illustrated in Figure C1. Figure C2 illustrates the evolution of the degree distribution.

GREEDY_REWIRING(G)

    1. while (CV < 1)

    2.   select a random node, i of G

    3.   select a random edge, e_{ij}, of node i such that degree(j) > 1

    4.   select a random edge, e_{jm}, of node j, with probability $k_m \big/ \sum\limits_{\{e_{jm}\}} k_m$,

                where k_m = degree(m), such that $m \neq i$ and $m \notin$ Neighbors(i)

    5.   remove edge e_{ij}

    6.   add edge e_{im}
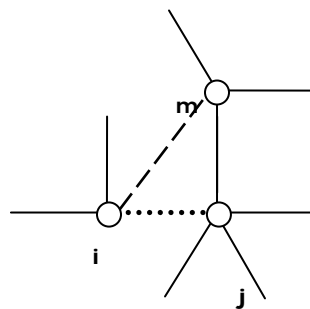
Figure C1: One step of the greedy rewiring process: node i is chosen at random, edge e_{ij} is selected at random from the edges of node i, and edge e_{jm} is selected at random from the edges of node j, with probabilities proportional to the degree of node m. Edge e_{ij} (shown with dotted line) is removed and edge e_{im} (shown with a dashed line) is added.
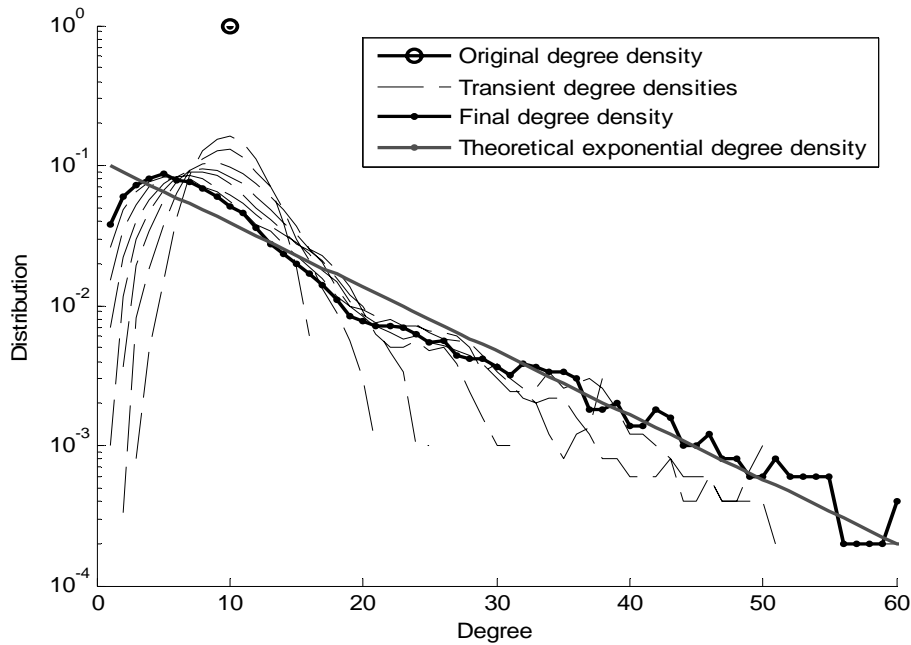
Figure C2: The evolution of the degree distribution in the greedy rewiring procedure (shown on a linear-log scale). We start with a regular random network of mean 10 (in fact, all the nodes have degree 10) and rewire the network until its degree distribution is approximately exponential mean 10 (degree distribution shown with thick black line). A theoretical exponential degree distribution (of mean 10) is also shown (with a gray line).

Appendix D: Analytical Methods for Epidemic Dynamics

We review the equations of each of the analytical models (Section 5.1) and modifications to the homogenous-mixing model (Section 5.2) for a population of size N. See the original papers on these models for complete details (references below). We have standardized some of the notation used in the models.

All models were parameterized such that the probability of transmission per contact during the infectious period, T, is constant across all models, and the infectious rate is equal to 0.1.

A) <u>Percolation Model</u> (Newman, 2002)

$$\xi = 1 - \sum_k p_k \left(1 + (u-1)T\right)^k$$

$$u = \frac{\sum_k k p_k \left(1 + (u-1)T\right)^{k-1}}{\sum_k k p_k}$$

ξ: final size of epidemic

p_k degree distribution of contact network

u: probability that a node at the end of a random edge is uninfected

T: per contact probability of transmission

B) <u>Pair-approximation Model</u> (Keeling, 1999)

$$\frac{dS}{dt} = -\tau [SI]$$

$$\frac{dI}{dt} = \tau [SI] - gI$$

$$\frac{d[SI]}{dt} = \frac{\tau [SI]}{\alpha S}\left((\alpha-1)[SS] - (\alpha-1)[SI] - \alpha S\right) - \gamma [SI]$$

$$\frac{d[SS]}{dt} = -2 \frac{\tau [SI]}{\alpha S}(\alpha-1)[SS]$$

$$S + I + R = N$$

S: number of nodes susceptible at time t

I: number of nodes infectious at time t

R: number of nodes recovered at time t

τ: constant rate at which infectious nodes infect neighbors

γ: constant rate at which infectious nodes become recovered

α average number of contacts between susceptible and infectious nodes

$[SI]$:   number of contacts between susceptible and infectious nodes

$[SS]$:   number of contacts between susceptible nodes

C) Heterogeneous-Mixing Model (Moreno et al, 2002)

$$\frac{dS_k}{dt} = -\tau k S_k(t)\Theta(t)$$

$$\frac{dI_k}{dt} = \tau k S_k(t)\Theta(t) - \gamma I_k(t)$$

$$\Theta(t) = \frac{\sum_k k p_k I_k(t)}{\sum_k k p_k}$$

$$S_k(t) + I_k(t) + R_k(t) = 1$$

$$S(t) = \sum_k p_k S_k(t)$$

$$I(t) = \sum_k p_k I_k(t)$$

$$R(t) = \sum_k p_k R_k(t)$$

S_k, S : fraction of nodes of degree k susceptible at time t; total fraction of susceptible nodes

I_k, I : fraction of nodes of degree k infectious at time t; total fraction of infectious nodes

R_k, R : fraction of nodes of degree k recovered at time t; total fraction of recovered nodes

τ : constant rate at which susceptible nodes are infected

γ : constant rate at which infectious nodes become recovered

$\Theta(t)$: probability of contact with an infectious node

p_k : probability that a node in the network has degree

D) Dynamical PGF Model (Volz, in review)

$$\frac{d\theta}{dt} = -\tau p_I \theta$$

$$\frac{dp_I}{dt} = \tau p_S p_I \theta \frac{g''(\theta)}{g'(\theta)} - \tau p_I (1 - p_I) - p_I \gamma$$

$$\frac{dp_S}{dt} = \tau p_S p_I \left(1 - \theta \frac{g''(\theta)}{g'(\theta)}\right)$$

$$\frac{dS}{dt} = -\tau p_I \theta g'(\theta)$$

$$\frac{dI}{dt} = \tau p_I \theta g'(\theta) - \gamma I$$

$$S(t) + I(t) + R(t) = 1$$

S : fraction of nodes susceptible at time t

I : fraction of nodes infectious at time t

R : fraction of nodes recovered at time t

τ : constant rate at which infectious nodes infect a neighbor

γ : constant rate at which infectious nodes become recovered

θ : The fraction of degree one nodes that remain susceptible at time t

p_I : probability that a susceptible node has an infectious neighbor

p_S : probability that a susceptible node has an susceptible neighbor

$g(x)$: probability generating function for the degree distribution of the contact network

E) R_0-based Modification (Aparicio & Pascual, 2006)

$$\frac{dS}{dt} = -\gamma_e R_0^{net}(S/N)I$$

$$\frac{dI}{dt} = \gamma_e R_0^{net}(S/N)I - \gamma_e I$$

$$\frac{dY}{dt} = \gamma_e I - gY$$

$$S + I + Y + R = N$$

$$\gamma_e = \tau + \gamma$$

$$R_0^{net} = \frac{\tau}{\tau + \gamma} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$$

S : number of nodes susceptible at time t

I : number of nodes infected and infectious at time t

Y : number of nodes infected but inactive at time t

R : fraction of nodes recovered at time t

τ : constant rate at which infectious nodes infect susceptible neighbors

γ : constant rate at which infected nodes become recovered. ($1/\gamma$ is thus the infectious period).

γ_e : constant rate at which infectious nodes become inactive. ($1/\gamma_e$ is thus the effective portion of the infectious period)

g : constant rate at which inactive nodes recover. ($1/g$ is thus the ineffective portion of the infectious period)

α : average degree of nodes in network

R_0^{net} : network reproductive ratio (average number of secondary cases early in the outbreak)

F) Power-law Infection Rate Modification (Stroud et al, 2006)

$$\frac{dS}{dt} = -\tau \alpha (S/N)^{\nu} I$$

$$\frac{dI}{dt} = \tau \alpha (S/N)^{\nu} I - \gamma I$$

$$S + I + R = N$$

S : number of nodes susceptible at time t

I : number of nodes infected and infectious at time t

R : fraction of nodes recovered at time t

τ : constant rate at which infectious nodes infect susceptible neighbors

α : average degree of nodes in network

γ : constant rate at which infected nodes become recovered. ($1/\gamma$ is thus the infectious period).

ν : power law exponent specific to each network; estimated from simulation results.