Does gravitational gossip weigh heavy on your local area network?

Anthony Billups *,Reynaldo Castro-Estrada †,Wilbert Fernández ‡
Tairi Roque-Urrea § Anthony Tongen ¶
Ariel Cintrón-Arias ‖ Carlos Castillo-Chávez **

August 13, 2004

Abstract

The World Wide Web has proven useful in disseminating information. However, it is extremely difficult to track its spread in such a complex network. One simplification is to model it as a small Local Area Network. Previous work has shown that a gossip-based multicast can be an effective tool for modeling the flow of information on a local area network. We examine the effect of different networks within this mathematical framework. We propose a general mathematical model that incorporates the susceptibility and infectivity of a given machine in a local area network. We also use the concept of proportionate mixing to model the spread of information among heterogeneous patches, where each patch consists of users and non-users of an information router.

*Northeastern University, Boston (billups.a@neu.edu).
†Arizona State University, Tempe (Reynaldo.Castro@asu.edu).
‡Arizona State University, Tempe (wfc151986@yahoo.com)
§Binghamton University (roque@math.binghamton.edu)
¶University of Arizona (tongen@math.arizona.edu)
‖Cornell University (ariel@cam.cornell.edu)
**Arizona State University (chavez@math.la.asu.edu)

1

Contents

1 Introduction

Before the World Wide Web, the spread of information through phones, word of mouth, and letters was difficult to quantify as well as to model. If we receive information that we find interesting, we can choose to pass the information on to our circle of friends and family. Once they receive it, they can also make a decision on whether or not to pass it on. This process continues over time and can be tracked for small populations. This tracking will look into how far the information has traveled over time. It will also be possible to see how many people are "infected" with the information over time. This process is very similar to that of tracking a virus in a population. Information spreads in ways that resemble the transmission dynamics of viruses. The literature on this latter topic includes [3], [2], [1], [4].

Now, the spread of information is a more complex process. With the help of the World Wide Web, an individual's circle of friends has greatly increased. Information can be spread across the world in a matter of seconds. However, before we can pursue the study of the global spread of information, we proceed to study it on smaller networks.

A very interesting idea to model the flow of information is to use a gossip protocol for a subgroup multicast [12]. An example of a multicast process is when someone sends out a mass e-mail to everyone in his or her address book. This framework may be helpful in the study of the spread of information in a small network. The gossip protocol represents the process that information is spread through a system. The protocol idea is to pay close attention to each node in the system and monitor the connections that node makes. This is done so you can have reliable and scalable results for the entire system based on each node. There is a system here in which a person generates some information (the gossip) and sends this gossip via the router to machines registered to receive messages in the local area network. This total group is broken up into patches based on different parameters which are explained in Section 2.1. The process by which the gossip is sent by a machine to multiple machines is known as the gossip-based multicast process. Other references about protocols include [8], [9], [10], [11], [13], [14].

In this paper, we review the Jenkins (et al.) [12] model and improve upon their work. In Section 2.1.1, we explain the results and add to the framework on gossip protocols for subgroup multicast. In Section 2.2, we introduce a new multigroup model. We first discuss this continuous-time model for m-patches based on subgroup multicast. We analyze the new model when only one patch is involved (with all users of the router) in Section 2.2.2 and 2.2.3. In Sections 2.2.4 and 2.2.5 we focus on a two-patch of users model. In Section 3 we summarize our results and in section 4 we outline future work.

2 Gravitational Gossip in Multi-patches

2.1 Discrete-Time Proportional Mixing Model

Jenkins (et al.)(2001) discussed a gossip-based protocol for a subgroup multicast [12]. The local area network consists of users that are broken up into subgroups. The subgroup is determined by two parameters: rating and timeout. The members set these parameters based on their personal preferences. The rating is between 0 and 1. A member with rating r wants to only receive a fraction r of the gossip being sent out over the whole system. The timeout is a fixed time in which the member is willing to receive gossip. For instance, a member who has set his timeout to 5 minutes and rating to .5 will only receive about 50 percent of the gossip overhead that is 5 minutes old or newer. Thus, members with extra time on their hands might set their parameters higher to receive more information. Likewise, a member who is busy might not have that much time to read gossip, so they set their parameters low in order to receive a small amount of the gossip overhead. Gossip overhead is the expected number of gossip messages initiated and received by a machine in a gossip round. We have included a figure based on the flow of gossip between subgroups in Figure 1.
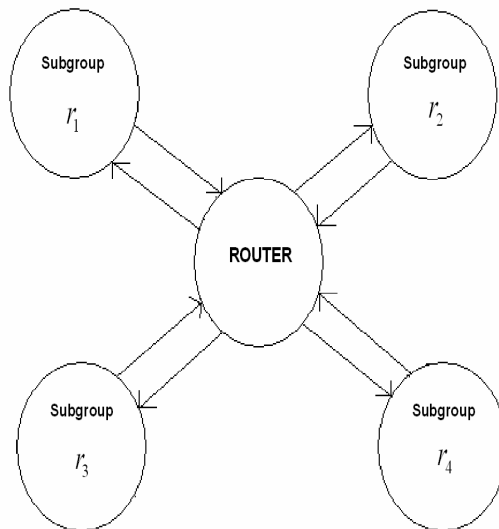

Figure 1: Flow of Gossip through the Router

Parameters	Description
$x_i(t)$	Proportion of susceptible machine at rating r_i at time t
$1 - x_i(t)$	The fraction of infective machines at rating r_i and time t
S_i	Susceptibility for each machine in subgroup r_i
I_i	Infectivity for each machine in subgroup r_i
N_i	Number of machines with rating r_i
$N_i S_i I_j N_j (1 - x_j(t))$	Total number of infections in neighborhood r_i due to contacts with subgroup r_j
$\sum_j N_i S_i I_j N_j (1 - x_j(t))$	Total number of infections in r_i due to contacts with all neighborhoods r_j including r_i

Table 1: Definitions of parameters for the model

With the understanding of the division of the network into subgroups, we can examine two parameters of each subgroup. Every machine in the subgroup has an *infectivity* (I) and *susceptibility* (S). These two parameters are based on the machines ability to send and receive gossip. Since every machine in the subgroup has the same rating and timeout, they will also have the same I and S. Let I_i and S_i denote the infectivity and susceptibility for subgroup i, which has rating r_i. Where r_i is the rating of subgroup i. The sequence of ratings is assumed to be ordered as follows:

$$0 \leq r_1 < r_2 < r_3 ... < r_{m-1} ... < r_m = 1 \tag{1}$$

In this model, we refer to a machine that has not received the gossip update as being *susceptible* and a machine that has received the update as *infected*. In epidemic modeling, this is known as an *SI* model ([3],[15]). First, let us assume there is a large number of machines in each subgroup. Second, let us assume the fraction of machines infected by each round of gossip is very close to the expected value. Given these assumptions, we create deterministic recurrence relations for the fraction of susceptible machines in each subgroup as a function of time (measured in rounds). In Table 1 we have included definitions for the parameters used in the equations.

$$x_i(t+1) = \text{P(machine } u \text{ at rating } r_i \text{ is susceptible at time } (t+1))$$

$$x_i(t+1) = \text{P}(u \text{ was susceptible at time } t)\text{P(no infected machines gossiped to } u \text{ at time } t+1)$$

$$x_i(t+1) = x_i(t)\text{P(each machine that gossiped to machines of rating } r_i \text{ at time } t+1 \text{ missed } u)$$

$$x_i(t+1) = x_i(t)\left(1 - \tfrac{1}{N_i}\right)^{(\text{number of gossips to rating } r_i \text{ at time } t+1)}$$

$$x_i(t+1) \approx x_i(t)\left(1 - \tfrac{1}{N_i}\right)^{N_i S_i \sum_{j=1}^{m} I_j N_j (1 - x_j(t))}$$

$$x_i(t+1) \approx x_i(t)\exp(-S_i\alpha(t)), \text{for } N_i \text{ large}$$

$$(2)$$

where

$$\alpha(t) = \sum_{j=1}^{m} I_j N_j (1 - x_j(t))$$

.

$\alpha(t)$ function is often, in the epidemiology literature, referred to as the *force of infection*. Notice that in the model (2), all subgroups have the same force of infection. We propose the following force of infection for subgroup i:

$$\alpha_i(t) = \sum_{j=1, j\neq i}^{m} I_j N_j (1 - x_j(t)), i = 1, 2, 3..., m$$

In words, the force of infection $\alpha_i(t)$ excludes *intra-neighborhood* infectivities, unlike, $\alpha(t)$, where all *intra-* and *inter*-neighborhood infectivities contribute to the spread of infection.

2.1.1   Numerical Simulations

Figure 2 displays the fraction of infected machines at each of the three quality ratings as a function of time. This plot was generated for one subgroup containing 100 machines at rating 1, 100 at rating 0.5, and 200 at rating 0.25, with a timeout value of 15. We used $\delta = 0.01$, and set each machine's infectivity proportional to its rating.
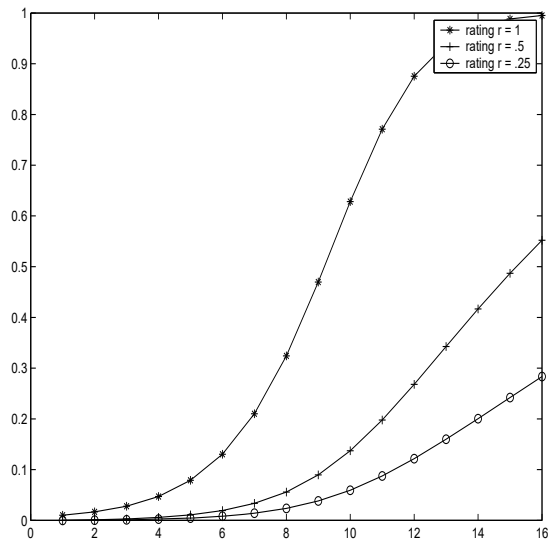
Figure 2: Plot of the fraction of infected machines vs. number of gossip rounds

In the numerical simulations, we also considered two other cases for the local area network. For both cases, we calculate the fraction of infections with different values for $\alpha(t)$ and $\alpha_i(t)$. $\alpha(t)$ is the force of infection of the local area network including the patch where the gossip originated. When a gossip is sent by a machine, it goes through the router and then it is filtered to all subgroups including the subgroup where the gossip originated. in $\alpha_i(t)$ is the force of infection of the local area network that excludes the patch where the gossip originated. In the latter example, when a gossip is sent by a machine it goes through the router but doesn't return to the original subgroup. A machine can not infect machines in their own neighborhood. For case 1 we assume the infection rate $I_m = 1$ for both α_1 and α_i . I_m is the highest infectivity in the local area network from subgroup m. In Figure 3 we can see that the fraction of infections with rating 1 can be reduced by 90 percent with $\alpha_i(t)$ getting less infective members in the local area network. This is due to the fact that we have a small number of subgroups. There is a big change if a machine can not spread gossip to machines in its own patch.
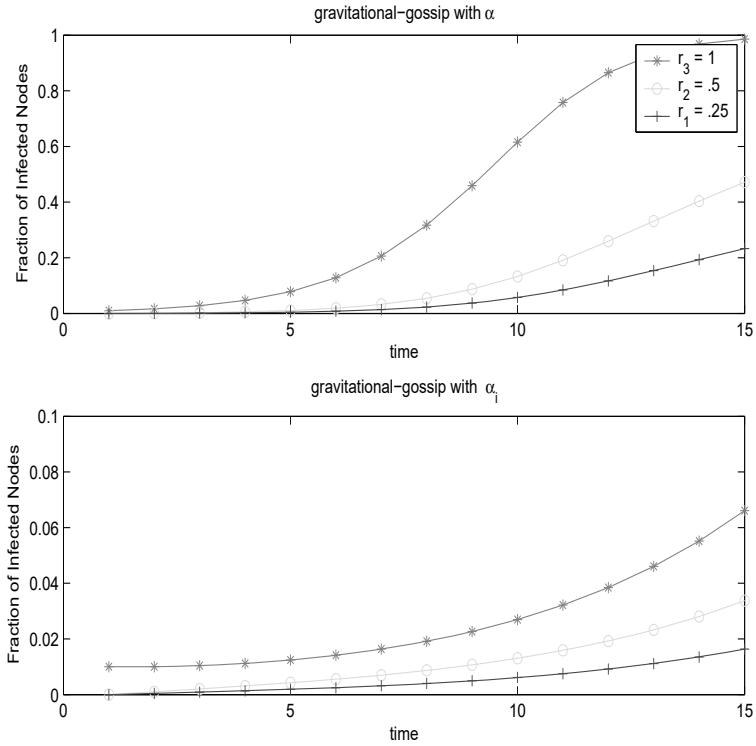
Figure 3: Plot of fraction of infected nodes with $\delta = 0.01$, $N_j = 100$ for $j \in \{1, 2, 3\}$

In the next simulation we calculate the infection rate in each patch for $\alpha(t)$ and $\alpha_i(t)$ assuming the same fraction of infected nodes in each patch. Figure 4 shows that it takes a higher infection rate when a gossip is not sent to the patch where the gossip originated. We calculate $I_j = r_j I_m$.
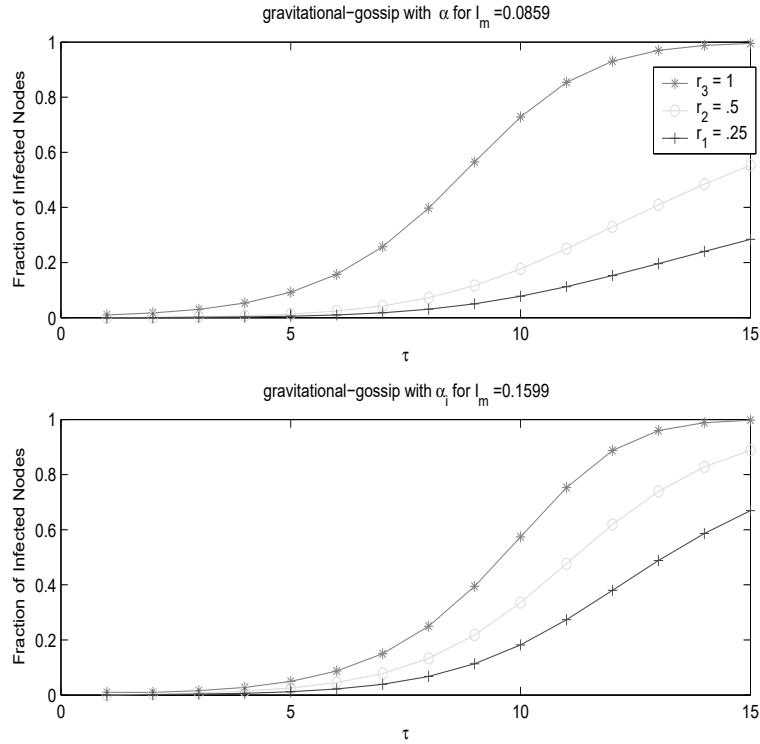
Figure 4: Shows the I_m needed to have the desired fraction of infected nodes for a given timeout

The numerical results in Figure 5 show that the infection rate I_m is a non-linear term that decay exponentially with increasing timeouts. This result indicates that for a given timeout the infection rate for $\alpha_i(t)$ needs to be stronger if it is to infect the same amount of members as in the case of $\alpha(t)$. Also we can see that for a large timeout, it is not necessary to have a high infection rate to infect most of the members in the Local Area Network.
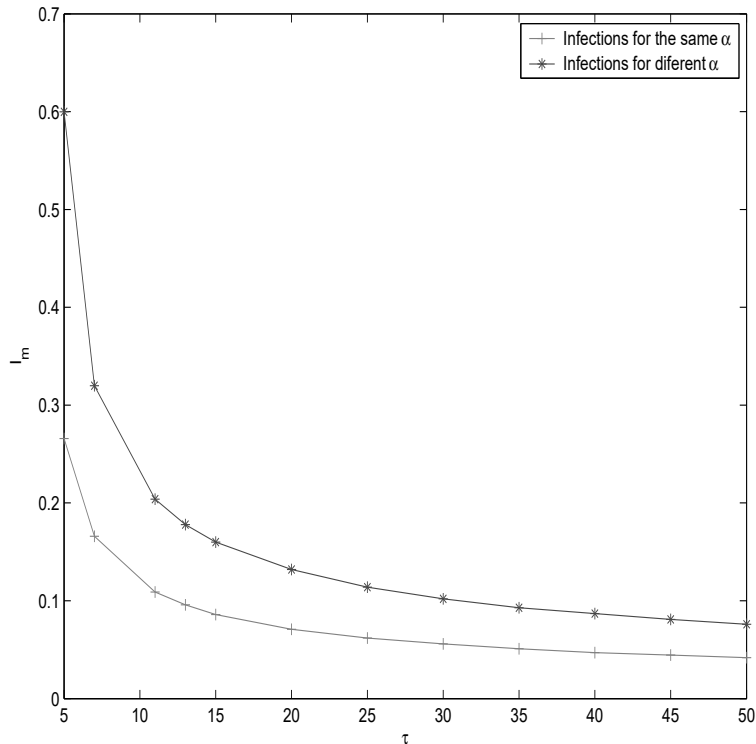
Figure 5: Plot of infection rate for a given timeout

2.2 Continuous-Time Proportional Mixing Model

Since the beginning of our history people have always communicated with other people. Now we use more communication ways such as phones, internet, tv, newspaper, magazine, etc. Today the computer is a very useful tool for communication. We are able to use the internet to communicate to other computers (i.e. to other people.) To access this structure of communication, computers need to be connected to some internet server provider (i.e MSN, AOL, Verizon, etc.) called the information router. Also computers can communicate each other just by being connected with each other without using the router. From this communication structure we can use the of proportional mixing from [6] Castillo-Chavez and Song, B. The idea of this model is having a N population of computers where each computer can communicate with each other in the same neighborhood being a user or non-user of the router. Computers can also communicate to other computers in different neighborhoods while accessing the router.

2.2.1   Framework: Model for m-Patches

We feel this new model is relevant because it makes sense if everyone does not have access to the router. One example of the model would be if it is expensive to access the router. We assume there are users (U) and non-users (NU) of the router in each of the m different neighborhoods. Divide each group of machines into m neighborhoods according to different rating r_i $1 \leq i \leq m$. Each neighborhood population is divided into subgroups of users and non-users. Assume users and non-users have contact within their own neighborhood. Also, users can also have contacts with other users of another neighborhood. A user from another neighborhood can send a message to the router and the router sends it to users from the m different neighborhoods. The newly infected users can leave the router and infect their own neighborhood, both users and non-users. The subgroups fall into one of two classes according to their epidemiological status. S_i and I_i denote the population of users and non-users in the neighborhood i which are susceptible and infectious, respectively; X_i and Y_i denote the corresponding "epidemiological" classes for users. The total population of the two groups is $Q_i = S_i + I_i$ and $T_i = X_i + Y_i$. The constants a_i and b_i denote the per-capita contact rates of users and non-users in neighborhood i. In addition, $\omega_i = \frac{\rho_i}{\sigma_i + \rho_i}$ and $\tau_i = \frac{\sigma_i}{\sigma_i + \rho_i}$, where ρ_i and σ_i denote the rates at which users get on and off the router.
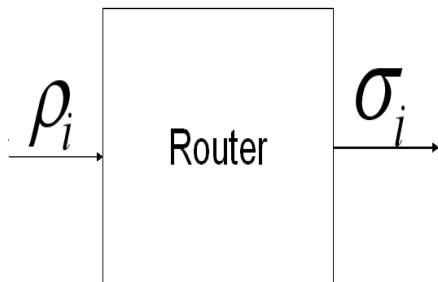

Figure 6: The rate ρ_i is the rate at which users get on the router. σ_i is the rate at which users get off the router.

In the discrete-time proportional mixing model, we stated that:
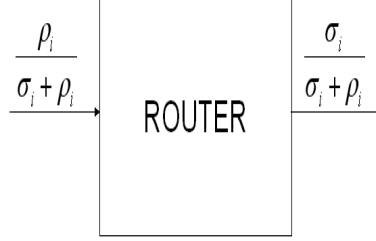
$$0 < r_1 < r_2 < r_3 ... < r_i ... < r_m \tag{3}$$

11

Figure 7: $\omega_i = \frac{\rho_i}{\sigma_i+\rho_i}$ which is the proportion of time users spent on the router. $\rho_i = \frac{\sigma_i}{\sigma_i+\rho_i}$ which is the proportion of time users spent off the router.

Where the rating of machines in subgroup m is equal to one. One is the highest rating a machine can have. Machines in subgroup m on average will receive the highest amount of gossip overhead.

Similarly, assume:

$$\omega_1 < \omega_2 < \omega_3 ... < \omega_i ... < \omega_m \tag{4}$$

Where ω_m corresponds to the rate at which machines in subgroup m access the router. This assumption states that machines in the m subgroup have the highest proportion of time on the router.

These assumptions are a main part of the motivation to connect the discrete and continuous-time proportional mixing models. Here are the proportional mixing parameters and description:

1. $P_{a_i a_i} = \tilde{P}_{a_i} = \frac{a_i Q_i}{a_i Q_i + b_i \tau_i T_i}$ is the mixing probability between NU from the same neighborhood i.

2. $P_{a_i b_i} = \tilde{P}_{b_i} = \frac{b_i \tau_i T_i}{a_i Q_i + b_i \tau_i T_i}$ is the mixing probability of NU and U from the same neighborhood i.

3. $P_{b_i a_i} = \bar{P}_{a_i} = (\frac{a_i Q_i}{a_i Q_i + b_i \tau_i T_i})\tau_i$ is the mixing probability of U and NU from the same neighborhood i.

4. $P_{b_i b_i} = \bar{P}_{b_i} = (\frac{b_i \tau_i T_i}{a_i Q_i + b_i \tau_i T_i}) \tau_i$ is the mixing probability between U from the same neighborhood i.

5. $P_{b_i b_j} = \bar{P}_{b_j^i} = (\frac{b_j \omega_j T_j}{\sum_{k=1}^{m} b_k \omega_k T_k}) \omega_i$ is the mixing probability between U from neighborhoods i and j.

6. $P_{a_i a_j} = 0$ means NU from neighborhoods i and j do not have contacts assuming $i \neq j$.

7. $P_{a_i b_j} = 0$ means NU from neighborhood i and U from neighborhood j have no contacts assuming $i \neq j$.

For each neighborhood, the following two "conditional probability" identities hold:

$$\tilde{P}_{a_i} + \tilde{P}_{b_i} = 1, i = 1, 2, ..., m, \qquad (5)$$

$$\bar{P}_{a_i} + \bar{P}_{b_i} + \sum_{j=1}^{m} \bar{P}_{b_j} = \tau_i + \omega_i = 1, i = 1, 2, ..., m. \qquad (6)$$

Parameters	Description
Λ_i	recruitment rate of U
A_i	recruitment rate of NU
μ	rate at which NU and U leave the system
σ_i	the rate at which a U leaves the router
ρ_i	the rate at which a U enters the router
a_i	average number of contacts of NU per unit time
b_i	average number of contacts of U per unit time
β_i	transmission rate per contact
$\frac{1}{\rho_i}$	the average time spent on the router
$\tau_i = \frac{\sigma_i}{\sigma_i + \rho_i}$	the proportion of time spent off the router (U)
$\omega_i = \frac{\rho_i}{\sigma_i + \rho_i}$	the proportion of time spent on the router (U)

Table 2: Definitions of parameters. i refers to the index of a neighborhood.

where the (force of) infection rate for NU is

$$B_i(t) = \beta_i a_i S_i \left[\tilde{P}_{a_i} \frac{I_i}{T_i \tau_i + Q_i} + \tilde{P}_{b_i} \frac{Y_i \tau_i}{T_i \tau_i + Q_i} \right] \tag{7}$$

and the (force of) infection rate for U is

$$V_i(t) = \beta_i b_i X_i \left[\bar{P}_{a_i} \frac{I_i}{T_i \tau_i + Q_i} + \bar{P}_{b_i} \frac{Y_i \tau_i}{T_i \tau_i + Q_i} + \sum_{j=1}^{m} \bar{P}_{b_j^i} \frac{Y_j \omega_j}{T_j \omega_j} \right] \tag{8}$$

2.2.2 Numerical Simulation for m-Patches

The simulation were carried out using the software Berkeley Madonna. The graphs represent the iteration between users and non-users of the router using two patches of subgroups, under the proportional mixing model. In figures 8-9 we can see that an endemic equilibrium exists. From both figures show that the dynamics of infection are logistic, that is they seem to read a carrying capacity in each patch and the level of total infection is different for each class.

Figure 8: Dynamics of Users and non-Users in the first patch.
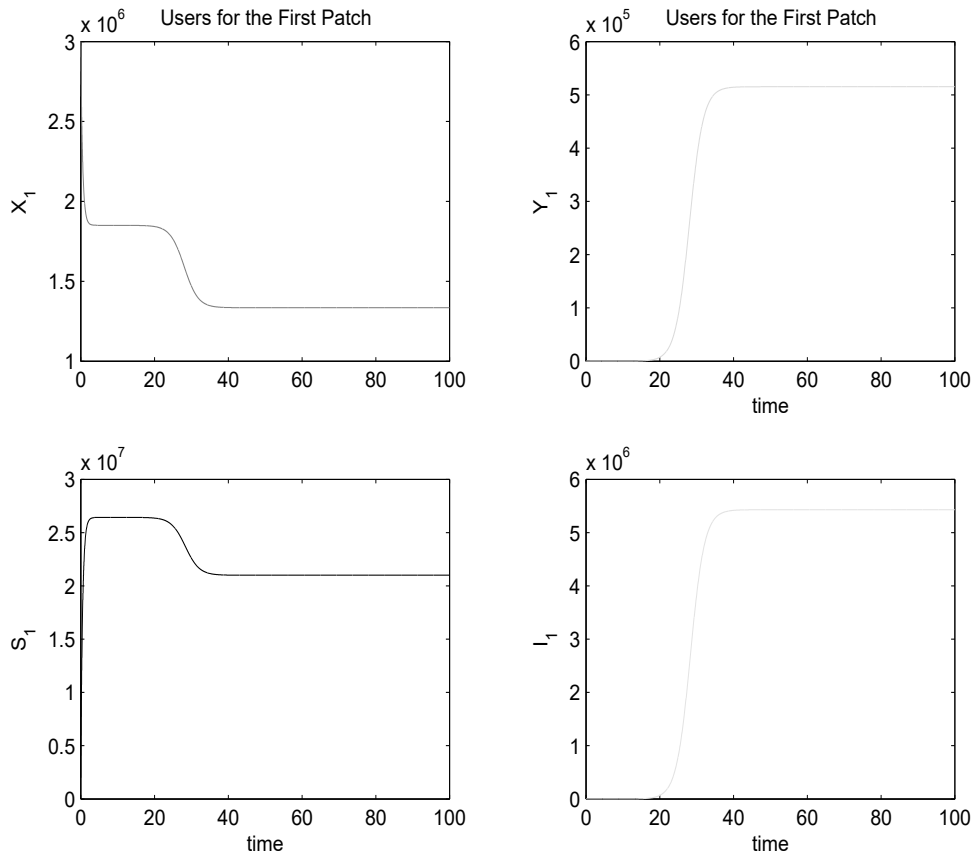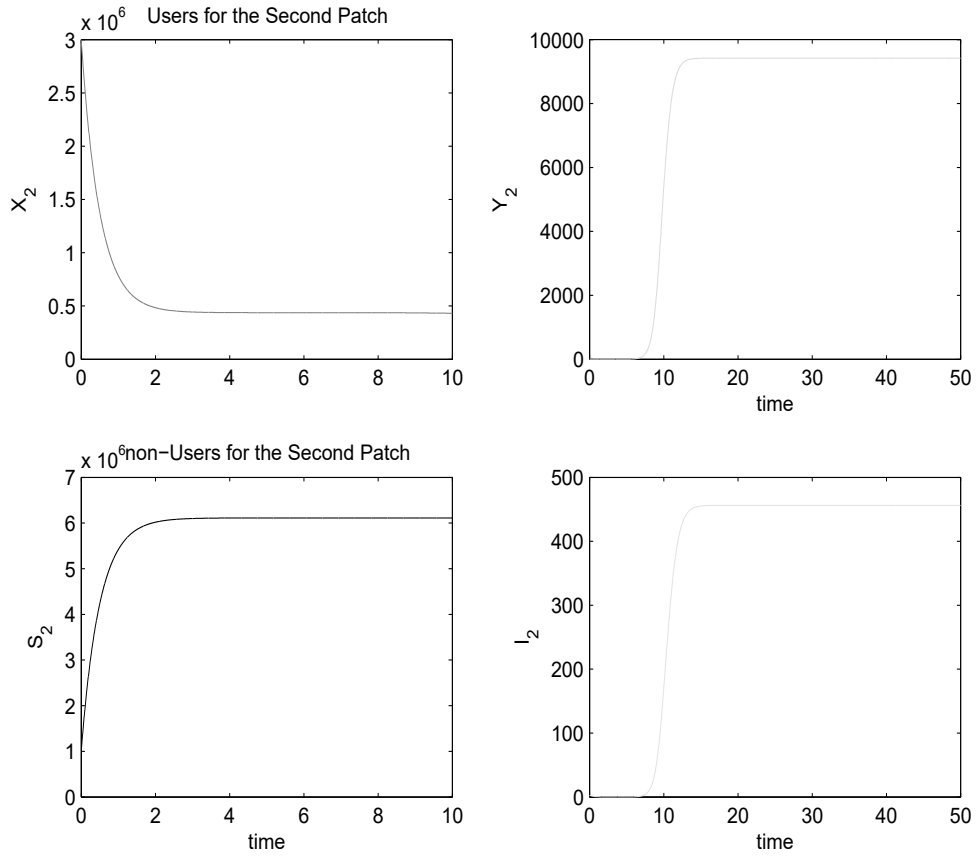
Figure 9: More dynamics of Users and non-Users in the second patch.

2.2.3   Model for One Single Patch of All Users

In this analysis, we use the $Susceptible - Infective(SI)$ model. Let $X(t)$ and $Y(t)$ denote the number of susceptible and infective machines at time t. Moreover, let $N(t) = X(t) + Y(t)$ denote the size of the total population at time t.
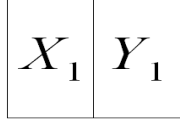
Figure 10: One patch which consist of all users.

$$\begin{aligned}
\frac{dX}{dt} &= \Lambda - \beta X \frac{Y}{N} - \mu X \\
\frac{dY}{dt} &= \beta X \frac{Y}{N} - \mu Y \\
\frac{dN}{dt} &= \Lambda - \mu N
\end{aligned}$$

(9)

It follows from $\frac{dN}{dt} = \Lambda - \mu N$ that $N \to \frac{\Lambda}{\mu}$ as $t \to \infty$, where β is the rate of infection of each computer.

Let $c = \frac{\Lambda}{\mu}$ and $y \approx \frac{Y}{c}$. Consider the *asymptotic limiting system* given by:

$$\frac{dy}{dt} = (\beta - \mu)y \left[1 - \frac{\beta}{\beta - \mu} y \right]$$

(10)

Equation (10) is the well known logistic equation [4]. The two equilibria are $y^* = 0$ (disease-free) and $y^* = \frac{\beta - \mu}{\beta}$ (endemic).

If $\beta - \mu > 0$, then $y^* = 0$ is unstable and $y^* = \frac{\beta - \mu}{\beta}$ is stable. Otherwise if $\beta - \mu < 0$, then $y^* = 0$ is stable and $y^* = \frac{\beta - \mu}{\beta}$ is unstable.

2.2.4 One Patch R_0 Analysis

Equation (9) has fixed points $(X^*, Y^*) = (\frac{\Lambda}{\mu}, 0)$ and $(X^*, Y^*) = (\frac{\mu T}{\beta}, \frac{\Lambda}{\mu} - \frac{\mu T}{\beta})$. Since T is constant, we have that the limit of $T \to \frac{\Lambda}{\mu}$ is constant which is considered the asymptotic limit system. The previous equations can be reduced to:

$$\dot{Z} = \beta(l - Z)\frac{Z}{l} - \mu Z$$

$$\dot{Z} = \beta(1 - \frac{Z}{l})Z - \mu Z$$

$$\dot{V} = V\left[(\beta - \mu) - \beta V\right] \tag{11}$$

$$\dot{V} = (\beta - \mu)V\left[1 - \frac{\beta V}{\beta \mu}\right]$$

$$\dot{V} = V[1 - dV]$$

V implies two fixed points, $V = 0$ and $V = \frac{\beta - \mu}{\beta}$. The first is the disease free and the second is the endemic equilibrium.

At $\beta - \mu$, there is a transcritical bifurcation where the stability of two fixed points inter changes. As you can see below we have an example of the bifurcation, where the horizontal line is the zero fixed point.

R_o is equal to $\frac{\beta}{\mu}$ where β is the average number of contacts leading to infection and $\frac{1}{\mu}$ is the average time spent on the router. Shows that when R_o is greater than one, the fixed point is the endemic equilibrium. Figure 2.2.4, R_o less than one corresponds to the disease free fixed point.
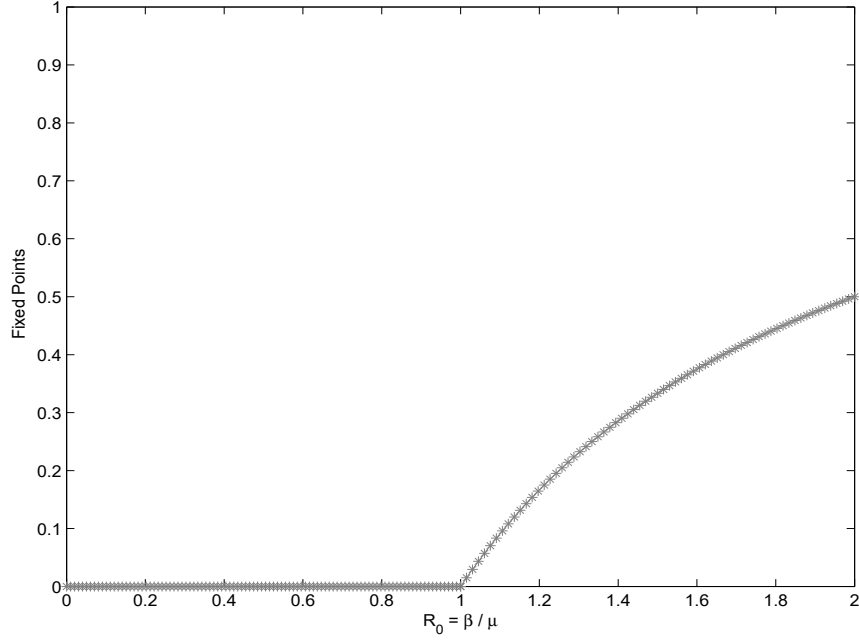
Figure 11: Bifurcation Diagram.

2.2.5 Model for Two-Patches of All Users

This model is very similar to the discrete-time model. The only difference between the two is that now users can interact with users in their own subgroup with out accessing the router. We assume constant population for this model to simplify calculations.

$$\dot{X}_1 = \mu \bar{T}_1 - V_1(t) - \mu X_1 \tag{12}$$

$$\dot{Y}_1 = V_1(t) - \mu Y_1 \equiv F_1(X_1, Y_1, X_2, Y_2) \tag{13}$$

$$\dot{X}_2 = \mu \bar{T}_2 - V_2(t) - \mu X_2 \tag{14}$$

$$\dot{Y}_2 = V_2(t) - \mu Y_2 \equiv F_2(X_1, Y_1, X_2, Y_2) \tag{15}$$

Assuming constant population simplifies the system of equations. We can reduce the system of four equations to two, by setting:

$$X_1 = \bar{T}_1 - Y_1 \tag{16}$$

19

Figure 12: Two patches which consist of all users.

$$X_2 = \bar{T}_2 - Y_2 \tag{17}$$
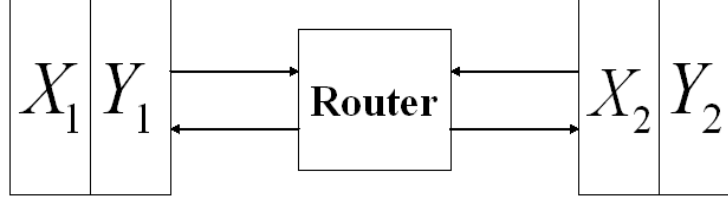
We substitute equations (16) and (17) into our new $V_i(t)$

$$V_1(t) = \beta_1 b_1 (\bar{T}_1 - Y_1) \left[\frac{Y_1 \tau_1}{\bar{T}_1} + \frac{(b_1 \omega_1^2 Y_1) + (b_2 \omega_2 Y_2 \omega_1)}{N} \right] \equiv F_1 \tag{18}$$

$$V_2(t) = \beta_2 b_2 (\bar{T}_2 - Y_2) \left[\frac{Y_2 \tau_2}{\bar{T}_2} + \frac{(b_2 \omega_2^2 Y_2) + (b_1 \omega_1 Y_1 \omega_2)}{N} \right] \equiv F_2 \tag{19}$$

where $N = b_1 \omega_1 \bar{T}_1 + b_2 \omega_2 \bar{T}_2$

We find the partial derivatives of equations (14) and (16) for two subgroups evaluated at the Disease Free Equilibrium (DFE).
(X_1, X_2, Y_1, Y_2)
$DFE \rightarrow (\bar{T}_1, \bar{T}_2, 0, 0)$

$$\frac{dF_1}{dY_1} = -\mu + \beta_1 b_1 \bar{T}_1 \left[\frac{\tau_1}{\bar{T}_1} + \frac{(b_1 \omega_1^2)}{N} \right] = -\mu + K_{11} \tag{20}$$

$$\frac{dF_1}{dY_2} = \beta_1 b_1 \bar{T}_1 \frac{b_2 \omega_2 \omega_1}{N} = k_{12} \tag{21}$$

$$\frac{dF_2}{dY_1} = \beta_2 b_2 \bar{T}_2 \frac{b_1 \omega_1 \omega_2}{N} = k_{21} \tag{22}$$

20

$$\frac{dF_2}{dY_2} = -\mu + \beta_2 b_2 \bar{T}_2 \left[\frac{\tau_2}{\bar{T}_2} + \frac{b_2 \omega_2{}^2}{N} \right] = -\mu + k_{22} \tag{23}$$

K_{11} and K_{22} stand for the local interactions that go on with-in the subgroup./

K_{12} and K_{21} stand for the global interactions that go on when users access the router.

We use the *Next Generation Approach* (see Appendix) to solve for R_0 by the following Jacobian:

$$J = \left(\begin{array}{cc} \Delta F_1 & \Delta F_2 \end{array} \right) \tag{24}$$

$$J = \left(\begin{array}{cc} k_{11} - \mu & k_{21} \\ k_{12} & k_{22} - \mu \end{array} \right) \tag{25}$$

J=M-D

$$J = \left(\begin{array}{cc} k_{11} & k_{21} \\ k_{12} & k_{22} \end{array} \right) - \left(\begin{array}{cc} \mu & 0 \\ 0 & \mu \end{array} \right) \tag{26}$$

where

$$M = \left(\begin{array}{cc} k_{11} & k_{21} \\ k_{12} & k_{22} \end{array} \right) \tag{27}$$

and

$$D = \left(\begin{array}{cc} \mu & 0 \\ 0 & \mu \end{array} \right) \tag{28}$$

$$D^{-1} = \left(\begin{array}{cc} \frac{1}{\mu} & 0 \\ 0 & \frac{1}{\mu} \end{array} \right) \tag{29}$$

$M2 = MD^-1$

$$M2 = \left(\begin{array}{cc} k_{11} & k_{21} \\ k_{12} & k_{22} \end{array} \right) \left(\begin{array}{cc} \frac{1}{\mu} & 0 \\ 0 & \frac{1}{\mu} \end{array} \right) \tag{30}$$

$$M2 = \left(\begin{array}{cc} \frac{k_{11}}{\mu} & \frac{k_{21}}{\mu} \\ \frac{k_{12}}{\mu} & \frac{k_{22}}{\mu} \end{array} \right) \tag{31}$$

$$\sigma(MD^{-1}) = \text{eigenvalues of } MD^{-1}$$

$$\sigma(MD^{-1}) = \{\lambda_1, \lambda_2\} \tag{32}$$

$$\lambda_1 = \frac{trace - \sqrt{trace^2 - 4det}}{2} \tag{33}$$

$$\lambda_2 = \frac{trace + \sqrt{trace^2 - 4det}}{2} \tag{34}$$

where trace$= \frac{k_{11}+k_{22}}{\mu}$

and

det$= \frac{k_{11}k_{22}}{\mu^2} - \frac{k_{12}k_{21}}{\mu^2}$

We want max $R_0 \equiv max\{|\lambda_1||\lambda_2|\}$

After some algebra we find λ_2 to be the max so:
$R_0 \equiv \frac{k_{11}+k_{22}+\sqrt{(k_{11}+k_{22})^2-4(k_{11}k_{22}-k_{12}k_{21})}}{2\mu}$

2.2.6 Two Patch R_0 Analysis and Numerical Results

In order to analyze R_0, we use the equivalent equation for (R_1).
$R_1 = \frac{\mu(k_{11}+k_{22})-k_{11}k_{22}}{\mu^2-k_{21}k_{12}}$

We found if $R_1 < 1$ then $R_0 < 1$

Rearranging R_1 implies:
$R_1 = \left(\frac{\tilde{k_{11}}\tilde{k_{22}}}{\tilde{k_{12}}\tilde{k_{21}}}\right)\left(\frac{\frac{1}{\tilde{k_{11}}\tilde{k_{22}}}-1}{\frac{1}{\tilde{k_{12}}\tilde{k_{21}}}-1}\right)\left(\frac{1}{\tilde{k_{11}}\tilde{k_{22}}}\right) < 1,$
Where $\tilde{k_y} = \frac{\tilde{k_y}}{\mu}$.

We interpret R_1 it can be said that $\tilde{k_{11}}$ and $\tilde{k_{22}}$ are the driving force for R_1 to be less than one.

Therefore, the interactions with users in the same neighborhood determine whether or not our R_0 will be less than one. If $R_0 < 1$ than there is no epidemic and if $R_0 > 1$ than an epidemic occurs where a vast majority of the population will be infected with

the gossip.

Numerical Results

In the Figure 13 we see that there is an endemic equilibrium for a system for all users. There is a rapid outbreak of the gossip from a small initial source, then the gossip appears to be of epidemic proportions (for $R_0 > 1$). In this graph, the susceptibles start having interaction with infected users approximate at 10 minutes and become infected users of the router.
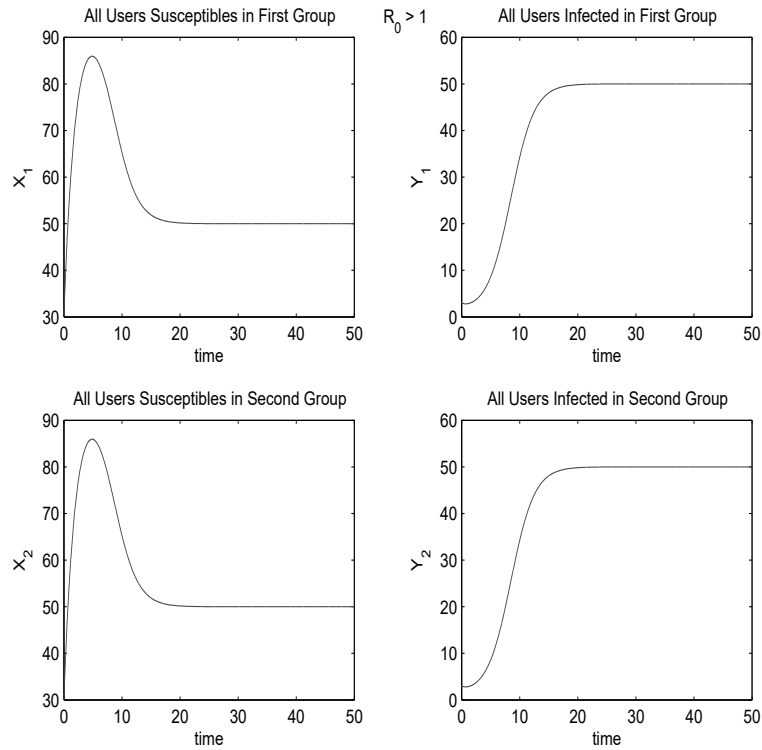

Figure 13: System of all users for 2 patches when $R_0 > 1$

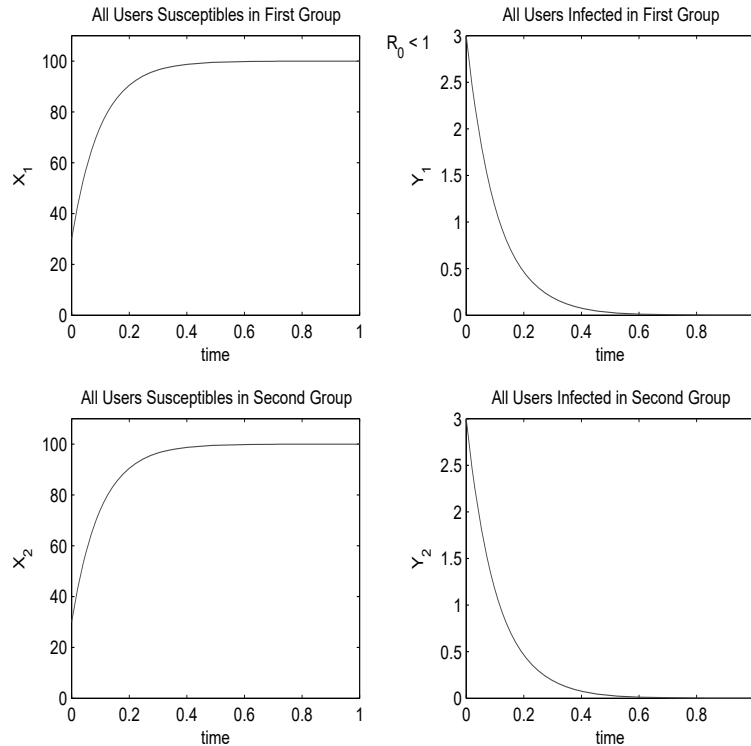Figure 14 displays disease free behavior which corresponds to $R_0 < 1$.

Figure 14: System of all users for 2 patches when $R_0 < 1$

3 Concluding Remarks

Does gravitational gossip weigh heavy on your local area network ? We have presented a number of models to answer this question. Based on what we have shown in the previous sections the answer is yes. The Local Area Network plays a major role on how a machine receives and sends information through the router. Our results show that if some gossip is started in the Local Area Network, then machines may get infected in a short period of time (depending on the network they belong to). Different networks produce very different outcomes.

A network that consists of only users seems to be a very effective way to spread information. We found that when users interact with users inside their subgroup they spread the gossip faster. make more contacts with users in their own group. These assumptions are approximately correct for a small number of subgroups in the Local Area Network.

In a network where there is a large number of subgroups, local interactions do

24

not play a major role. In fact, comparison between a large Local Area Network that does not allow local interactions and large Local Area Network that does show "no" difference. The number of infected machines over time are just about equal.

If you want to model the flow of information received by your machine. The details of the local area network to which you belong are extremely important. You have to play close attention how information spreads between machines.

4 Future Work

It would be a good idea to model a system where each machine can be a part of multiple subgroups, which is a very realistic assumption. This is specially true for laptops. Wireless technology allows a single user to access multiple networks. We think this will be a great future direction for our paper.

We are very interested in the spread of computer viruses on networks. We plan to model how computer viruses spread and attempt to quantify an ideal network. If we show what networks can be the most susceptible, we can suggest plans to reduce the chance of having an infected network.

5 Acknowledgements

6 Appendix

A   Next Generation Operator

R_0 is often found through the study and computation of the eigenvalues of the Jacobian at the disease- or infectious-free equilibrium. Diekmann et al. 1990 follow a different approach: *the next generation approach*. They define R_0 as the spectral radius of the "next generation operator". The details of this approach are outlined

in the rest of this section. First, we consider the case where heterogeneity is discrete, that is, the case where heterogeneity is defined using groups defined by fixed characteristics, that is, for epidemiology models that can be written in the form:

$$\frac{dX}{dt} = f(X, Y, Z), \tag{35}$$

$$\frac{dY}{dt} = g(X, Y, Z), \tag{36}$$

$$\frac{dZ}{dt} = h(X, Y, Z), \tag{37}$$

where $X \in \mathbb{R}^r$, $Y \in \mathbb{R}^s$, $X \in \mathbb{R}^n$, r, s, ≥ 0, and h(X,0,0)=0. The components of X denote the number of susceptibles, recovered, and other classes of non-infected individuals. The components of Y represent the number of infected individuals who do not transmit the disease (various latent or non-infectious stages). The components of Z represent the number of infected individuals capable of transmitting the disease (e.g., infectious and non-quarentined individuals).

Let $U_0 = (X^*, 0, 0) \in \mathbb{R}^{r+s+n}$ denote the disease-free equilibrium, that is,

$$f(X^*, 0, 0) = g(X^*, 0, 0) = h(X^*, 0, 0) = 0. \tag{38}$$

Assume that the equation $g(X^*, Y, Z) = 0$ implicitly determines a function $Y = \tilde{g}(X^*, Z)$. Let $A = D_z h(X^*, \tilde{g}(X^*, 0), 0)$ and further assume that A can be written in the form $A = M - D$, with M ≥ 0 (that is, $m_{ij} \geq 0$) and $D > 0$, a diagonal matrix.

The spectral bound of matrix B is denoted by $m(B) = sup\{\Re\lambda : \lambda \in \sigma(B)\}$, where $\Re\lambda$ means the real part of λ, while $\rho(B) = \lim_{n\to\infty} \|B^n\|^{1/n}$ denotes the spectral radius of B. The proof of the following theorem involving matrix A found in Diekmann et al., (1990):
Either

$$m(A) < 0 \iff \rho(MD^-1) < 1 \tag{39}$$

or

$$m(A) > 0 \iff \rho(MD^-1) > 1 \tag{40}$$

The *basic reproductive number* is defined as the spectral radius (dominant eigenvalue) of the matrix MD^1, that is,

$$R_0 = \rho(MD^-1). \tag{41}$$

26

B   MATLAB Code

```
****************************************************************************
MTBI 2004, Project: The Spread of Computer Information on Networks.
Authors:  Reynaldo Castro; Anthony Billups ; Wilbert Fernandez
****************************************************************************

******************** Input by the user ******************************
***************** Computer Virus program **************************
disp('Note: If you do not want to enter each rating the program');
disp('      will generate equidistant rating for you. '); disp(' ');
rb    input('Do you want to enter each rating ? [Y=1,N=0]: '); if
rbr input('Enter the ratings in accending between 0 and 1
(e.g.[r1;r2;..;rf] ): '); else ri = input('Enter rating interval
between 0 and 1 (e.g [r1, r2] ), r : '); rs = input('Enter the
number of partitions you want to have: '); r = linspace
(ri(1),ri(2),rs); end delta = input('Enter the value of \delta : ');
N = input('Enter the number of machines corresponding to each
subgroup (e.g. [N1; N2; ..; Nf] ): '); tOut = input('Enter the
TimeOuts (e.g.[tOut(1); ..; tOut(f)]: '); disp(' ************* SEE
THE PLOTS *********************************' );
%************* Inital Values *********************************************
%C = [ strcat('*','-','b'); strcat('+','-','g'); strcat('o','-','r')];
rn = length(r); S=zeros(rn,1); I=zeros(rn,1);
Im=zeros(length(tOut),1); Im2=zeros(length(tOut),1); S(end) = 1;
I(end)= 1; Im(1) = 0.6; Im2(1) = 0.6; initIm=Im(1); initIm2=Im(1);
S(1:end-1) = log2(1-r(1:end-1))/log2(delta); S0 = S; I(1:end-1) =
r(1:end-1)*I(end);
%****************************************************************************
for i = 1: length(tOut)
    loop = 1;
    loop2 = 1;
    Im(i) = initIm;
    inc=.001;
    while(loop)
        S=S0*Im(i);
        x=[];
        x(1:rn-1,1) = 1.0;
        x(rn,1) = 0.99;
        %*****************Calculate The Fraction of Infections*****************
```

```matlab
        [x1,y] = groupAlpha7(x,delta,I,N,S,r,rn,tOut(i),loop,Im(i),inc);
        loop = y(1);
        Im(i) = y(2);
        inc=y(3);
    end
%    temp1 = Im(i);

    Im2(i) = initIm2;
    inc=.001;
    while(loop2)
        S2=S0*Im2(i);
        x2=[];
        x2(1:rn-1,1) = 1.0;
        x2(rn,1) = 0.99;
        %******************Calculate The Fraction of Infections******************
        [x2,y2] = group2Alpha7(x2,delta,I,N,S2,r,rn,tOut(i),loop2,Im2(i),inc);
        loop2 = y2(1);
        Im2(i) = y2(2);
    end
%    temp2 = Im2(i);
%****************   Plotting the Results %*****************************
    figure(i);
    subplot(2,2,1); plot(1-x1(3,:),'-*r');
    hold on
    plot(1-x1(2,:),'-*g');
    hold on
    plot(1-x1(1,:),'-*b');

    hold on
%       for j = 1:rn
%           plot(1:max(size(x1)),r(j),strcat(C(j,1),C(j,2),C(j,3)));
%       end
    axis([0 tOut(i) 0 1])
    ylabel('Fraction of Infected Nodes');
    xlabel('time');
    title(strcat('gravitational-gossip with  \alpha for I_m = ',  num2str(Im(i))));
    hold off
    subplot(2,2,2);  plot(x1');
    axis([0 tOut(i) 0 1])
    ylabel('Fraction of Susceptibles Nodes');
```

```
        xlabel('time');
        title(strcat('gravitational-gossip with \alpha for I_m = ',  num2str(Im(i))));
        hold off
        %----------------------------------------------------------------
        subplot(2,2,3); plot(1-x2');
        axis([0 tOut(i) 0 1])
        hold on
%        for j = 1:rn
%             plot(1:max(size(x2)),r(j),strcat(C(j,1),C(j,2),C(j,3)));
%        end
        ylabel('Fraction of Infected Nodes');
        xlabel('time');
        title(strcat('gravitational-gossip with \alpha_i for I_m =  ',  num2str(Im2(i))
        hold off
        subplot(2,2,4);  plot(x2');
axis([0 tOut(i) 0 1]) ylabel('Fraction of Susceptibles Nodes');
xlabel('time'); title(strcat('gravitational-gossip with  \alpha_i
for I_m = ', num2str(Im2(i))));
 hold off
end figure(i+1) plot(tOut, Im, '+-r') hold on plot(tOut, Im2,
'*-b'); hold off legend('Infections for the same \alpha',
'Infections for diferent \alpha'); xlabel('\tau'); ylabel('I_m');


function [g,y] = groupAlpha7(x,delta,I,N,S,r,rn,tOut,loop,Im,inc)
% This function Calculates the Fraction of susceptibles, with the total
% of infections including the group where virus was originate.
j=1; y(1) = loop; y(2) = Im; while(x(end,j) > delta)
  alpha=y(2)*dot(r.*N,1-x(:,j));
     x(:,j+1)=x(:,j).*exp(-S.*alpha);
     j=j+1;
end
%********************************
j if j==tOut
     x(end,tOut)
     if x(end,tOut) < delta
          y(1)=0;
     end
       y(2) = y(2) - inc/10
   elseif j>tOut
```

```
        y(2) = y(2) +.1
        inc=inc/10
   else
y(2) = y(2) - inc end
%*******************************
g = x; y(3)=inc;
%******* End of the Function ******

function [g2,y2]=group2Alpha7(x2,delta,I,N,S,r,rn,tOut,loop2,Im2,inc
)
% This function Calculates the Fraction of susceptibles, with the total
% of infections excluding the group where virus was originate.
j=1; y2(1) = loop2; y2(2) = Im2; while(x2(end,j) > delta) for i =
1:rn
    switch i
        case 1
            alphai = y2(2)*dot(r(2:end).*N(2:end),1-x2(2:end,j));
        case rn
            alphai = y2(2)*dot(r(1:end-1).*N(1:end-1),1-x2(1:end-1,j));
        otherwise
            alphai = y2(2)*dot(r.*N,1-x2(:,j))-Im2*r(i)*N(i)*(1-x2(i,j));
    end
            x2(i,j+1)=x2(i,j).*exp(-S(i)*alphai);
end j=j+1; end
  %*******************************
j if j==tOut
    x2(end,tOut)
     if x2(end,tOut) < delta
        y2(1)=0;
     end
  y2(2)=y2(2) - inc/10;
elseif j>tOut
  y2(2) = y2(2)+.1
  inc=inc/10
  disp('oops')
else
  y2(2)=y2(2)- inc
end
%*******************************
g2 = x2; y(3)=inc;
```

```
% ******** End of the Function *****************************
```

C   Berkeley Madonna Code

```
METHOD RK4

STARTTIME = 0 STOPTIME=50 DT = 1
;------------------------------------------------------------------------
1st Group Non Users
------------------------------------------------------------------------
d/dt(S1) = At1 - Bt1 - mu*S1 d/dt(I1) = Bt1 - mu*I1
;------------------------------------------------------------------------
Users
------------------------------------------------------------------------
d/dt(X1) = Lambda1 - V1 - mu*X1 d/dt(Y1) = V1 - mu*Y1
;------------------------------------------------------------------------
2nd Group Non-Users
------------------------------------------------------------------------
d/dt(S2) = At2 - Bt2 - mu*S2 d/dt(I2) = Bt2 - mu*I2
;------------------------------------------------------------------------
Users
------------------------------------------------------------------------
d/dt(X2) = Lambda2 - V2 - mu*X2 d/dt(Y2) = V2 - mu*Y2
;------------------------------------------------------------------------
Conditions---------------------------------------------------------------
init S1 = 2*10^6 init I1 = 3 init X1 = 3*10^6 init Y1 = 3 init S2 =
10^6 init I2 = 3 init X2 = 3*10^6 init Y2 = 3
;------------------------------------------------------------------------
Parameters
------------------------------------------------------------------------
mu = 2 a1 = 2/3 a2 = 1/6 rho1 =8 rho2 = 20 sigma1 =52 sigma2 = 40 T1
= 25.204*10^6 T2 = 5.8286*10^6 Q1 = 31.4352*10^6 Q2 = 7.26939*10^6
beta1 = 10 ;beta2 = 2 beta2 = 5 b1 = 2 b2 = 1/2 Lambda1 = 3.7*10^6
Lambda2 = .87429*10^6 At2 = 12.22371*10^6 At1 = 52.85928*10^6 ;
-------------------------------------------------- Force of
infection for the two groups
```

;%%%%%%%%%%%%%%%%%%%   Force of the infections for USERS %%%%%%%%%%%%%%%%%%%%%%%%%%%

```
V1 = beta1*b1*X1*(barPa1*(I1/(T1*tau1+Q1)) +
barPb1*(Y1*tau1/(T1*tau1+Q1)) + barPb11*Y1/T1 + barPb12*Y2/T2)
barPa1 = (a1*Q1/(a1*Q1+b1*tau1*T1))*tau1 barPb1 =
(b1*tau1*T1/(a1*Q1+b1*tau1*T1))*tau1 barPb11 = (b1*w1*T1/( b1*w1*T1
+ b2*w2*T2 ))*w1 barPb12 = (b2*w2*T2/( b1*w1*T1 + b2*w2*T2 ))*w1 w1
= sigma1/( rho1 + sigma1)

V2 = beta2*b2*X2*(barPa2*(I2/(T2*tau2+Q2)) +
barPb2*(Y2*tau2/(T2*tau2+Q2)) + barPb21*Y1/T1 + barPb22*Y2/T2)
barPa2 = (a2*Q2/(a2*Q2+b2*tau2*T2))*tau2 barPb2 =
(b2*tau2*T2/(a2*Q2+b2*tau2*T2))*tau2 barPb21 = (b1*w1*T1/( b1*w1*T1
+ b2*w2*T2 ))*w2 barPb22 = (b2*w2*T2/( b1*w1*T1 + b2*w2*T2 ))*w2 w2
= sigma2/( rho2 + sigma2)
```

;%%%%%%%%%%%%%%%%%%%   Force of the infection for Non-Users %%%%%%%%%%%%%%%%%%%%%%%%%

```
Bt1 =
beta1*a1*S1*(tildePa1*(I1/(T1*tau1+Q1))+tildePb1*(Y1*tau1/(T1*tau1+Q1)))
tildePa1 = a1*Q1/(a1*Q1+b1*tau1*T1) tildePb1 =
b1*tau1*T1/(a1*Q1+b1*tau1*T1) tau1 = rho1/(sigma1+rho1)


Bt2 =
beta2*a2*S2*(tildePa2*(I2/(T2*tau2+Q2))+tildePb2*(Y2*tau2/(T2*tau2+Q2)))
tildePa2 = a2*Q2/(a2*Q2+b2*tau2*T2) tildePb2 =
b2*tau2*T2/(a2*Q2+b2*tau2*T2) tau2 = rho2/(sigma2+rho2)

METHOD RK4

STARTTIME = 0 STOPTIME=1 DT = 0.02 ;    R0 < 1 -----------------
;------------------------------ Parameters ---------------------
mu = 10 beta1 = 1 beta2 = 1 b1 = 1 b2 = 1 w1 = .99 w2 = abs(w1^2 -
w1) barT1 = 100 barT2 = 100 tau1 = abs(1 - w1) tau2 =abs(1 - w2) N
=b1*w1*barT1+b2*w2*barT2
;--------------------------------------------------------------
```

```
;V1t = beta1*b1*X1*(Y1*tau1/barT1 + (b1*w1*Y1*w1 + b2*w2*Y2*w1)/N)
;V2t = beta2*b2*X2*(Y2*tau2/barT2 + (b1*w1*Y1*w2 + b2*w2*Y2*w2)/N)

d/dt(X1) = mu*barT1 - beta1*b1*X1*(Y1*tau1/barT1 + (b1*w1*Y1*w1 +
b2*w2*Y2*w1)/N)  - mu*X1 init X1 = 3*10

d/dt(Y1) = beta1*b1*X1*(Y1*tau1/barT1 + (b1*w1*Y1*w1 +
b2*w2*Y2*w1)/N) - mu*Y1 init Y1 = 3

d/dt(X2) = mu*barT2 - beta2*b2*X2*(Y2*tau2/barT2 + (b1*w1*Y1*w2 +
b2*w2*Y2*w2)/N)  - mu*X2 init X2 = 3*10 d/dt(Y2) =
beta2*b2*X2*(Y2*tau2/barT2 + (b1*w1*Y1*w2 + b2*w2*Y2*w2)/N) - mu*Y2
init Y2 = 3
```

References

[1] Andersson, H. and Britton, T.(2000). *Stochastic Epidemic Models and Their Statistical Analysis,* Lecture Notes in Statistics 151. Springer-Verlag, New York.

[2] Banks,H., and Castillo-Chávez, C. (2003). *Bioterrorism in Mathematical Modeling Applications in Homeland Security.* Society for Industrial and Applied Mathematics, Philadelphia, 2003.

[3] Brauer, F., and Castillo-Chávez, C. (2001). *Mathematical Models in Population Biology and Epidemiology.* Springer-Verlag, New York.

[4] Castillo-Chávez, C., Blower, S., Kirschner, D, van den Driessche, P., and Yakubu, A-A.(Editors)(2002). *Mathematical Approaches for Emerging and Re-emerging Infectious Diseases, An Introductions*, Vol. 125, Vol. 126, IMA Series in Mathematics and Its Applications. Springer-Verlag, New York.

[5] Castillo-Chávez, C., Feng, Z., and Huang, W. (2002), On the Compuation of $\mathcal{R}_0$ and Its Role on Global Stability. *Mathematical Approaches for Emerging and Re-emerging Infectious Diseases, An Introductions*, Vol. 125, IMA Series in Mathematics and Its Applications, pp. 229-250.

[6] Castillo-Chávez, C., Song, B., and Zhang. (2003). An epidemic Model with Virtual Mass-Transpotation: The case of Small-Pox in a Large City. *Bioterrorism Mathematical Modeling Applications in Homeland Security*, (Eds.), Banks, H., and

Castillo-Chávez,C., SIAM Frotiers in Applied Mathematics. Springer-Verlag, New York.

[7] Diekmann, O., and Heesterbeek(2000). *Mathematical Epidemiology of Infectious Diseases.* Wiley, New York.

[8] Feige, U., Peleg,D., Raghavan, P., and Upfal, E.(1990). Randomized Broadcast in Networks. *Lecture Notes In Computer Science*,450,128-137.

[9] Gibson,D., Kleinberg,J., Raghavan,P. *Inferring Web Communities from Link Topology.* The 9th ACM Conference on Hypertext and Hypermedia,1998.

[10] Gonzalez,T.F.(2003). An efficient algorithm for gossiping in the multicasting communication environment *IEEE Transactions on Parallel and Distributed Systems*, 14, 7, 701-8

[11] Gupta,I.,Kermarrec,A.M., and Ganesh,A.J.(2002). Efficient epidemic-style protocols for reliable and scalable multicast. *Proceedings 21st IEEE Symposium on Reliable Distributed Systems*,180-9

[12] Jenkins, K., Hopkinson, K., and Birman K.(2001). Reliable Group Communication with Subgroups. *IEEE International Workshop on Applied Reliable Group Communication within the International Conference on Distributed Computing Systems 16-19 of April 2001, Pheonix, USA.*

[13] Karp,R.,Schindelhauer,C.,Shenker,S., and Vocking,-B.(2000). Randomized rumor spreading. *Proceedings 41st Annual Symposium on Foundations of Computer Science*,565-74.

[14] Kempe,D.,Dobra,A.,Gehrke,J.(2003). Gossip-based computation of aggregate information *Proceedings 44th IEEE Symposium on Foundations of Computer Science FOCS 2003*, 482-91.

[15] Murray, J.(2002) and (2003). *Mathematical Biology I: An Introduction and II: Spatial Models and Biomedical Applications.* Springer-Verlag, New York.