Mathematical model complex contagion

Aart van Bochove

May 2020

This mathematical model only treats edges within a single community. Note that time is discrete in this model, it goes in steps of one.

For a random edge, we define $\theta_{ij}(t)$ to be the probability that a certain node i on level L_i spreads the information to another node j on level L_j at time step t. In the designed model, L_j can only take the values of L_{i-1}, L_i and L_{i+1}. The probability $\phi_{m,L_i,j}(n_{ij}, t)$ for a certain node j to receive m pieces of information from nodes on level L_i at time step t behaves according to the binomial distribution and can be calculated by

$$\phi_{m,L_i,j}(n_{ij}, t) = C^m_{n_{ij}}[\theta_{ij}(t)]^m[1 - \theta_{ij}(t)]^{n_{ij}-m}. \tag{1}$$

Here n_{ij} is the number of edges between nodes on level L_i and node j. This number is equal for all of the nodes j on a certain level L_j. $C^m_{n_{ij}}$ is the binomial coefficient. Of course m can not be bigger then n_{ij}, every convinced neighbor spreads the information just once every time step.

The next step is to calculate the probability $\Phi_{M,L_i,j}(n_{ij}, t)$ of node j recieving M pieces of information from nodes on level L_i during and before time t. For $M = 0$ this would be equal to

$$\Phi_{0,L_i,j}(n_{ij}, t) = \prod_{t'=0}^{t} \phi_{0,L_i,j}(n_{ij}, t'). \tag{2}$$

For $M = 1$ this would be equal to

$$\begin{aligned} \Phi_{1,L_i,j}(n_{ij}, t) = {} & \phi_{1,L_i,j}(n_{ij}, 0) \prod_{t'=1}^{t} \phi_{0,L_i,j}(n_{ij}, t') \\ & + \phi_{0,L_i,j}(n_{ij}, 0)\phi_{1,L_i,j}(n_{ij}, 1) \prod_{t'=2}^{t} \phi_{0,L_i,j}(n_{ij}, t') + \cdots . \end{aligned} \tag{3}$$

If $\phi_{m,L_i,j}(n_{ij}, t')$ would be the same for each time step t', this would lead to the binomial distribution again. However, equation (1) shows that $\phi_{m,L_i,j}(n_{ij}, t')$ has a time dependence. Because of this, $\Phi_{M,L_i,j}(n_{ij}, t)$ is given by

$$\Phi_{M,L_i,j}(n_{ij}, t) = \sum_k s_k, \tag{4}$$

with

$$s_k \in \left\{ \prod_{t'=0}^{t} \phi_{m_{t'},L_i,j}(n_{ij}, t') \mid \sum_{t'=0}^{t} m_{t'} = M \right\}. \tag{5}$$

The total amount of information μ_{ij} from level L_i to node j is equal to M times the weight w_{ij} from L_i to j. That means that the probability to get μ_{ij} information is only defined for $\mu_{ij} = w_{ij} \cdot M$, with $M \in \mathbb{N}$. The total amount of information ν received by node j is then given by $\nu = \sum_{L_i=L_j-1}^{L_j+1} \mu_{ij} = \sum_{L_i=L_j-1}^{L_j+1} w_{ij} \cdot M$. Note that this sum starts at 0 if L_j is the lowest level, and ends at L_j if L_j is the highest level. The same applies to the sum and product in equation (7).

This total amount of information μ_{ij} allows us to talk about a probability $\Phi_{\mu_{ij}}(n_{ij}, t)$ instead of a probability $\Phi_{M,L_i,j}(n_{ij}, t)$.

The probability $\psi_{\nu,j}(t)$ for node j to have received less then ν information in total before and during time t can be calculated very similarly to equations (4) and (5), and is given by

$$\psi_{\nu,j}(t) = \sum_k S_k, \tag{6}$$

with

$$S_k \in \left\{ \prod_{L_i=L_j-1}^{L_j+1} \Phi_{\mu_{ij}}(n_{ij}, t) \mid \sum_{L_i=L_j-1}^{L_j+1} \mu_{ij} < \nu \right\}. \tag{7}$$

In the network which is created, there is also a loss of information from the nodes. This loss of information λ from node j is constant, except if there is no information in the node (then there is nothing to lose). The total amount of information lost in one time step by node j is given by Λ. This loss of information has to be combined with equation (6) to get the probability $\Psi_{\nu,j}(t)$ that a node j possesses ν or more information at time t:

$$\Psi_{\nu,j}(t) = \begin{cases} 1 - \psi_{(\nu-\Lambda),j}(t) & \text{if } \nu - \Lambda > 0, \\ 1 - \psi_{0,j}(t) & \text{if } \nu - \Lambda \leq 0. \end{cases} \tag{8}$$

The probability of losing information is given by the probability that a node possesses more then 0 information *before* time step t (so before or during time step $t-1$), which is $1 - \Psi_{0,j}(t-1)$. The amount of information Λ lost in a time step is approximated by the expectation value of λ, so

$$\Lambda = \lambda \cdot (1 - \Psi_{0,j}(t-1)). \tag{9}$$

The probability that a node j is convinced at a certain time step t is equal to $\Psi_{\alpha,j}(t)$, with α the information threshold for being convinced.

If level L_j consists of N_{L_j} nodes, the fraction of convinced individuals $F_{L_j}(t)$ at level L_j at time t is equal to

$$F_{L_j}(t) = \frac{1}{N_{L_j}} \sum_{j=1}^{N_{L_j}} \Psi_{\alpha,j}(t) = \Psi_{\alpha,j}(t), \tag{10}$$

since in this model all nodes on a level have the same properties. The total fraction of convinced individuals $f(t)$ at time t can then be calculated as

$$f(t) = \frac{1}{N} \sum_{L_j=1}^{N_L} N_{L_j} \cdot F_{L_j}(t), \tag{11}$$

with N the total amount of nodes in the network and N_L the total amount of levels in the network.

To finalize the calculations, $\theta_{ij}(t)$ needs to be analyzed. The probability $\theta_{ij}(t)$ that a node i from level L_i spreads the information to another node j on level L_j at time step t is equal to the probability that node i has been convinced somewhere during previous time steps times the probability that the convinced node i spreads the information to node j. The first probability is given by equation (8) and the second probability is equal to 2 divided by the number of neighbors k_j of node j. This means that $\theta_{ij}(t)$ actually only depends on the properties of node i, and is given by

$$\theta_{ij}(t) = \Psi_{\alpha,i}(t-1) \cdot \frac{2}{k_i}. \tag{12}$$

To be able to calculate $f(t)$, two initial values should be set:

$$\Psi_{\nu,i}(0) = \begin{cases} \rho_0(L_i) & \text{if } \nu \leq \alpha, \\ 0 & \text{if } \nu > \alpha, \end{cases} \tag{13}$$

$$\theta_{ij}(0) = \Psi_{\alpha,i}(0) \cdot \frac{2}{k_i}. \tag{14}$$

with $\rho_0(L_i)$ the density of seeded nodes per level.