Stochastic Information Flow in an Inhomogeneous Network

Bret Hanlon*, Clyde Martin†, Yishao Zhou‡

*Department of Statistical Science, Cornell University Ithica, New York
Email: bmh35@cornell.edu
†Department of Mathematics and Statistics, Texas Tech University
Lubbock, Texas
Email: clyde.f.martin@ttu.edu
‡Department of Mathematics, Stockholm University
Stockholm Sweden
Email yishao@math.su.se

Abstract—This paper develops a stochastic model of information flow in a network consisting of homogeneous groups of nodes. Communication between the groups is uniform but different than communication within a group. The model is modified to consider the spread of diseases and computer viruses.

Two recent papers, [6], [7], develop stochastic models of information flow in a network. The first paper, [6], considers a network with nearest neighbor connections. The information starts in one node and eventually spreads to all of the nodes in the network. The second paper, [7], is an extension of [6], where nodes can lose the information and can become blocked so that they cannot receive it. These models are used to analyze genetic diversity in the playa system, which is a communicating ecosystem.

This paper generalizes the model in [6] presenting a stochastic model of information spread through a network consisting of bundles, i.e. a group of nodes that communicate. Information passes from node to node. We say that a bundle has the information if at least one node in the bundle has the information. A bundle that has the information can spread the information to other bundles. Thus, there is both intra- and inter-bundle communication. Some fixed number of nodes starts with the information, which then spreads to communicating nodes. Generally there are two quantities of interest: the time until all of the bundles have received the information and the time until all of the nodes have received the information. Depending on the application, one might also be interested in the time until a specific bundle or node has received the information.

We consider two applications of the model: the spread of diseases and computer viruses. In the context of disease spread this model is a generalization of [5]. Paper [5] develops and simulates a model of disease spread in a single bundle of nodes. We extend this model to allow for interaction between separate but connected bundles. For example, in a multi-floor building, each floor is a bundle and each individual is a node. One then models how the disease spreads on a given floor (intra-bundle communication) and how the disease spreads between floors (inter-bundle communication). Another example is a different classrooms in a school.

For the spread of computer viruses, an institution’s network of computers is a bundle and each computer is a node. Once one computer on a given network is infected, the virus will spread to the other computers on that network. Also, certain computers will have connections outside of their network. Thus the virus can move to another institution’s network. The simulation results from this model suggest that the damage of viruses can be greatly reduced with increased communication between systems administrators of different networks.

I. GENERAL MODEL

This section describes the general stochastic model. Nodes take the value one or zero, and accordingly are called either one-nodes or zero-nodes. One-nodes correspond to nodes which posses the information. If a one-node communicates with a zero-node, the zero-node receives the information and becomes a one-node. The only possible transitions in the system are zero-nodes becoming one-nodes. Connected nodes communicate after an exponential amount of time. Nodes within a given bundle are connected in a uniform complete graph: every node communicates with all other nodes in its bundle at the same rate. Inter-bundle communication is more flexible as described below.

There are \( n \) total bundles and bundle \( i \) has size \( N_i \) (it contains \( N_i \) nodes), \( 1 \leq i \leq n \). Bundle \( i \) communicates with bundle \( j \) at rate \( \lambda_{ij} \), where \( \lambda_{ij} = \lambda_{ji} \). More precisely, each bundle \( i \) node communicates with each bundle \( j \) node after an independent exponential rate \( \lambda_{ij} \) amount of time. Notice that communication within bundle \( i \) occurs at rate \( \lambda_{ii} \). Because of the exponential rates of communication this system is a continuous time Markov chain (CTMC), denoted \( X(t) \). The state of the system is the \( n \)-tuple which gives the number of one-nodes in each bundle. Let \( X_i(t) \) be the number of one-nodes in bundle \( i \) at time \( t \), then \( X(t) = (X_1(t), X_2(t), ..., X_n(t)) \). Again, the only transitions in the system are zero-nodes becoming one-nodes. Therefore, \( X(t) \)
is described by the rates of increase in one-nodes in bundle $i$, which are given by:

$$q[x, (x_1, ..., x_i+1, ..., x_n)] = (N_i - x_i) \sum_{j=1}^{n} \lambda_{ij} x_j, \quad 1 \leq i \leq n.$$  

We comment on the derivation of the above rates. Let $x_j$ be the number of one-nodes in bundle $j$. Each of the $x_j$ one-nodes in bundle $j$ is communicating with each of the $N_i - x_i$ zero-nodes in bundle $i$ at rate $\lambda_{ij}$. Therefore the contribution from bundle $j$ nodes transmitting the information to bundle $i$ nodes is $\lambda_{ij} x_j (N_i - x_i)$.

We are interested in the following two first passage times:

$$T_1 = \min\{\tau | X_i(\tau) \geq 1, \forall i\},$$  

$$T_2 = \min\{\tau | X_i(\tau) = N_i, \forall i\}.$$  

$T_1$ is the time until the information has spread to all of the bundles, i.e. at least one node in each bundle has the information. $T_2$ is the time until every node in the network has received the information.

Finally, we discuss inter-bundle communication. The values of $\lambda_{ij}$, $i \neq j$, determine the connections between the bundles. For example, if $\lambda_{ij} > 0$ for $1 \leq i, j \leq n$, then the bundles are connected in a complete graph and all bundles can directly communicate. Three examples for inter-bundle communication are described below.

**Example 1. Uniform Complete Graph**

In this situation, all of the bundles are connected in a uniform complete graph where the communication rates between bundles are $\lambda_2$. Then

$$\lambda_{ij} = \begin{cases} 
\lambda_{ii}, & \text{if } i = j \\
\lambda_2, & \text{if } i \neq j.
\end{cases}$$

**Example 2. Rectangular Grid with Nearest Neighbor Connections**

The bundles are located in a rectangular grid where bundle $i$ has coordinates $(a_i, b_i)$. For example, bundle 1 is located in the bottom left corner of the grid and has coordinates $(1, 1)$. Bundles are only connected to their nearest neighbors in the grid. A given non-edge bundle is connected to four other bundles: the one above it, the one below it, the one to the right of it, and the one to the left of it. Each connected bundle communicates at rate $\lambda_2$. The communication rates are thus given by

$$\lambda_{ij} = \begin{cases} 
\lambda_{ii}, & \text{if } |a_i - a_j| = 0 \text{ and } |b_i - b_j| = 0 \\
\lambda_2, & \text{if } |a_i - a_j| = 1 \text{ and } |b_i - b_j| = 0 \\
\lambda_2, & \text{if } |a_i - a_j| = 0 \text{ and } |b_i - b_j| = 1 \\
0, & \text{otherwise}.
\end{cases}$$

Note that the condition $\{a_i - a_j = 0, b_i - b_j = 0\}$ is equivalent to $i = j$.

**Example 3. Completely Connected Rectangular Grid**

The bundles are arranged on a grid as in example 2; however, in this situation all of the bundles are connected. Let $d$ be a metric on the grid, where $d(i, j)$ represents the distance between bundles $i$ and $j$. The bundles communicate at a rate inversely proportional to their distance:

$$\lambda_{ij} = \begin{cases} 
\lambda_{ii}, & \text{if } i = j \\
\frac{\lambda_2}{d(i, j)}, & \text{if } i \neq j.
\end{cases}$$

The three above examples can be simplified to two parameter models by assuming that all of the intra-bundle communication rates are equal, i.e. $\lambda_{ii} = \lambda_1$, $1 \leq i \leq n$.

In the sequel, it is useful to define the communication matrix $\Lambda = (\lambda_{ij})$.

**REFERENCES**


