Marketing Models of SIR and SIS Type on Random Networks

Reinhard Illner, Victoria // joint work with J. Ma, M. Li, R. Edwards

Hong Kong, March 2016

Synopsis

1. Random Graphs of Configuration Type
2. The Miller-Volz Model
3. Marketing: Generalized Bass Models
4. Marketing with Product Fatigue: SIS models
5. Tests

Random graphs of configuration type

Graph with N nodes (typically, N large); specified degree distribution, $\{P_k\}$, where P_k is the probability that a randomly chosen vertex (node) has degree k.

Random graphs of configuration type

Graph with N nodes (typically, N large); specified degree distribution, $\{P_k\}$, where P_k is the probability that a randomly chosen vertex (node) has degree k.

Construction principle: assume N nodes; draw a degree sequence $\{k_i\}$ from the given distribution;

Random graphs of configuration type

Graph with N nodes (typically, N large); specified degree distribution, $\{P_k\}$, where P_k is the probability that a randomly chosen vertex (node) has degree k.

Construction principle: assume N nodes; draw a degree sequence $\{k_i\}$ from the given distribution; assign k_i "stubs" to vertex i; choose pairs of stubs randomly, and connect these stubs to get edges; multiple connections are not allowed. Finally, discard leftover stubs.

Some terminology

$$\langle k^n \rangle = \sum_{k=0}^{\infty} k^n P_k$$

$(n = 1 :$ average number of stubs). Variance:

$$Var[k] = \langle k^2 \rangle - \langle k \rangle^2$$

Some terminology

$$\langle k^n \rangle = \sum_{k=0}^{\infty} k^n P_k$$

$(n = 1 :$ average number of stubs). Variance:

$$Var[k] = \langle k^2 \rangle - \langle k \rangle^2$$

Generating function $\Psi(x) = \sum P_k x^k$. Then

$$\Psi(1) = 1, \quad \Psi'(1) = \langle k \rangle,$$

Ψ encodes all we know about the graph.

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz: SIR Dynamics in random networks, JM Biol. 62

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz:
SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz: SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

A susceptible node with degree k:

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz:
SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

A susceptible node with degree k:

remains susceptible while none of its k edges has transmitted infection.

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz:
SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

A susceptible node with degree k:

remains susceptible while none of its k edges has transmitted infection.
Transmission is a necessary condition for infection.

If $\theta(t)$ is the probability that a random edge has not transmitted by time t, then the probability that the node remains susceptible is θ^k.

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz: SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

A susceptible node with degree k:

remains susceptible while none of its k edges has transmitted infection.
Transmission is a necessary condition for infection.

If $\theta(t)$ is the probability that a random edge has not transmitted by time t, then the probability that the node remains susceptible is θ^k.

Speed at which the susceptible node becomes infected is determined by the dynamics of θ.

An Epidemiology Precursor: The Miller-Volz Model

Ref.: Miller JC (2011) A note on a paper by E. Volz:
SIR Dynamics in random networks, JM Biol. 62

Volz, E.M. (2008) SIR dynamics in random networks, JM Biol. 56.

A susceptible node with degree k:

remains susceptible while none of its k edges has transmitted infection.
Transmission is a necessary condition for infection.

If $\theta(t)$ is the probability that a random edge has not transmitted by time t, then the probability that the node remains susceptible is θ^k.

Speed at which the susceptible node becomes infected is determined by the dynamics of θ.

Let $\{P_k\}$ be the degree distribution, with generating function

$$\Psi(x) = \sum_{n=0}^{\infty} P_k x^k \,. \tag{1}$$

Let $\{P_k\}$ be the degree distribution, with generating function

$$\Psi(x) = \sum_{n=0}^{\infty} P_k x^k \,. \tag{1}$$

The probability that a randomly selected node remains susceptible at time t is

$$\Psi(\theta) = P_0 + \theta P_1 + \theta^2 P_2 + \dots$$

(because... : Let's draw the possibilities!)

Let $\{P_k\}$ be the degree distribution, with generating function

$$\Psi(x) = \sum_{n=0}^{\infty} P_k x^k \,. \tag{1}$$

The probability that a randomly selected node remains susceptible at time t is

$$\Psi(\theta) = P_0 + \theta P_1 + \theta^2 P_2 + \dots$$

(because... : Let's draw the possibilities!)
$\implies$ the fraction of nodes that are susceptible at time t is

$$S = \Psi(\theta)\,. \tag{2}$$

The fraction of infectious nodes $I(t)$ increases because susceptible nodes get infected, and decreases because infected nodes recover (with a constant rate γ per node). Thus,

$$\frac{dI}{dt} = -\frac{dS}{dt} - \gamma I = -\Psi'(\theta)\frac{d\theta}{dt} - \gamma I\,.$$

To describe the dynamics of θ, let $P_I(t)$ be the probability that the neighbour connected by this edge is infectious at time t, and β be the transmission rate along an random edge.

To describe the dynamics of θ, let $P_I(t)$ be the probability that the neighbour connected by this edge is infectious at time t, and β be the transmission rate along an random edge.
Then βP_I is the attack rate on the edge, and θ is the survival probability $\Longrightarrow$

$$\theta' = -\beta P_I \theta \,.$$

Let $\phi = P_I \theta$: the probability that a random edge of class θ connects a (degree-1, see above) susceptible node to an infectious node. Then, the above equation becomes

$$\theta' = -\beta\phi \,. \tag{3}$$

We treat ϕ as a second dependent variable!

We treat ϕ as a second dependent variable!
An edge leaves class ϕ either because transmission occurred along it (at rate β),

We treat ϕ as a second dependent variable!
An edge leaves class ϕ either because transmission occurred along it (at rate β), or because the infectious node recovers (at rate γ).

We treat ϕ as a second dependent variable!
An edge leaves class ϕ either because transmission occurred along it (at rate β), or because the infectious node recovers (at rate γ). Further, an edge of a susceptible node enters class ϕ because its other neighbour becomes infected:

We treat ϕ as a second dependent variable!
An edge leaves class ϕ either because transmission occurred along it (at rate β), or because the infectious node recovers (at rate γ). Further, an edge of a susceptible node enters class ϕ because its other neighbour becomes infected:
this happens at a rate $-h'(t)$, where $h(t)$ is the probability that we arrive at a susceptible node when following a random edge that has not transmitted disease.
Thus,

$$\phi' = -\beta\phi - \gamma\phi - h'(t)\,.$$

Finally, we model $h(t)$: The probability that we arrive at a degree-k node when following a random edge is

$$q_k = kP_k / \sum_{k=0}^{\infty} kP_k = kP_k / \Psi'(1).$$

The probability that this node is susceptible is $\theta^{k-1} q_k$, and we arrive at

$$h(t) = \sum_{k=0}^{\infty} \theta^{k-1} \frac{kP_k}{\Psi'(1)} = \frac{\Psi'(\theta)}{\Psi'(1)}. \tag{4}$$

The equation for ϕ' can now be rewritten as

$$\phi' = -(\beta + \gamma)\phi - \frac{\Psi''(\theta)}{\Psi'(1)}\theta' = -(\beta + \gamma)\phi + \beta\phi\frac{\Psi''(\theta)}{\Psi'(1)}\,. \qquad (5)$$

The dynamics of the disease are thus determined by (3) and (5). The fraction of nodes which are infectious at time t changes according to

$$I'(t) = -S' - \gamma I = \beta\phi\Psi'(\theta) - \gamma I\,. \qquad (6)$$

Homogeneous Limit

In a homogeneously mixed population, seen as a contact network on a complete graph, the Miller-Volz model becomes the classic SIR model

$$\frac{dS}{dt} = -qSI\,,$$
$$\frac{dI}{dt} = qSI - \gamma I\,,$$

where $q = (N-1)\beta$ is the per capita transmission rate in a population of N individuals.

Marketing: Generalized Bass models

Cllassical Bass model (Frank Bass, 1969) assumes that a well-mixed potential buyer population divides into a fraction which has bought a product, $F(t)$, and $1 - F(t)$, the fraction that has not bought but consists of potential buyers (called "susceptibles", for obvious reasons):

$$\frac{dF}{dt} = \alpha(1 - F) + qF(1 - F)\,. \tag{7}$$

α is the rate of spontaneous conversion into buyers due to advertising; q represents the adoption rate of the product due to word-of-mouth recruitment of a potential buyer.

Marketing: Generalized Bass models

Cllassical Bass model (Frank Bass, 1969) assumes that a well-mixed potential buyer population divides into a fraction which has bought a product, $F(t)$, and $1 - F(t)$, the fraction that has not bought but consists of potential buyers (called "susceptibles", for obvious reasons):

$$\frac{dF}{dt} = \alpha(1 - F) + qF(1 - F)\,. \tag{7}$$

α is the rate of spontaneous conversion into buyers due to advertising; q represents the adoption rate of the product due to word-of-mouth recruitment of a potential buyer.

If we set $F = I, 1 - F = S, \alpha = 0$, the equation is

$$\frac{dI}{dt} = qIS,$$

exactly the SIR model for $\gamma = 0$. "Word of mouth" propagation of a product "is" an infectious disease without recovery.

Adapting the Miller-Volz ideas, I: No advertising

Identical to the Miller-Volz model, but NO recovery. Use subscript W to label the random "word-of-mouth" network. Variables will be θ_W, ϕ_W; the equations will be

$$\frac{d\theta_W}{dt} = -\beta\phi_W \,, \tag{8a}$$

$$\frac{d\phi_W}{dt} = -\beta\phi_W + \beta\phi_W \frac{\Psi''(\theta_W)}{\Psi'(1)} \,, \tag{8b}$$

$$S = \Psi(\theta_W) \,, \tag{8c}$$

$$\frac{dI}{dt} = \beta\phi_W \Psi'(\theta_W) \,. \tag{8d}$$

Adapting Miller-Volz ideas, II: marketing one product

We incorporate an (outside) marketing node A in the model, connecting to all nodes in W.

Adapting Miller-Volz ideas, II: marketing one product

We incorporate an (outside) marketing node A in the model, connecting to all nodes in W. Denote by $\theta_A(t)$ the fraction of all **edges with source** A which have not transmitted by time t.

Adapting Miller-Volz ideas, II: marketing one product

We incorporate an (outside) marketing node A in the model, connecting to all nodes in W. Denote by $\theta_A(t)$ the fraction of all **edges with source** A which have not transmitted by time t. A node in the word-of-mouth network W remains susceptible if information has neither been transmitted by word of mouth nor by A-edges, so Equation (8c) must now be modified to become

$$S(t) = \Psi(\theta_W)\theta_A \,. \qquad (9)$$

Adapting Miller-Volz ideas, II: marketing one product

We incorporate an (outside) marketing node A in the model, connecting to all nodes in W. Denote by $\theta_A(t)$ the fraction of all **edges with source** A which have not transmitted by time t. A node in the word-of-mouth network W remains susceptible if information has neither been transmitted by word of mouth nor by A-edges, so Equation (8c) must now be modified to become

$$S(t) = \Psi(\theta_W)\theta_A \,. \tag{9}$$

Edges in the class θ_A only leave their class because of transmission. Thus,

$$\frac{d\theta_A}{dt} = -\alpha\theta_A \,. \tag{10}$$

Probability that the source node of a θ_W edge is susceptible should now be modelled as

$$h(t) = \frac{\Psi'(\theta_W)}{\Psi'(1)}\theta_A \,.$$

Reduction of this probability by the infection of a susceptible source causes a θ_W edge to enter ϕ_W. Thus,

$$\frac{d\phi_W}{dt} = -\beta\phi_W - h'(t) = -\beta\phi_W + \beta\phi_W \frac{\Psi''(\theta_W)}{\Psi'(1)}\theta_A + \alpha\theta_A \frac{\Psi'(\theta_W)}{\Psi'(1)} \,.$$

The A-edges have no direct influence on the θ_W edges.

Initial Conditions

Initially, every node is susceptible, and no edge has transmitted. The complete network marketing model is:

$$S(t) = \Psi(\theta_W)\theta_A\,, \tag{11a}$$

$$\frac{d\theta_A}{dt} = -\alpha\theta_A\,, \tag{11b}$$

$$\frac{d\theta_W}{dt} = -\beta\phi_W\,, \tag{11c}$$

$$\frac{d\phi_W}{dt} = -\beta\phi_W + \beta\phi_W\frac{\Psi''(\theta_W)}{\Psi'(1)}\theta_A + \alpha\theta_A\frac{\Psi'(\theta_W)}{\Psi'(1)}\,, \tag{11d}$$

with initial conditions $S(0) = 1$, $\theta_W(0) = 1$, $\phi_W(0) = 0$, $\theta_A(0) = 1$.

Two competing products

In addition to A (e.g., "Apple") there is now a second external node B (e.g., "Microsoft"), also connected by an edge to each node in W. B competes with A to place their product.

Two competing products

In addition to A (e.g., "Apple") there is now a second external node B (e.g., "Microsoft"), also connected by an edge to each node in W. B competes with A to place their product.

Susceptibles (initially all nodes in W) can turn into two different kinds of buyers, I_A and I_B. Probability θ_W is defined as before, but now there are fractions ϕ_A and ϕ_B of edges in W which have not transmitted but originate in an I_A or I_B, respectively.

Two competing products

In addition to A (e.g., "Apple") there is now a second external node B (e.g., "Microsoft"), also connected by an edge to each node in W. B competes with A to place their product.

Susceptibles (initially all nodes in W) can turn into two different kinds of buyers, I_A and I_B. Probability θ_W is defined as before, but now there are fractions ϕ_A and ϕ_B of edges in W which have not transmitted but originate in an I_A or I_B, respectively.

Allow two possibly different word-of-mouth transmission rates β_A and β_B. By θ_A and θ_B we denote the fractions of edges from A into W (and B into W) which have not transmitted.
As before, we assume that, initially, every node is susceptible, and no edge has transmitted. This leads to the following equations:

$$
\begin{aligned}
S(t) &= \Psi(\theta_W)\theta_A\theta_B\,, && \text{(12a)}\\
\theta_A' &= -\alpha_A\theta_A\,, && \text{(12b)}\\
\theta_B' &= -\alpha_B\theta_B\,, && \text{(12c)}\\
\theta_W' &= -\beta_A\phi_A - \beta_B\phi_B\,, && \text{(12d)}\\
\phi_A' &= -\beta_A\phi_A + \frac{\Psi''(\theta_W)}{\Psi'(1)}\beta_A\phi_A\theta_A\theta_B + \alpha_A\frac{\Psi'(\theta_W)}{\Psi'(1)}\theta_A\theta_B && \text{(12e)}\\
\phi_B' &= -\beta_B\phi_B + \frac{\Psi''(\theta_W)}{\Psi'(1)}\beta_B\phi_B\theta_A\theta_B + \alpha_B\frac{\Psi'(\theta_W)}{\Psi'(1)}\theta_A\theta_B && \text{(12f)}
\end{aligned}
$$

Initial conditions as in (11), except $\phi_A(0) = \phi_B(0) = 0$, $\theta_A(0) = \theta_B(0) = 1$.

The gain terms in the third and fourth equations add up to $-h'$, where $h(t)$ now is given by

$$h(t) = \frac{\Psi'(\theta_W)}{\Psi'(1)} \theta_A \theta_B.$$

The gain terms in the third and fourth equations add up to $-h'$, where $h(t)$ now is given by

$$h(t) = \frac{\Psi'(\theta_W)}{\Psi'(1)}\theta_A\theta_B.$$

This h is the probability of reaching a susceptible node if one follows a W-edge from a susceptible inside W. The rate of change of h tells us how fast this susceptible turns into an infected, and we have to distinguish whether it turns into an I_A or an I_B node. The four terms in h' are split in just the right way to reflect this.

Equations (12) allow to compute the fractions I_A, I_B of users who have bought products A or B, respectively: Compute

$$\begin{aligned} S' =& \Psi'(\theta_W)\theta'_W\theta_A\theta_B + \Psi(\theta_W)\theta'_A\theta_B + \Psi(\theta_W)\theta_A\theta'_B \\ =& -\beta_A\Psi'(\theta_W)\phi_A\theta_A\theta_B - \alpha_A\Psi(\theta_W)\theta_A\theta_B \\ & -\beta_B\Psi'(\theta_W)\phi_B\theta_A\theta_B - \alpha_B\Psi(\theta_W)\theta_A\theta_B\,. \end{aligned}$$

Equations (12) allow to compute the fractions I_A, I_B of users who have bought products A or B, respectively: Compute

$$\begin{aligned} S' =& \Psi'(\theta_W)\theta'_W\theta_A\theta_B + \Psi(\theta_W)\theta'_A\theta_B + \Psi(\theta_W)\theta_A\theta'_B \\ =& -\beta_A\Psi'(\theta_W)\phi_A\theta_A\theta_B - \alpha_A\Psi(\theta_W)\theta_A\theta_B \\ & -\beta_B\Psi'(\theta_W)\phi_B\theta_A\theta_B - \alpha_B\Psi(\theta_W)\theta_A\theta_B \,. \end{aligned}$$

The first two terms on the right clearly generate A-buyers, the last two terms generate B-buyers. Hence,

$$I'_A = \beta_A\Psi'(\theta_W)\phi_A\theta_A\theta_B + \alpha_A\Psi(\theta_W)\theta_A\theta_B \,, \tag{13a}$$

$$I'_B = \beta_B\Psi'(\theta_W)\phi_B\theta_A\theta_B + \alpha_B\Psi(\theta_W)\theta_A\theta_B \,. \tag{13b}$$

Some Tests

Numerical tests gave perfect agreement of the behaviour predicted by this model with microscopic simulations performed using Gillespie's algorithm. We compare our models with the underlying stochastic marketing process on two types of networks: a Poisson network and a scale-free network. For a Poisson degree distribution

$$P_k = \frac{\lambda^k e^{-\lambda}}{k!},$$

and

$$\Psi(x) = e^{\lambda(x-1)}.$$

A scale-free network has a power-law degree distribution

$$P_k \propto k^{-r}$$

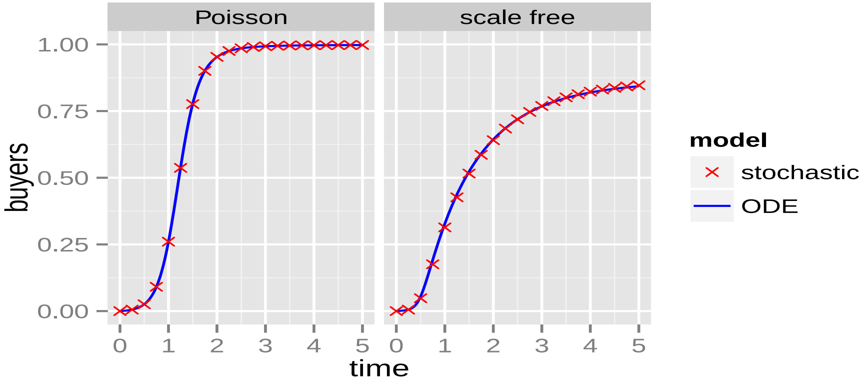

Figure : Comparison of the single-product model (11) with stochastic simulations (average of 500 runs) on a Poisson and a scale-free network. Network size for both networks $N = 20,000$. Poisson network: average degree $\langle k \rangle = 6$. Scale-free network: $P_k \propto k^r$ where $r = -2$, with maximum degree $k_{\max} = 66$. Word-of-mouth transmission rate $\beta = 1$, Advertisement rate $\alpha = 0.01$.

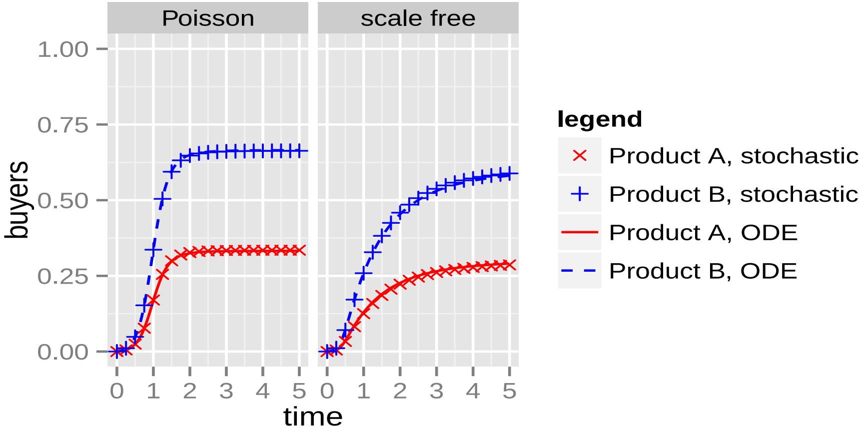

Figure : Comparison of the two-product model (12) with stochastic simulations on a Poisson and a scale-free network. The networks are the same as in Figure 1. The word-of-mouth transmission rates are $\beta_A = \beta_B = 1$, the advertisement rate $\alpha_A = 0.01$, $\alpha_B = 0.02$.

Some additional observations

If the transmission rates $\beta_A = \beta_B$ are equal, it is rather straightforward to predict the market shares companies A and B will achieve:

Theorem

Let $\beta_A = \beta_B = \beta$ *and* $\phi_A(0) = \phi_B(0) = 0$. *Then, for all* $t > 0$,

$$\frac{I_A}{I_B} = \frac{\alpha_A}{\alpha_B}. \tag{14}$$

This means that relative market share is proportional to relative advertising effort, regardless of the underlying network.

Compare with the previous figures!

A random network may have disconnected components.
For example: two degree one nodes may form an isolated pair.

A random network may have disconnected components.
For example: two degree one nodes may form an isolated pair.
More prevalent on a scale free network, which has many degree-one nodes. Thus, word of mouth may not be able to reach everyone on the social network.

A random network may have disconnected components.
For example: two degree one nodes may form an isolated pair. More prevalent on a scale free network, which has many degree-one nodes. Thus, word of mouth may not be able to reach everyone on the social network.

Take 1 product, no advertising: The expected final fraction of buyers can be computed. Specifically, as $\theta_W(t)$ is a positive and decreasing function, $\theta_W(\infty)$ exists. The fraction of susceptible nodes that never become buyers as time $t \to \infty$ is $S(\infty) = \Psi(\theta_W(\infty))$. To compute $\theta_W(\infty)$, we first simplify the equations: Dividing ϕ'_W by θ'_W yields

$$\frac{d\phi_W}{d\theta_W} = 1 - \frac{\Psi''(\theta_W)}{\Psi'(1)}\,.$$

Integrating on both sides, with $\phi_W(\theta_W(0)) \approx \phi_W(1) \approx 0$, leads to

$$\phi_W = \theta_W - \frac{\Psi'(\theta_W)}{\Psi'(1)}\,.$$

Substituting into the equations results in

$$\theta_W' = -\beta\theta_W + \beta\frac{\Psi'(\theta_W)}{\Psi'(1)}\,.$$

Thus, $\theta_W(\infty)$ is the interior root (i.e., strictly between 0 and 1) of

$$\theta_W = \frac{\Psi'(\theta_W)}{\Psi'(1)}\,. \tag{15}$$

If there is outside advertising as used in our model, it will reach everyone. Revisit the picture for one product and a scale-free network, which typically has many nodes in disconnected components; the product first diffuses quickly through word of mouth and advertisement in the connected components, then diffuses slowly in disconnected components through advertisement only.

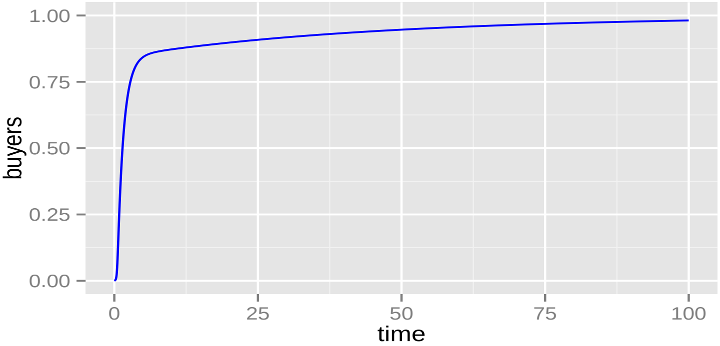

Figure : The fraction of buyers on a scale-free network converges to unity in two stages: it increases quickly and spreads through the large connected component, then approach unity exponentially through advertisement in disconnected components. The network and disease parameters are the same as before.

Recent work: SIS marketing models

... are much harder, because the independence assumptions one can make in the SIR case are no longer true.

Recent work: SIS marketing models

... are much harder, because the independence assumptions one can make in the SIR case are no longer true. Some terminology:

$$S(t) + A(t) + B(t) = N$$

Recent work: SIS marketing models

… are much harder, because the independence assumptions one can make in the SIR case are no longer true. Some terminology:

$$S(t) + A(t) + B(t) = N$$

$$S = \sum_{k=0}^{\infty} S_k$$

Recent work: SIS marketing models

... are much harder, because the independence assumptions one can make in the SIR case are no longer true. Some terminology:

$$S(t) + A(t) + B(t) = N$$

$$S = \sum_{k=0}^{\infty} S_k$$

$$S_k + A_k + B_k = N_k = NP_k.$$

(note that edge distribution appears here)
M_{SA}, M_{SB} etc. : number of **directed** edges with source A and target S.
$M_S = \sum kS_k = M_{SS} + M_{SA} + M_{SB}$. M_A, M_B similar.

We define

$$p_A := \frac{M_{SA}}{M_S},\ p_B := \frac{M_{SB}}{M_S}.$$

Probabilities that an edge with target S has source in A or B, respectively.

We define

$$p_A := \frac{M_{SA}}{M_S},\ p_B := \frac{M_{SB}}{M_S}.$$

Probabilities that an edge with target S has source in A or B, respectively.

$$\frac{d}{dt}S_k = -\beta_A p_A k S_k - \beta_B p_B k S_k - \alpha_A S_k - \alpha_B S_k + \gamma(A_k + B_k).$$

Linear in terms of the unknowns S_k, A_k, B_k, but the appearance of the p_A, p_B, defined in terms of the unknowns M_{SA} etc. This introduces nonlinearity into the system.

We define

$$p_A := \frac{M_{SA}}{M_S},\ p_B := \frac{M_{SB}}{M_S}.$$

Probabilities that an edge with target S has source in A or B, respectively.

$$\frac{d}{dt}S_k = -\beta_A p_A k S_k - \beta_B p_B k S_k - \alpha_A S_k - \alpha_B S_k + \gamma(A_k + B_k).$$

Linear in terms of the unknowns S_k, A_k, B_k, but the appearance of the p_A, p_B, defined in terms of the unknowns M_{SA} etc. This introduces nonlinearity into the system. It also assumes, **recklessly!** that the p_A, p_B can be used for any node with any degree k.

Similarly,

$$\begin{aligned} \frac{d}{dt}A_k &= \beta_A p_A k S_k + \alpha_A S_k - \gamma A_k, \\ \frac{d}{dt}B_k &= \beta_B p_B k S_k + \alpha_B S_k - \gamma B_k. \end{aligned}$$

Summing over k and using $\sum A_k = A$, $\sum B_k = B$, $p_A M_S = M_{SA}$, and $p_B M_S = M_{SB}$ give

$$\frac{d}{dt}A = \beta_A M_{SA} - \gamma A + \alpha_A S, \tag{16}$$

$$\frac{d}{dt}B = \beta_B M_{SB} - \gamma B + \alpha_B S. \tag{17}$$

Further

$$\frac{d}{dt}M_A = \sum_k k\frac{dA_k}{dt} = \beta_A p_A \sum_k k^2 S_k + \alpha_A M_S - \gamma M_A, \quad (18)$$

$$\frac{d}{dt}M_B = \sum_k k\frac{dB_k}{dt} = \beta_B p_B \sum_k k^2 S_k + \alpha_B M_S - \gamma M_B. \quad (19)$$

with $M_A(0) = M_B(0) = 0, \quad M_S(0) = \sum_k kS_k(0) = N\sum_k kP_k.$

Further

$$\frac{d}{dt}M_A = \sum_k k\frac{dA_k}{dt} = \beta_A p_A \sum_k k^2 S_k + \alpha_A M_S - \gamma M_A, \quad (18)$$

$$\frac{d}{dt}M_B = \sum_k k\frac{dB_k}{dt} = \beta_B p_B \sum_k k^2 S_k + \alpha_B M_S - \gamma M_B. \quad (19)$$

with $M_A(0) = M_B(0) = 0, \quad M_S(0) = \sum_k kS_k(0) = N\sum_k kP_k.$

We need equations for M_{SA}, M_{SB} etc.! This is somewhat challenging.

Average excess degree

$\sum_{k=0}^{\infty} kS_k$, isthe total number of edges starting from a susceptible.

Average excess degree

$\sum_{k=0}^{\infty} kS_k$, isthe total number of edges starting from a susceptible. Thus, if we follow an edge and arrive at a node of class S, the probability that this node has j edges is $jS_j / \sum_{k=0}^{\infty} kS_k$.
(it is proportionally more likely to reach nodes with more edges if one follows an edge). The *average excess degree* E of such nodes is then

$$E := \sum_{j=1}^{\infty} (j-1) \frac{jS_j}{\sum_{k=0}^{\infty} kS_k}. \tag{20}$$

This is needed in the sequel.

The dynamics of M_{SA} involves 10 terms on the right hand side. Here is the equation.

$$\frac{d}{dt}M_{SA} = -\underbrace{\beta_A M_{SA}}_{1} - \underbrace{\gamma M_{SA}}_{2} + \underbrace{\gamma M_{AA}}_{3} + \underbrace{\gamma M_{BA}}_{4} + \underbrace{\beta_A E p_A M_{SS}}_{5} - \underbrace{\beta_B E p_B M_{SA}}_{6} - \underbrace{\beta_A E p_A M_{SA}}_{7} + \underbrace{\alpha_A M_{SS}}_{8} - \underbrace{\alpha_A M_{SA}}_{9} - \underbrace{\alpha_B M_{SA}}_{10}. \quad (21)$$

Terms 5,6 and 7 contain an implicit assumption of **EDGE CHAOS!** Note that the p_A, p_B are taken as independent of k. The approximate validity of this assumption must depend on the type of graph.

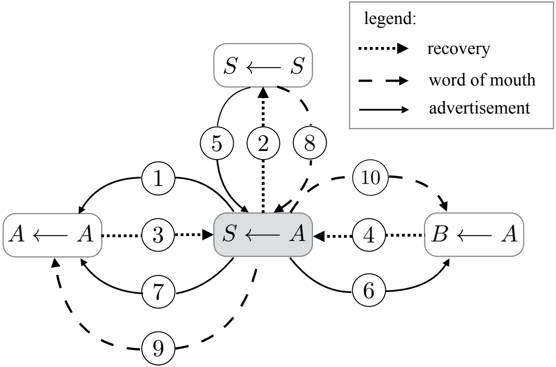

Figure : Gains and losses of the edges in the M_{SA} class, whose dynamics is given in (21). Circled number on each flow corresponds to the such numbered term in (21). Flows 2, 3, and 4 represent gains and losses due to recovery of a buyer on one end of an edge; flows 8, 9, and 10 represent the gains and losses due to the conversion of a susceptible to a buyer by advertisement; the other flows represent conversion by word-of-mouth.

The full model contains more equations for edges (for M_{SB}, M_{AB}, etc. In summary, the full model consists of the following equations:

$$M_S = \sum_{k=0}^{\infty} kS_k \, ; \;\; E = \frac{\sum_{j=1}^{\infty}(j-1)jS_j}{M_S} \, ; \tag{22a}$$

$$\frac{d}{dt}S_k = -\beta_A p_A k S_k - \beta_B p_B k S_k - \alpha_A S_k - \alpha_B S_k + \gamma(N_k - S_k) \, ; \tag{22b}$$

$$\frac{d}{dt}A = \beta_A M_{SA} - \gamma A + \alpha_A S \, ; \tag{22c}$$

$$\frac{d}{dt}B = \beta_B M_{SB} - \gamma B + \alpha_B S \, ; \tag{22d}$$

$$\frac{d}{dt}M_A = \beta_A p_A \sum_k k^2 S_k + \alpha_A M_S - \gamma M_A \, ; \tag{22e}$$

$$\frac{d}{dt}M_B = \beta_B p_B \sum_k k^2 S_k + \alpha_B M_S - \gamma M_B \, ; \tag{22f}$$

$$\frac{d}{dt}M_{SA} = -\beta_A M_{SA} - \gamma M_{SA} + \gamma(M_A - M_{SA}) + \beta_A E p_A M_{SS} - \beta_B E p_B M_{SA} - \beta_A E p_A M_{SA} + \alpha_A M_{SS} - \alpha_A M_{SA} - \alpha_B M_{SA}; \tag{23a}$$

$$\frac{d}{dt}M_{SB} = -\beta_B M_{SB} - \gamma M_{SB} + \gamma(M_B - M_{SB}) + \beta_B E p_B M_{SS} - \beta_A E p_A M_{SB} - \beta_B E p_B M_{SB} + \alpha_B M_{SS} - \alpha_B M_{SB} - \alpha_A M_{SB} \tag{23b}$$

$$\frac{d}{dt}M_{AB} = -2\gamma M_{AB} + \beta_A E p_A M_{SB} + \beta_B E p_B M_{SA} + \alpha_B M_{SA} + \alpha_A M_{SB}\,. \tag{23c}$$

A few simulations:

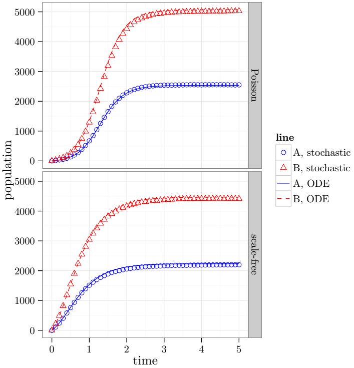

Figure : The comparison of the ensemble average of 100 stochastic simulations (symbols) and the solution of the ODE model (lines) on a Poisson network with average degree $\langle k \rangle = 5$, and a scale-free network with exponent $r = 3$ and minimum degree 2 (to avoid isolated pairs). Both networks have 10^4 nodes. The parameters are $\beta_A = \beta_B = 1$, $\gamma = 1$, $\alpha_A = 0.01$, $\alpha_B = 0.02$.

A warning:

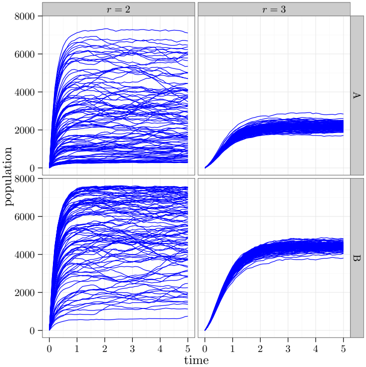

Figure : The comparison of 100 stochastic simulations on two scale-free networks. Both networks contain 10^4 nodes, and have minimum degree 2 to avoid isolated components. The exponents of the power-law degree distributions are $r = 2$ and $r = 3$ respectively. The marketing parameters are the same as in Figure 5.

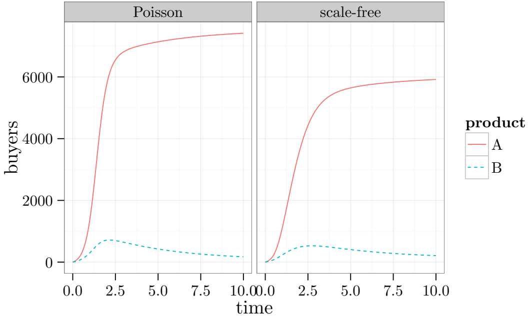

Figure : The time evolution of market shares of products A and B, with $\beta_A = 1$, $\beta_B = 0.7$, $\alpha_A = 0.02$, and $\alpha_B = 0.01$. The networks are the same as in Figure 7.

The homogeneous mixing limit is

$$S' = -\lambda_A \frac{AS}{N} - \lambda_B \frac{BS}{N} - \alpha_A S - \alpha_B S + \gamma A + \gamma B, \tag{24a}$$

$$A' = \lambda_A \frac{AS}{N} + \alpha_A S - \gamma A, \tag{24b}$$

$$B' = \lambda_B \frac{BS}{N} + \alpha_B S - \gamma B. \tag{24c}$$

where $\lambda_A = \beta_A(N-1), M_{SA} = SA$, etc.

We have additional analytical results regarding market share, and many more numerical tests. The greatest weakness is in the assumptions on the p_A, p_B.

On Edge Chaos

Our model is closed by using the edge chaos assumption and we checked the validity of this assumption.

On Edge Chaos

Our model is closed by using the edge chaos assumption and we checked the validity of this assumption.
Check whether the probability that a neighbour of a susceptible is susceptible (or a buyer) depends on the degree of the source node, i.e., whether $p_S = M_{SS}/M_S$, $p_A = M_{SA}/M_S$ and $p_B = M_{SB}/M_S$ will in microscopic simulations depend on the degree of the source node (and hence contradict the edge chaos assumption: the independence of k of these terms is a necessary, but not sufficient conditions for edge chaos).

On Edge Chaos

Our model is closed by using the edge chaos assumption and we checked the validity of this assumption.

Check whether the probability that a neighbour of a susceptible is susceptible (or a buyer) depends on the degree of the source node, i.e., whether $p_S = M_{SS}/M_S$, $p_A = M_{SA}/M_S$ and $p_B = M_{SB}/M_S$ will in microscopic simulations depend on the degree of the source node (and hence contradict the edge chaos assumption: the independence of k of these terms is a necessary, but not sufficient conditions for edge chaos).

To do so, we compute the average fraction of susceptible neighbours across all susceptible nodes with a degree k, namely, $p_S(k) = M_{SS}^{(k)}/(kS_k)$ where $M_{SS}^{(k)}$ is the total number of susceptible neighbours of degree-k susceptible nodes.

Time evolution of $p_S(k)$ in microscopic simulations is shown in Figure 8. As the diffusion process reaches equilibrium, $p_S(k)$ decreases with the degree k, i.e., there is an anti-correlation between the degree of a susceptible source node and the probability that its neighbour is susceptible.

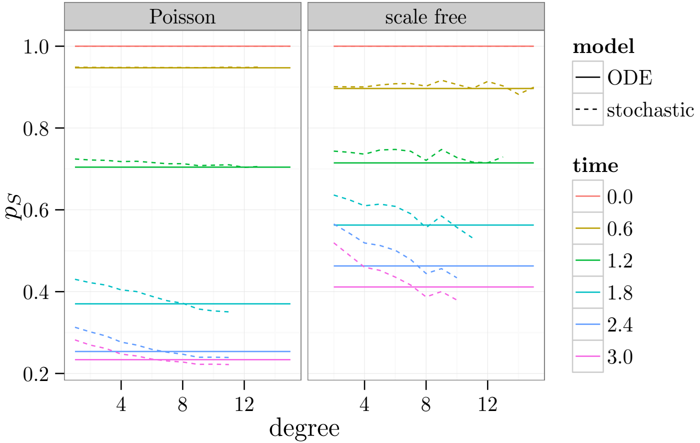

Figure : The time evolution of p_S as a function of the degree of the source node, for the simulations shown in Figure 5. Each curve represents $p_S(k)$ at a specific time.

Figure 9 shows $p_S(k)$ for the older SIR model. Here, edge chaos assumption is valid for all time. Thus, repeat purchase will gradually introduce correlation between the degree of a susceptible node and the probability that its neighbours are susceptible.

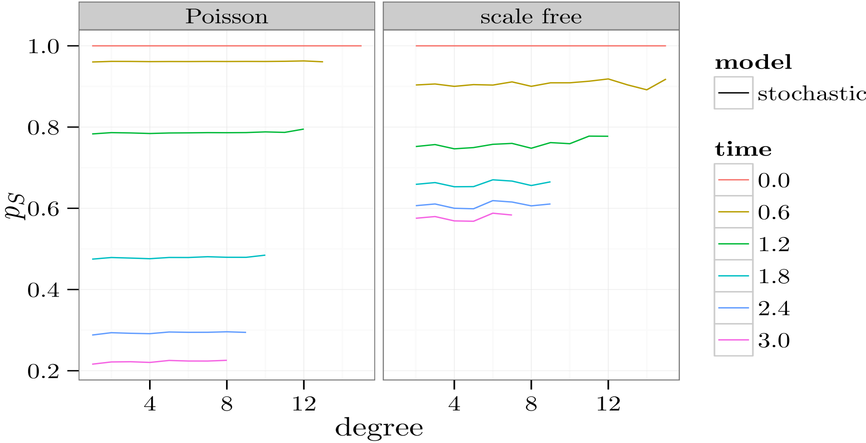

Figure : The time evolution of p_S as a function of the degree of the source node, for the SIR model. Each curve represents $p_S(k)$ at a specific time. Marketing parameters and the networks are identical to Figure 5, except that $\gamma = 0$.

Thank You Very Much